Essays in Entrepreneurship and Financial Economics

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:41128477">http://nrs.harvard.edu/urn-3:HUL.InstRepos:41128477</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Essays in Entrepreneurship and Financial Economics

A dissertation presented

by

Cheng Luo

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University
Cambridge, Massachusetts
April 2018
Abstract

The chapters in this dissertation study entrepreneurship activity and capital market behavior. In Chapter 1, I ask whether the opportunity cost of marriage affects female entrepreneurship. I use World War II casualties as exogenous shocks to local marriage markets across the US and test whether women in high-casualty regions were more active in starting new businesses than women in low-casualty regions. In Chapter 2, I examine hedge funds’ strategic behaviors at investment conferences. I evaluate performances of their stock pitches through event studies and analyze the behaviors and motives of various types of investors. In Chapter 3, my coauthors and I compare two asset pricing tests, the Fama-MacBeth cross-section test versus the Jensen’s alpha time-series test. We study their relevance to a risk-averse investor facing transaction costs as well as their statistical power of detecting anomalies in capital markets.
Contents

Abstract ................................................................. iii
Acknowledgments ....................................................... x

Introduction ............................................................ 1

1 The Other Gender Gap: Female Entrepreneurship after World War II 3
  1.1 Introduction ......................................................... 3
  1.2 Related Literature .................................................. 9
  1.3 US Mobilization for World War II .................................. 12
  1.4 Data ................................................................. 15
  1.5 Empirical Methodology .......................................... 16
    1.5.1 Baseline Estimation ........................................ 18
    1.5.2 Instrument for Marriage Market .......................... 19
  1.6 Empirical Evidence on War Impact .............................. 26
    1.6.1 Impact on Sex Ratio ........................................ 27
    1.6.2 Impact on Entrepreneurship ............................... 29
  1.7 Mechanism ........................................................ 32
    1.7.1 Marriage Market ............................................. 32
    1.7.2 Substitution Effect .......................................... 36
  1.8 Long-Term Effects .............................................. 37
  1.9 Conclusion ....................................................... 38

2 Talking Your Book: Evidence from Stock Pitches at Investment Conferences 41
  2.1 Introduction ....................................................... 41
  2.2 Related Literature ............................................... 45
  2.3 Data ............................................................... 47
  2.4 Stock Pitch Performances ....................................... 52
    2.4.1 Event Study Methodology ................................. 52
    2.4.2 Short-Term Market Reaction ............................ 53
    2.4.3 Long-Term Risk-Adjusted Returns ..................... 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5 Investor Behaviors</td>
<td>63</td>
</tr>
<tr>
<td>2.5.1 Holdings by Funds That Pitched Stocks</td>
<td>64</td>
</tr>
<tr>
<td>2.5.2 Why Funds Pitch Stocks</td>
<td>65</td>
</tr>
<tr>
<td>2.5.3 Holdings by Other Investors</td>
<td>71</td>
</tr>
<tr>
<td>2.6 Conclusion</td>
<td>75</td>
</tr>
<tr>
<td>3 Detecting Anomalies: The Relevance and Power of Asset Pricing Tests</td>
<td>77</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>77</td>
</tr>
<tr>
<td>3.2 The Relevance of Standard Asset Pricing Tests</td>
<td>81</td>
</tr>
<tr>
<td>3.2.1 The Return Generating Process: Scores and Factor Returns</td>
<td>83</td>
</tr>
<tr>
<td>3.2.1.1 Exposures to Raw and Unit Factor Portfolios</td>
<td>85</td>
</tr>
<tr>
<td>3.2.1.2 Computing Portfolio Expected Return and Variance</td>
<td>86</td>
</tr>
<tr>
<td>3.2.2 The Investor’s Security Selection Problem</td>
<td>87</td>
</tr>
<tr>
<td>3.2.3 Optimal Weights</td>
<td>88</td>
</tr>
<tr>
<td>3.2.3.1 Risk Neutral, Constant Transaction Costs</td>
<td>89</td>
</tr>
<tr>
<td>3.2.3.2 Risk Averse, No Transaction Costs</td>
<td>92</td>
</tr>
<tr>
<td>3.2.3.3 Risk Averse, Constant Transaction Costs</td>
<td>96</td>
</tr>
<tr>
<td>3.2.4 Summary</td>
<td>98</td>
</tr>
<tr>
<td>3.3 The Power of Standard Asset Pricing Tests</td>
<td>99</td>
</tr>
<tr>
<td>3.3.1 Asymptotic Power Curves</td>
<td>99</td>
</tr>
<tr>
<td>3.3.1.1 Data Mining</td>
<td>100</td>
</tr>
<tr>
<td>3.3.1.2 The Cross Section Test</td>
<td>101</td>
</tr>
<tr>
<td>3.3.1.3 The Time Series Test</td>
<td>104</td>
</tr>
<tr>
<td>3.3.1.4 A Comparison of the Cross Section and Time Series Tests</td>
<td>107</td>
</tr>
<tr>
<td>3.3.1.5 The Sharpe Ratio in US Data</td>
<td>108</td>
</tr>
<tr>
<td>3.3.1.6 Choosing Return Horizon</td>
<td>109</td>
</tr>
<tr>
<td>3.3.2 Small Sample Power Curves</td>
<td>116</td>
</tr>
<tr>
<td>3.4 Conclusion</td>
<td>122</td>
</tr>
<tr>
<td>References</td>
<td>123</td>
</tr>
<tr>
<td>Appendix A Appendix to Chapter 1</td>
<td>133</td>
</tr>
<tr>
<td>A.1 Supplementary Tables and Figures</td>
<td>133</td>
</tr>
<tr>
<td>Appendix B Appendix to Chapter 2</td>
<td>135</td>
</tr>
<tr>
<td>B.1 Description of the Sohn Investment Conference</td>
<td>135</td>
</tr>
<tr>
<td>B.2 Supplementary Tables and Figures</td>
<td>137</td>
</tr>
</tbody>
</table>
## Appendix C  Appendix to Chapter 3

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C.1</strong></td>
<td>Extensions to the Basic Model</td>
<td>141</td>
</tr>
<tr>
<td><strong>C.1.1</strong></td>
<td>Variable Transaction Costs</td>
<td>141</td>
</tr>
<tr>
<td><strong>C.1.2</strong></td>
<td>Capacity Constraints</td>
<td>143</td>
</tr>
<tr>
<td><strong>C.1.3</strong></td>
<td>Dynamic Trading</td>
<td>145</td>
</tr>
<tr>
<td><strong>C.2</strong></td>
<td>Derivations</td>
<td>148</td>
</tr>
<tr>
<td><strong>C.2.1</strong></td>
<td>Derivation of Equation 3.11</td>
<td>148</td>
</tr>
<tr>
<td><strong>C.2.2</strong></td>
<td>Solution for Equation C.2 with a Budget Constraint</td>
<td>148</td>
</tr>
<tr>
<td><strong>C.2.3</strong></td>
<td>Derivation of $dC(A)/dA&lt;0$ in Equation C.3</td>
<td>149</td>
</tr>
<tr>
<td><strong>C.2.4</strong></td>
<td>Derivation of Equation 3.16 for the Cross Section Test</td>
<td>151</td>
</tr>
<tr>
<td><strong>C.2.5</strong></td>
<td>GMM Derivation of Equation 3.17 for the Time Series Test</td>
<td>152</td>
</tr>
<tr>
<td><strong>C.2.6</strong></td>
<td>Using an Efficient Factor Portfolio in Equation 3.17</td>
<td>156</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Pre-War County Characteristics by Casualty Rate ........................................ 17
1.2 Determinants of Enlistment Rate ................................................................. 22
1.3 Determinants of Casualty Rate ......................................................................... 24
1.4 Public Health Condition and Casualty Rate .................................................... 26
1.5 War Impact on Sex Ratio .................................................................................. 28
1.6 War Impact on Entrepreneurship .................................................................... 31
1.7 Cohort Table ..................................................................................................... 33
1.8 Impact on Marital Outcomes ............................................................................ 34
1.9 Mechanism: Marriage Market and Gender Substitution .................................. 35

2.1 Number of Investment Conferences and Stock Pitches .................................. 48
2.2 Characteristics of Funds that Pitched and Pitched Stocks ................................ 50
2.3 Shareholder Composition of Pitched Stocks ..................................................... 51
2.4 Market-Adjusted Returns of Pitched Stocks .................................................... 59
2.5 Risk-Adjusted Returns of Pitched Stocks ......................................................... 60
2.6 Factor Tilts of Pitched and Non-Pitched Stocks ................................................. 61
2.7 Jensen’s Alpha Test for Calendar-Time Portfolios ............................................ 63
2.8 Determinants of Stock Pitch Decisions ............................................................. 68
2.9 Pitch Performances and Fund Raising Activities .............................................. 70
2.10 Post-Pitch Returns of Pitched Stocks and Newly-Bought Stocks .................... 72

3.1 Empirical Moments for Factor Returns (1963:07 to 2016:03) ......................... 110
3.2 Time Series Tests: Monthly, Weekly, Daily ..................................................... 111
3.3 Power Difference between the CS and TS Test ................................................. 121

B.1 Conference List .............................................................................................. 137
B.2 Selected List of Funds that Pitched at Investment Conferences ....................... 138
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Female Share of Entrepreneurship and Labor Force</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Mechanism</td>
<td>7</td>
</tr>
<tr>
<td>1.3</td>
<td>Military Enlistment</td>
<td>13</td>
</tr>
<tr>
<td>1.4</td>
<td>Sex Ratio by Birth Cohort</td>
<td>14</td>
</tr>
<tr>
<td>1.5</td>
<td>Distribution of County-level Casualty Rate</td>
<td>20</td>
</tr>
<tr>
<td>1.6</td>
<td>State-level Enlistment Rate and Casualty Rate</td>
<td>25</td>
</tr>
<tr>
<td>1.7</td>
<td>War Impact on Female Share of Entrepreneurship</td>
<td>38</td>
</tr>
<tr>
<td>2.1</td>
<td>Short-Term Market Reaction to Pitched Stocks</td>
<td>54</td>
</tr>
<tr>
<td>2.2</td>
<td>Trading Volumes of Pitched Stocks</td>
<td>55</td>
</tr>
<tr>
<td>2.3</td>
<td>Market-Adjusted Returns of Pitched Stocks</td>
<td>56</td>
</tr>
<tr>
<td>2.4</td>
<td>Risk-Adjusted Returns of Pitched Stocks</td>
<td>57</td>
</tr>
<tr>
<td>2.5</td>
<td>Calendar-Time Portfolios</td>
<td>62</td>
</tr>
<tr>
<td>2.6</td>
<td>Portfolio Weights of Pitched Stocks</td>
<td>65</td>
</tr>
<tr>
<td>2.7</td>
<td>Holdings of Pitched Stocks using Alternative Measures</td>
<td>66</td>
</tr>
<tr>
<td>2.8</td>
<td>Trading Behaviors of Active Institutional Investors</td>
<td>74</td>
</tr>
<tr>
<td>2.9</td>
<td>Trading Behaviors of Retail Investors and Corporate Insiders</td>
<td>75</td>
</tr>
<tr>
<td>3.1</td>
<td>Asymptotic Power Curve of The Cross Section Test</td>
<td>103</td>
</tr>
<tr>
<td>3.2</td>
<td>Asymptotic Power Curve of The Time Series Test</td>
<td>106</td>
</tr>
<tr>
<td>3.3</td>
<td>Asymptotic Power Curve of the CS and TS Test at Various Sharpe Ratios of the Incumbent Factor Model</td>
<td>108</td>
</tr>
<tr>
<td>3.4</td>
<td>Monthly, Weekly, and Daily TS Test Statistics</td>
<td>113</td>
</tr>
<tr>
<td>3.5</td>
<td>Bias and Power, Annualized Alphas By Return Horizon</td>
<td>115</td>
</tr>
<tr>
<td>3.6</td>
<td>Small Sample and Simulated Power Curve of the CS and TS Test</td>
<td>116</td>
</tr>
<tr>
<td>3.7</td>
<td>Small Sample Power Curve of the CS Test</td>
<td>118</td>
</tr>
<tr>
<td>3.8</td>
<td>Small Sample Power Curve of the TS Test</td>
<td>119</td>
</tr>
<tr>
<td>3.9</td>
<td>Small Sample Power Curve of the CS and TS Test at Various Sharpe Ratios of the Incumbent Factor Model</td>
<td>120</td>
</tr>
<tr>
<td>A.1</td>
<td>Cross-County Variation in Casualty Rate</td>
<td>133</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>A.2</td>
<td>Examples of Historical Documents</td>
<td>134</td>
</tr>
<tr>
<td>B.1</td>
<td>Registration Page of Sohn Conference</td>
<td>136</td>
</tr>
<tr>
<td>B.2</td>
<td>Distribution of Money Raised</td>
<td>140</td>
</tr>
</tbody>
</table>
Acknowledgments

I am immensely grateful to my advisers Lauren Cohen, Christopher Malloy and Josh Lerner for their invaluable guidance and support. Their creativity, dedication and wealth of knowledge has been a tremendous inspiration. I thank Malcolm Baker, John Campbell, Claudia Goldin, Robin Greenwood, William Kerr, Scott Kominers, Ramana Nanda, Tom Nicholas, Andrei Shleifer, Christopher Stanton, Jeremy Stein and Luis Viceira for detailed comments and helpful suggestions. I thank William Skimmyhorn and John Vigna at West Point for sharing their deep knowledge on the history of World War II. I have had the great opportunity to work with Malcolm Baker and Ryan Taliaferro on Chapter 3 in this thesis.

I am grateful to my close friends and my classmates in the economics program for many thoughtful discussions and encouragement, especially Gordon Liao, Yosub Jung, Huaizhi Chen, Ashvin Gandhi, Yizhou Jin, Andrea Passalacqua, Yang He, Sophie Wang and Weiling Liu.

Finally, I am above all grateful to my family. Without the love and support of my wife and parents, I would not have accomplished this.
To Sarah
Introduction

This thesis consists of three essays in entrepreneurship and financial economics. In Chapter 1 of this dissertation, “The Other Gender Gap: Female Entrepreneurship after World War II,” I exploit exogenous variation in the marriage market across the US caused by World War II casualties to provide causal evidence on how opportunity cost influences women’s entrepreneurship decisions. I show that marriage is an important form of opportunity cost hindering women from starting new businesses. World War II casualties affected the local marriage market for single women and access to capital for war widows. Using novel business registration and individual-level census data, I find that single women are more active in starting new businesses when they face worse prospects in the marriage market. As a result, US counties with heavier casualties had a higher female share of entrepreneurs. This difference persists to this day. Evidence in favor of the marriage-market channel suggests that reducing opportunity cost is more effective in encouraging female entrepreneurship than merely providing financial subsidies.

In Chapter 2, “Talking Your Book: Evidence from Stock Pitches at Investment Conferences,” I show that hedge funds take advantage of the publicity of these conferences to strategically release their book information to drive market demand using a novel dataset drawn from investment conferences from 2008 to 2013, Specifically, hedge funds sell pitched stocks after the conferences to take profit and create room for better investment opportunities. However, the pitched stocks still perform better than non-pitched stocks in the funds’ portfolios afterwards. Hedge funds do not pitch obviously bad stocks because maintaining a good reputation helps them raise money. Pitched stocks earn a cumulative abnormal return of 20% over 18 months
before the pitch and continue to outperform the benchmark by 7% over 9 months afterwards. Post-conference abnormal return reverts partially after another 9 months. Moreover, mutual funds exhibit opposite trading behaviors—selling before the pitches and buying afterwards—and may contribute to the post-pitch outperformance. Other hedge funds trade pitched stocks similarly to the funds that pitch them, suggesting that they either run correlated strategies or share information with each other.

In Chapter 3, “Detecting Anomalies: The Relevance and Power of Asset Pricing Tests,” (with Malcolm Baker and Ryan Taliaferro) we compare the two standard approaches for identifying capital market anomalies—cross-sectional coefficient tests, in the spirit of Fama and MacBeth (1973), and time-series intercept tests, in the spirit of Jensen (1968). A new signal can pass the first test, which we label a “score anomaly,” it can pass the second test as a “factor anomaly,” or it can pass both. We demonstrate the relevance of each to a mean-variance optimizing investor facing simple transaction costs that are constant across stocks. For a risk-neutral investor facing transaction costs, only score anomalies are relevant. For a risk-averse investor facing no transaction costs, only factor anomalies are relevant. In the more general case of risk aversion and transaction costs, both tests matter. In extensions, we derive modified versions of the basic tests that net out anomaly execution costs for situations where the investor faces capital constraints, a multi-period portfolio choice problem, or transaction costs that vary across stocks. Next, we measure the econometric power of the two tests. The relative power of time-series factor tests falls with the in-sample Sharpe ratio of the incumbent factor model, as in Shanken (1992). New factor anomalies can be successively harder to detect, leading to a lower natural limit on the number of anomalies. Meanwhile, for an investor facing transaction costs, where score anomalies are also applicable, there can be a higher natural limit on the number of anomalies that can be statistically validated as relevant.
Chapter 1

The Other Gender Gap: Female Entrepreneurship after World War II

1.1 Introduction

Over the last century, the gender gap in employment, education, and many other domains has narrowed considerably, thanks to more liberal social norms and legislation (Goldin, 1983, 1986, 1992). But there is still a significant gender gap in entrepreneurship. Women are half as active as men in starting new businesses and make up a smaller share of business owners. As of 2011, women represented 35 percent of total entrepreneurial activity and women-owned businesses were one-third as likely to grow to have more than $1 million revenue as men-owned businesses (Mitchell, 2011; Fairlie, Morelix, Reedy, and Russell, 2015). The gap is even more severe in the high-tech industry. Not only employment in the Silicon Valley is dominated by male but also the startup and venture capital space has a huge male bias. This is economically important because a significant share of the wealth in the society is created through entrepreneurship rather than through employment. The gap in business formation directly contributes to the gap in business leadership.
This figure shows female share of new businesses founded (green line) and female share of labor force (red line) in the US. The shaded period indicates World War II. Source: author’s calculation using business registration records and data from the Bureau of Labor Statistics.

Both the government and the private sector have taken numerous attempts to reduce this gender gap in entrepreneurship. While some polices aim to improve financial incentives for women-owned businesses such as targeted lending programs and tax deductions, many firms focus on making the workplace more family-friendly such as better maternity leave policies and childcare. The later kind of policies reduce the opportunity cost for women to pursue their professional and business careers. In this paper, I investigate one type of opportunity cost that hinders women from starting new businesses; namely, the marriage market. Specifically, I exploit variation in marriage market conditions across US counties to investigate whether prospects in the marriage market play a role in female entrepreneurship. I show that women in counties with a worse marriage market face a lower opportunity cost and are more likely to start new businesses.

On the supply side, whether a woman becomes an entrepreneur depends on whether she is prepared to run a business and is given an opportunity. Women still face discrimination in
higher education, employment, and access to financing, which prevents them from acquiring the human and financial capital necessary for entrepreneurship. On the demand hand, whether a woman becomes an entrepreneur depends on whether she wants to. Women face higher an opportunity cost in marriage than men do due to childrearing and housekeeping. They might also have a lower risk appetite; specifically, for the risk of starting a new business. Also, stereotypes of women’s economic roles can discourage female entrepreneurship. Among all the aforementioned factors, the opportunity cost of marriage stands out because entrepreneurship demands a substantial time commitment and carries nondiversifiable risk. It has family-unfriendly incentives similar to employment in business and law, where the gender gap is still large (Goldin, 2014).

Because the marriage market can be affected by economic development and industry structure, its effect on entrepreneurship is endogenously related to certain county characteristics and is challenging to test empirically. For instance, regions with agriculture or heavy industry have more men per woman and women therefore face not only lower competition in the marriage market but also fewer business opportunities. To address this endogeneity concern, I use US casualties during World War II as plausibly exogenous shocks to local marriage markets and show that these shocks have a large and significant impact on female entrepreneurship. I begin by calculating local military casualty rates using the US decennial censuses and military records from the National Archives. About 2.3 percent of the young male population in the US were war casualties in World War II. However, the casualty rates for individual counties varied from 10 percent to close to zero for all counties and from 5.5 percent to close to zero for counties with at least 50,000 population in 1940. Casualties among young men caused imbalanced sex ratios[^1] and led to unfavorable marriage markets for women of certain birth cohorts. Because women tend to marry men older than themselves, this imbalance in the marriage market could persist over time. Consistent with the literature, I find that counties with more casualties tend to have more single women and lower birth rates (Abramitzky, Delavande, and Vasconcelos, 2011; Brainerd, 2016).

[^1]: Sex Ratio = 100*Male Population/Female Population
However, due to draft procedures, the variations in casualty rate across counties may not be entirely independent of county characteristics and might therefore violate the exclusion restriction (Goldin [1991], Acemoglu, Autor, and Lyle [2004], Goldin and Olivetti [2013]). To establish conditional instrumental variable (IV) independence, I review historical military documents and identify covariates that can cause regional differences in enlistments and casualties. For instance, to ensure food supplies during the war, the government granted deferments to farmers. Therefore, counties with a higher share of agricultural population had lower enlistment rates. After controlling for these covariates, the residual variations in casualty rate across counties can be attributable to “luck” on the battlefield and are plausibly exogenous to county characteristics.

Having established the conditional exogeneity of imbalanced sex ratios, I examine their effect on female entrepreneurship after World War II using an IV strategy. To measure entrepreneurship, I exploit novel business registration data collected through public records made available by state governments. I measure a county’s female entrepreneurship participation rate ($EPR_{Female}$) as the number of new businesses started by women divided by the size of the female labor force in that county. The key feature of the business registration data that makes my analysis possible is that for each registered company, the registry provides the owner’s name and address, allowing me to infer gender from first name and to link the owner to a particular region.

The identification faces a major challenge. High- and low-casualty counties might differ in unobserved dimensions that cause different female entrepreneurship rates. Although I cannot completely rule out this alternative interpretation, I provide evidence suggesting that they are not the primary sources of my findings. My results are robust to including a variety of control variables on county characteristics. Moreover, female entrepreneurship rate and the sex ratio did not vary systematically between high- and low-casualty counties prior to the war. This suggests that these counties were initially broadly comparable with each other.

Using the military death data and the business registration data, I ask whether women in regions with more imbalanced sex ratios start more new businesses due to a lower opportunity
cost of marriage. I find evidence consistent with the conjectured negative impact of opportunity cost on female entrepreneurship. My baseline estimates suggest that for a decrease of 10 men per 100 women in the adult population, women are 19 percent more active in starting new businesses and the female share of new businesses increases by 14 percent. I further show that the impact of war casualties on female entrepreneurship was persistent over time.

There could be multiple possible channels through which World War II casualties affected female entrepreneurship, as illustrated in Figure 1.2. First is easing of financial constraints. War widows received a life insurance payout and death gratitude from the government equivalent to about 100 thousand dollars today. This could serve as important startup capital, especially for those without dependents, at a time when women still could not borrow from banks easily. Second is lower opportunity cost of marriage. With marriage market conditions worsening, single women faced a lower opportunity cost of staying single and starting their own businesses. Third is less competition from men. With fewer men starting new businesses, women had more business opportunities and faced less competition.

This diagram demonstrates the possible mechanisms through which World War II casualties could impact female entrepreneurship.

To shed further light on the mechanism behind my findings, I next test the prediction that war casualties had a more pronounced effect on entrepreneurship for single than for married women because the marriage market only mattered for single women. To do so, I merge the business registration data with the individual-level 1940 census from IPUMS (Integrated...
Public Use Microdata Series) and obtain the marital status of women entrepreneurs. I do find that the marriage market has a strong and significant impact on entrepreneurship for single women but a smaller and less-significant impact for married women. This result provides further evidence against the access-to-capital channel for war widows.

Finally, I show that the effect of the opportunity cost of marriage on female entrepreneurship is not solely caused by the mechanical substitution between men and women. When men did not come back home from the war, women would take up the vacant business opportunities. To provide evidence against the mechanical substitution channel, I estimate the war impact on female entrepreneurship in different business sectors. I classify companies into business sectors based on the company name in the business registration record and divide the sectors into female-dominated, gender-neutral, and male-dominated. If the postwar increase in female entrepreneurship is purely due to substitution, it is the gender-neutral sectors that would have the most new female entrepreneurs. In fact, I find that women started more businesses in the female-dominated sectors.

Numerous papers have investigated the gender gap in the labor market and the impact of war on female labor participation. However, to the best of my knowledge, this is the first to use business registration data to study historical female entrepreneurship in the US and to exploit the cross-county variation caused by World War II in a quasi-random experiment. I document that the gender gap in entrepreneurship was even more severe during the early twentieth century. The share of new companies founded by women was consistently between 10 and 15 percent and did not start to rise until the 1970s, when Congress passed major legislation outlawing gender discrimination. Using business registration data not only enables me to trace female entrepreneurship to the beginning of the twentieth century but also allow me to focus on companies with higher growth potentials. Registering a company provides many practical benefits such as judicial protection and ability to use formal labor contracts and is crucial to company growth.\footnote{The Equal Pay Act in 1963, the Civil Rights Act in 1964, the Fair Housing Act in 1968, the Education Amendments in 1972, and the Women’s Educational Equity Act and the Equal Credit Opportunity Act in 1974.}

\footnote{Klapper, Amit, and Guillen (2010) Guzman and Stern}
The remainder of the paper is organized as follows. Section 2 relates my paper to current literature on the gender gap in entrepreneurship and on the social impacts of war. Section 3 provides historical context on US engagement in World War II. Section 4 describes the data. Section 5 lays out the empirical strategy and explores World War II casualty rates as an instrumental variable. Section 6 establishes a causal link between the opportunity cost of marriage and female entrepreneurship and examines long-term effects. Section 7 investigates the mechanisms by which the war affected female entrepreneurship. Section 8 concludes.

1.2 Related Literature

This paper contributes to three main strands of literature. First, it contributes to the literature investigating the gender gap in entrepreneurship. Numerous papers have compared male and female entrepreneurs and have documented that men are more likely than women to pursue entrepreneurial activity, whether out of necessity or to pursue an opportunity (Fischer, Reuber, and Dyke 1993; Ayers and Siegelman 1995; Neumark, Bank, and Van Nort 1996; Ardagna and Lusardi 2008; Asoni 2010; Gompers and Wang 2017a). Buttner and Rosen (1989), Tinkler, Ku, Whittington, and Davies (2015) find that it is harder for women than for men to acquire either a bank loan or VC funding. There is also a large literature documenting the gender gap in leadership positions (Langowitz and Minniti 2007; Beaman, Chattopadhyay, Duflo, Pande, and Topalova, 2009; Bruhn, 2009; Minniti 2009; Field, Jayachandran, and Pande 2010; Minniti 2010; Minniti and Naudé 2010; Estrin and Mickiewicz 2011; Hausmann, Tyson, and Zahidi 2011; Pande and Ford 2011; Beaman, Duflo, Pande, and Topalova 2012; Duflo 2012; Ghani, Kerr, and O’Connell 2013). Other papers have examined the factors influencing women’s entrepreneurship decision and its transmission mechanism. Ghani, Kerr, and O’Connell (2014) find that political liberation encourages women to start new businesses. Macpherson (1988) examines entrepreneurship decisions by married women. Kacperczyk (2013) finds that women are more likely to become “intrapreneurs” than men. Moore (1983) examines the impact of gender discrimination in employment on entrepreneurship. Finally, several papers have
examined the benefits of gender diversity in entrepreneurship; for instance, Gompers and Wang (2017b) find that gender diversity improves venture capital funds’ performance.

More broadly, this paper contributes to the literature on women’s career choices. There is an extensive literature on the opportunity cost of marriage—specifically due to childrearing and housework—on a woman’s career (Mincer 1962; Hill 1979; Becker 1985; Bielby and Bielby 1988; Blau and Robins 1988; Joshi 1991; Korenman and Neumark 1992; Angrist and Evans 1998; Waldfogel 1998; Adda, Dustmann, and Stevens 2017; Zhang 2017a,b). Goldin (2014) shows that time commitment explains gender inequality in careers such as business and law. Goldin and Katz (2002) also find that birth control allows women to invest more in their human capital. My results provide additional evidence on the opportunity cost of marriage and are broadly consistent with the literature.

The second main strand of literature my paper contributes to is innovation and entrepreneurship. Numerous papers focus on the impact of access to capital on entrepreneurship decisions and outcomes. Personal wealth affects the propensity to engage in entrepreneurship (Hurst and Lusardi 2004a,b; Nanda 2010); receiving an inheritance or gift has a positive effect on that decision (Blanchflower and Oswald 1998; Cagetti and De Nardi 2006). Some find that liquidity constraints exert a noticeable influence on the viability of an entrepreneurial business (Evans and Jovanovic 1989; Holtz-Eakin, Joulfaian, and Rosen 1994), yet Astebro and Bernhardt (2003) find a negative relationship between bank loans and business survival. Hall and Woodward (2010) and Kerr, Nanda, and Rhodes-Kropf (2014) examine idiosyncratic risk in entrepreneurship. Others study the transmission of entrepreneurship and its long-term effect on economic growth (Knack and Keefer 1997; Audretsch, Keilbach, and Lehmann 2006; Iyer and Schoar 2010; Decker, Haltiwanger, Jarmin, and Miranda, 2014).

In particular, this paper adds to the growing literature that uses business registration data to study entrepreneurship with a focus on quality. Klapper et al. (2010) measures rate of business registration across countries and time. Hurst and Pugsley (2011) use business registration data to study what small businesses do. Recent papers such as Schoar (2010) and Levine and Rubinstein (2017) question the appropriate scope and quality of entrepreneurs.
Evans and Leighton (1990) and Poschke (2013) compare companies formed out of necessity by unemployed people to those formed by people with jobs to seize an opportunity. Guzman and Stern (2014, 2015) use business registration data to analyze both the quantity and quality of new companies. My own paper further shows that business registration data can be a useful measure of entrepreneurship.

Although there are many papers on the social impact of war, especially its impact on the marriage market and on minority communities, this is, to the best of my knowledge, the first to study the impact of World War II on female entrepreneurship in the US. War casualties create an imbalanced sex ratio and numerous papers have investigated its impact on women. Abramitzky et al. (2011) find that men in regions of France that had heavier casualties during World War I were more likely to marry a woman of a higher social class. Bethmann and Kvasnicka (2012) find that a lower sex ratio led to higher out-of-wedlock fertility rate in Germany after World War II. Brainerd (2016) finds similar results in Russia after World War II and further documents that women in regions with heavier casualties had less bargaining power within marriage. There is also a large literature on the effect of an imbalanced sex ratio on the marriage market in general that makes similar findings. Becker (1973, 1974, 1981) models the impact of sex ratio imbalance and makes theoretical predictions. Angrist (2002), Lafortune (2013) study the effect of an imbalanced sex ratio among immigrants in the US on the marriage market and labor market. Francis (2011), Chang and Zhang (2012), Grosjean and Khattar (2015) find similar effects among immigrants in Australia and Taiwan. Charles and Luoh (2010), Mechoulam (2011) examine the impact of high male incarceration rates on women’s martial outcomes.

Many papers have examined the social impact of war beyond the marriage market. Goldin (1991), Acemoglu et al. (2004), Goldin and Olivetti (2013) study the impact of America’s massive World War II mobilization on female labor force participation and labor market equilibriums. Gay and Boehmke (2017) examine the impact of World War I casualties on female labor force participation in France.

Researchers have also examined the long-term social and economic impact of historical
events and conditions, including wars. Cook-Stuntz (2015) documents the role model effect of Rosie the Riveter on daughters’ labor force decisions. Campante and Yanagizawa-Drott (2015) find that one generation’s military service affects the next generation’s decisions to serve. Glaeser, Kerr, and Kerr (2015) find that a region’s industry structure in the past can have a long-lasting impact on subsequent entrepreneurship outcomes. Cohen, Malloy, and Nguyen (2016) find that the 1930s drought in the US had a long-lasting impact on union formation. Cohen, Gurun, and Malloy (2017) find that the internment of Japanese-Americans during World War II affects companies’ trade, M&A, and other strategic behaviors to this day.

1.3 US Mobilization for World War II

The US government passed the Selective Service and Training Act in 1940, after the outbreak of World War II, and established a mandatory national draft registration. From 1940 until 1947—when the Act expired—16 million Americans were inducted and served in the war. The conscription almost exhausted the supply of men in certain birth cohorts and included 50 percent of men between age 20 and 49 in 1940. The Act initially registered all men from age 21 to 35. It later expanded to age 18 to 65, although only those between 18 to 45 were liable for military service. After each registration, lotteries determined the order of enlistment and local draft boards determined one’s eligibility for active duty. Volunteers were allowed only during the first year of the war. As a result, 67 percent of all those who served had been drafted.

Various historical factors influenced one’s likelihood of serving. First, there were deferred exemption categories in the conscription and local draft boards had considerable discretion in granting exemptions. Exemption categories included those with physical disabilities, workers in important war industries, men who were the primary support for children, and conscientious objectors. To ensure a sufficient food supply, the government also granted deferments to farmers. Furthermore, race affected one’s likelihood of being drafted. The military was still racially segregated in the 1940s and in order to minimize the cost of constructing separate
buildings for African Americans, draft boards simply avoided them. Finally, people of German, Italian, and Japanese origin may also have been avoided in the draft due to concerns about their loyalty.

The majority of the enlistments and casualties occurred towards the end of the war—around 1944—when the deadliest battles were fought. Statistics from the enlistment records show that men from certain birth cohorts experienced a higher enlistment rate than others. As shown in Figure 1.3, the peak cohort—those born around 1920—experienced an enlistment rate of 80 percent. As a result, they also experienced a higher casualty rate and a larger shock to the sex ratio. About 417 thousand US soldiers died during the war. Mortalities had a more noticeable impact on the sex ratio in the cohort born between 1915 and 1925.

![Figure 1.3: Military Enlistment](image)

This figure shows number of US enlisted soldiers during World War II by year of birth. Source: *World War II Army Enlistment Records, ca. 1938-1946* from the National Archives.

Figure 1.4 plots sex ratio against age for three birth cohorts: 1906–1915, 1916–1925 and 1926–1935. The 1916–1925 cohort had the largest decrease in sex ratio from age 20 to 30. The two-percent difference in sex ratio after age 30 is broadly in line with the overall casualty rate. The 1906–1915 cohort did not experience any war before age 30; they were too young for
World War I and too old for World War II. The older members of the 1926–1935 cohort served in World War II while the younger ones served in the Korean War, which was not as deadly. Thus, that cohort’s sex ratio did not decrease as much as that of the 1916–1925 cohort.

![Figure 1.4: Sex Ratio by Birth Cohort](image)

This figure plots sex ratio against age for three birth cohorts: people who were born during 1906–1915 (blue line), during 1916–1925 (red line) and during 1926–1935 (green line). Sex ratio is the ratio of male population over female population times 100. Most US soldiers in World War II belonged to the 1916–1925 cohort. As a result, among the three cohorts, the 1916–1925 cohort experienced the largest decrease in sex ratio from age 20 to 30. Source: the US decennial censuses.

As a result, World War II casualties had different implications for women from different birth cohorts. During the 1940s and 1950s, most women got married between age 20 and 25. Women who were born before 1920 were therefore likely to be already married in 1940 when the Selective Service Act was enacted. Women born after 1920 were likely to be single when the war began and would be looking for husbands after it ended in 1945. Because a woman was less likely at that time to marry a younger man, the localized shortage of men in the marriage market for those single women would not be replenished naturally and could have a long-lasting impact on their lives.
1.4 Data

To test my hypotheses regarding female entrepreneurship, I combine data from multiple historical sources:

**Military enlistment and casualties.** I hand-collect a novel dataset to measure county-level military enlistment and casualty rates, using records from the National Archives and Records Administration. The document titled *World War II Army Enlistment Records, ca. 1938–1946* contains individual-level information from enlistment cards[^3] for most of the US Army during World War II. The Department of War and the Department of the Navy compiled tabulations of casualties and lists of deceased soldiers, which are also accessible through the National Archives. The tabulations report for each county the number of casualties by cause. I download these tabulations and transcribe the scanned copies. Panel (b) in Figure A.2 shows an example. The lists of deceased soldiers provide individual-level detail[^4] I calculate a county’s military casualty rate as the ratio of the number of its deceased soldiers to the size of its young male population, defined as men between 20 and 40. I also calculate the casualty rate using other denominators, such as the total population; the main results do not change qualitatively.

**Entrepreneurship activity.** I construct a novel measure of county-level entrepreneurship by gender using business registration records. My dataset spans 1900 to 2015 and contains detailed registration information on individual businesses in 30 states[^5] Variables include business name, owner’s name, business type, address, and incorporation date. Although the earlier period is less well populated, the sample is still representative of the overall business patterns. In my analysis, I focus on businesses that were registered as either a limited liability[^3] In general, the enlistment records contain serial number, name, state and county of residence, place of enlistment, date of enlistment, and so on. See Panel (a) in Figure A.2 for an example card.

[^4]: In general, the casualty records contain name, rank, serial number, and type of casualty for the Army and the Army Air Force and name, rank, and next of kin for the Navy, Marine Corps, and Coast Guard. See Panel (c) in Figure A.2 for an example list.

[^5]: These states are Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Florida, Georgia, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Minnesota, Missouri, Nevada, New York, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Dakota, Tennessee, Texas, Vermont, Washington, Wisconsin, and Wyoming. The public data aggregator does not have data from other states’ commerce departments.
company (LLC) or a corporation (Corp). I infer the owner’s gender from the first name. I then calculate a county’s female (male) entrepreneurship participation rate as the ratio of the number of businesses registered by women (men) to the size of the female (male) labor force. I assign lower weights to unisex names and include them in my counts of both male and female entrepreneurs. I also calculate the percentage of new businesses that were registered by women.

In addition, I merge the business registration data with the individual-level 1940 census data using record linkage algorithms. The census contains detailed information for the 132 million individuals and 38 million households in 1940. Specifically, I obtain the marital status and date of birth of the entrepreneurs who registered a business during my sample period.

**County demographics.** I compile state- and county-level demographic statistics from the US decennial censuses from 1900 to 1990. I obtain population by age group, gender, education, employment status, and marriage status. Employment status was not reported before 1930 and in 1960. Education level was not reported before 1940 and in 1960. Per-capita income was only reported beginning in 1970. Definitions of age groups differ slightly across censuses; I expand the age boundaries so that the groups are consistent across time.

My analysis relies on a panel of US counties from 1910 to 1990. The sample includes counties that had at least 1,000 residents in 1940. Throughout the analysis, I exclude Alaska, Hawaii, Washington, DC, and Nevada from the analysis, consistent with Acemoglu et al. (2004). Alaska and Hawaii were still US territories in the 1940s and Nevada experienced a large population influx. Furthermore, I exclude counties whose geographic boundaries changed during the 1950s, that had a significant institutional population—such as prisons or barracks, or that had a majority of the population working in agriculture. Table 1.1 reports the summary statistics for the main variables used in the analysis.

### 1.5 Empirical Methodology

In this section, I describe the empirical strategy. I will first introduce the baseline estimating equation and describe the endogeneity issue. I will then formally introduce the instrumental
This table reports county characteristics before World War II. Column 1 reports mean values for all the US counties. Columns 2, 3 and 4 report the mean values for counties grouped by casualty rate, from low to high. *Casualty rate* is a county’s number of casualties normalized by its young male population in 1940. *Female share of entrepreneurship* is the percent of new businesses founded by women from 1930 to 1940. Other demographics and economics characteristics are from the 1940 census or the last decennial census when the data item is available. Standard deviations are in parentheses below the mean values.

<table>
<thead>
<tr>
<th></th>
<th>By Casualty Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (1)</td>
</tr>
<tr>
<td>Casualty rate (%)</td>
<td>2.32</td>
</tr>
<tr>
<td></td>
<td>(0.78)</td>
</tr>
<tr>
<td>Female share of entrepreneurship (%)</td>
<td>15.42</td>
</tr>
<tr>
<td></td>
<td>(24.56)</td>
</tr>
</tbody>
</table>

**Demographic Characteristics**

<table>
<thead>
<tr>
<th></th>
<th>By Casualty Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (1)</td>
</tr>
<tr>
<td>Population (‘000s)</td>
<td>53.05</td>
</tr>
<tr>
<td></td>
<td>(180.80)</td>
</tr>
<tr>
<td>Population density (per sq mile)</td>
<td>98.07</td>
</tr>
<tr>
<td></td>
<td>(319.00)</td>
</tr>
<tr>
<td>Percent young population (%)</td>
<td>18.51</td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
</tr>
<tr>
<td>Sex Ratio (per 100 women)</td>
<td>103.75</td>
</tr>
<tr>
<td></td>
<td>(6.43)</td>
</tr>
<tr>
<td>Percent rural population (%)</td>
<td>76.18</td>
</tr>
<tr>
<td></td>
<td>(25.29)</td>
</tr>
<tr>
<td>Percent white (%)</td>
<td>85.07</td>
</tr>
<tr>
<td></td>
<td>(21.26)</td>
</tr>
<tr>
<td>Percent college educated (%)</td>
<td>11.91</td>
</tr>
<tr>
<td></td>
<td>(4.65)</td>
</tr>
</tbody>
</table>

**Economic Characteristics**

<table>
<thead>
<tr>
<th></th>
<th>By Casualty Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (1)</td>
</tr>
<tr>
<td>Unemployment rate (%)</td>
<td>6.57</td>
</tr>
<tr>
<td></td>
<td>(3.19)</td>
</tr>
<tr>
<td>Percent in manufacturing (%)</td>
<td>4.34</td>
</tr>
<tr>
<td></td>
<td>(4.73)</td>
</tr>
<tr>
<td>Percent in agriculture (%)</td>
<td>10.35</td>
</tr>
<tr>
<td></td>
<td>(5.00)</td>
</tr>
</tbody>
</table>
variable for the marriage market; namely, World War II casualty rates. I discuss its potential correlation with unobserved county characteristics and explain why it is nevertheless valid.

1.5.1 Baseline Estimation

The goal of this paper is to investigate the effect of opportunity cost on female entrepreneurship in general. In the analysis, I focus on the marriage market as a specific form of opportunity cost and estimate its effect. My analysis focuses on the historical period before and after World War II because I will later use the war as a quasi-random experiment and exploit the casualty rate as exogenous variation. However, my results from this particular historical period can generalize to recent times because the tradeoff between marriage and career is still one of the most important life decisions for women today and the historical evidence is highly relevant for present-day policies.

To construct the sample, I include counties whose geographic boundaries did not change between 1940 and 1950 and follow the procedure described in the previous section. I calculate the sex ratio by dividing the female adult population by the male adult population times 100 and use it to measure local marriage market conditions. I then stack the data for the census years 1940 and 1950 and estimate the following the baseline equation:

\[ y_{it} = \phi R_{it} + X_{it}' \beta_t + \gamma_{1950} + \delta_s + u_{it} \]  

(1.1)

where \( y_{it} \) is the entrepreneurship variable in county \( i \) and year \( t \), \( R_{it} \) is sex ratio, \( X_{it}' \) are control variables interacted with the 1950 dummy for county, \( \delta_s \) denotes year fixed effects, and denotes state fixed effects. By including state fixed effects, I can estimate the coefficients by comparing female entrepreneurship in a county to that in other counties in the same state. The time dummy variable captures the nationwide increase in entrepreneurship from 1940 to 1950. I measure entrepreneurship activities from 1930 to 1940 for the 1940 census and from 1945 to 1955 for the 1950 census. I exclude the period when the US was engaged in World War II—from 1940 to 1945—when most men were deployed overseas and away from their civil lives. As a robustness check, I vary the size of the measurement window of entrepreneurship.
activities. The results do not change significantly.

The baseline estimating equation suffers from endogeneity. Regressing female entrepreneurship variables on sex ratio directly is problematic because of the confounding factors. Some unobserved county characteristics could be correlated with both the residual in the entrepreneurship outcome variable and the local sex ratio. Therefore, the OLS estimates do not reflect the causal effect and will include the endogenous effect from the unobserved county characteristics. For instance, women in more developed regions are better educated and more prepared for starting new businesses. However, people in these regions tend to have more liberal attitudes towards women in business and banks tend to be more open towards lending to women entrepreneurs. It is therefore hard to separate these effects. For another example, regions with agriculture or heavy industry have more men available for women. At the same time, industries in these regions are less suitable to women entrepreneurs. It is challenging to separate the effect of the marriage market from the effect of the industry structure. In either example, the OLS estimates will likely be biased.

1.5.2 Instrument for Marriage Market

To deal with the endogeneity problem discussed in the last subsection, I instrument marriage market conditions with World War II casualty rates and run a two-stage least square regression. On average, about 2.3 percent of a county’s young male population died during World War II, though there was significant variation across counties and states, as shown in Figure 1.5. The identification in this paper relies on the exogenous component in the cross-county variations in the casualty rate, which can be attributed to “luck” on the battlefield. For instance, Bedford, VA—with a population of just over three thousand at the time—suffered the most severe one-day loss in the US during World War II. Twenty-two men from the town died in the Normandy campaign because they were in the same company that was part of the first wave to hit the beaches in Normandy and got almost wiped out.

The exclusion restriction will be violated if there are unobserved county characteristics that systematically affected both the casualty rate and the female entrepreneurship rate. They will
Figure 1.5: Distribution of County-level Casualty Rate

This figure shows the distribution of county-level casualty rate, calculated as a county’s number of World War II casualties divided by its population between age 20 and 40 in 1940. The casualty rate in this histogram is winsorized at 5% on the upper bound. Source: author’s calculation using military records from the National Archives.

also affect long-term interpretation if they have persistent confounding effects on long-term change in female entrepreneurship and other outcomes of interest. A county’s casualty rate is a product of the enlistment rate and the individual casualty rate. For instance, because of draft procedures described in Section 3, certain historical factors affected both military enlistment and mortality differently across regions. Some of these historical factors are tightly related to economic outcomes such as percent of the population who are white or who are college-educated. To deal with such issues, I include potential covariates as control variables in the 2SLS regression and perform several robustness checks, which I describe below when presenting my results.

Table L.2 reports economic and noneconomic factors discussed in Section 3 that influenced enlistment rates. Results are consistent with Acemoglu et al. (2004). Columns 1 and 2 regress
county enlistment rate on the aforementioned factors, without and with state fixed effects, respectively. First, because the draft boards avoided fathers and preferred young single men, a high percentage of young males in a county’s population predicted a high enlistment rate. Second, a high percentage of college-educated people predicted a high enlistment rate. This could be caused by the voluntary component of the armed forces; for example, the National Guard. College-educated people might have been more aware of the dangerous consequences of the defeat of the Allies and therefore more likely to volunteer, whereas organized labor, such as unions, was mostly opposed to conscription during the mobilization period. Third, to ensure food supply during wartime, the government granted deferments to farmers. Thus, a high percentage of agricultural workers in a county’s population predicted a low enlistment rate. Fourth, due to racial segregation, whites were more likely to be drafted than African Americans. Thus, a high percentage of whites in a county’s population predicted a high enlistment rate.

Columns 3 to 6 in Table 1.2 report the demographics of draftees from different counties. As the local boards drafted more people, they tended to dip into older married men, as shown in Columns 3 and 4, and less educated men, as shown in Column 5. Counties with more college-educated men also contributed better-educated inductees. Better-educated and older men were more likely to enter the Air Force than the infantry, as shown in Column 6.

Table 1.3 reports factors affecting individual mortality rate. First, a county’s age structure determines its casualty rate. Not only did older counties send fewer men to the battlefield, but those men were also less likely to die, as shown in Columns 2 and 3. Second, service branch assignment had a significant impact on individual survival probability. Column 4 shows that better-educated inductees were less likely to die. Columns 5 and 6 show that men who joined the Air Force were less likely to die. Besides factors examined in Table 1.3, public health conditions, which were correlated with local economic conditions, could also affect soldiers’ survival probabilities. However, during World War II, because of penicillin and other medical advances, battle death was the main cause of death (Cirillo 2008). Table 1.4 shows that public health conditions did not have a significant effects on individual mortality rate.
Table 1.2: Determinants of Enlistment Rate

This table presents factors that influenced a county’s enlistment rate and its enlistees’ demographics during World War II. The dependent variables are a county’s enlistment rate (Columns 1 and 2) and the demographics of its enlistees: average age at enlistment (Column 3), percent of enlistees who were single (Column 4), average years of education obtained (Column 5), and percent of enlistees who went to the Air Force (Column 6). Enlistment Rate is a county’s number of enlisted men divided by its male population. Pct Young is the percentage of people between age 20 and 40 in a county’s population. Pct White is the percentage of whites in a county’s population. Pct College is the percentage of college-educated people in a county’s population. Pct Farm is the percentage of agricultural workers in a county’s population. State fixed effects are included in each regression except in Column 1. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Enlistment Rate</th>
<th>Age</th>
<th>Pct Single</th>
<th>Education (Yr)</th>
<th>Pct Air Corp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enlistment Rate</td>
<td>7.378***</td>
<td>-1.100***</td>
<td>-0.779**</td>
<td>-0.104**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.719)</td>
<td>(0.066)</td>
<td>(0.384)</td>
<td>(0.047)</td>
<td></td>
</tr>
<tr>
<td>Pct Young</td>
<td>0.711***</td>
<td>0.323***</td>
<td>-18.188***</td>
<td>-2.892***</td>
<td>-0.249***</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.049)</td>
<td>(1.882)</td>
<td>(0.639)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Pct White</td>
<td>-0.005</td>
<td>0.035***</td>
<td>1.927***</td>
<td>0.340***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.159)</td>
<td>(0.023)</td>
<td></td>
</tr>
<tr>
<td>Pct Farm</td>
<td>-0.041***</td>
<td>-0.039***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct College</td>
<td>0.023</td>
<td>0.034***</td>
<td></td>
<td>4.982***</td>
<td>0.340***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.011)</td>
<td></td>
<td>(0.185)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Counties</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
</tr>
<tr>
<td>R2</td>
<td>18.4%</td>
<td>76.1%</td>
<td>69.5%</td>
<td>50.5%</td>
<td>83.0%</td>
</tr>
</tbody>
</table>
I use infant death rate from the Vital Statistics and unemployment rate before the war to approximate for the health conditions of enlisted soldiers in a county.

Table 1.3 further shows that after controlling for the demographics and branch assignments of the enlistees, there was still significant residual variation in casualty rate across counties that could be attributed to battlefield luck. R-squared is around 42 percent across all regressions in Table 1.3. The remaining 58 percent of unexplained variation casualty is due to individual mortality and is plausibly random. The effect of omitted variables is likely small. Back-of-envelope calculation shows that unexplained variation in enlistment explains only 10 percent of casualties. About 24 percent of enlistment rates are unexplained. Multiplying it with R-squared in Table 1.3 one get $24\% \times 42\% \approx 10\%$. Empirical variations in casualty rates are also consistent with theoretical values assuming individual mortality follows independent Bernoulli distribution, as shown in Figure A.1.

There are also wide variations in enlistment and casualty rates at state level. Panel (a) of Figure 1.6 plots the casualty rate as a percentage of men eligible for military service in a state and shows that enlistment rate is the most important determinant of casualty rate. Table 1.3, too, shows that a higher enlistment rate predicts a higher casualty rate in a county; the coefficient, 2.8, percent is broadly in line with the overall casualty rate. In Panel (b), after controlling for enlistment rate, the relationship between the casualty rate (as a percentage of enlistees) and the enlistment rate is almost flat. However, there are still meaningful variations around the fitted regression line across states. These variations could be driven by luck on the battlefield as well as by the demographics of enlistees.

Although an average casualty rate of a little over two percent might seem small compared to that of some other countries during World War II, it could still have a significant impact on social outcomes for the following reasons. First, the impact on the marriage market is concentrated among certain birth cohorts. In the next section, I will show more detailed evidence on the age-specific impact on sex ratio. Second, because women are not likely to marry younger men, the shortage of men in the local marriage market would not be replenished naturally over time except through migration. This could have a long-lasting impact on the
Table 1.3: Determinants of Casualty Rate

This table presents factors that influenced a county’s casualty rates during World War II. The dependent variable is Casualty Rate, calculated as a county’s number of casualties divided by its population. Enlistment Rate is a county’s number of enlisted men divided by its male population. Enlistment: Age is the average age of enlistees at enlistment. Enlistment: Pct Single is the percent of enlistees who were single at enlistment. Enlistment: Educ. (Yr) is the average years of education enlistees had at enlistment. Enlistment: Pct Air Force is the percent of enlistees who went to the Air Force. State fixed effects are included in each regression. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: County Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct Young</td>
<td>0.005***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct White</td>
<td>0.002***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct Farm</td>
<td>-0.0003***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct College</td>
<td>0.001***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Enlistment Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlistment Rate</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Enlist.: Age</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlist.: Pct Single</td>
<td></td>
<td>0.0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlist.: Educ. (Yr)</td>
<td></td>
<td></td>
<td>-0.015*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.008)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlist.: Pct Air Force</td>
<td></td>
<td></td>
<td></td>
<td>-0.005***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observations</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
</tr>
<tr>
<td>R2</td>
<td>31.10%</td>
<td>42.10%</td>
<td>42.20%</td>
<td>42.10%</td>
<td>42.30%</td>
<td>43.70%</td>
</tr>
</tbody>
</table>
These figures show the relationship between casualty rate and enlistment rate across states. Panel (a) calculates casualty rate as a state’s number of casualties normalized by its male population who were eligible for military services. Panel (b) calculates casualty rate as a state’s number of casualties normalized by its number of enlisted men.

**Figure 1.6: State-level Enlistment Rate and Casualty Rate**
Table 1.4: Public Health Condition and Casualty Rate

This table reports the effect of public health condition on a county’s casualty rate during World War II. The dependent variable is Casualty Rate, calculated as a county’s number of casualties divided by its population. State fixed effects are included in each regression. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enlistment Rate</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(-0.001)</td>
<td>(-0.001)</td>
<td>(-0.001)</td>
<td>(-0.001)</td>
</tr>
<tr>
<td>Infant Death Rate</td>
<td>-0.001</td>
<td></td>
<td></td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(-0.001)</td>
<td></td>
<td></td>
<td>(-0.001)</td>
</tr>
<tr>
<td>Unemp. Rate</td>
<td></td>
<td>-0.0003</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.001)</td>
<td>(-0.001)</td>
<td></td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observations</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
</tr>
<tr>
<td>R2</td>
<td>42.1%</td>
<td>45.7%</td>
<td>42.1%</td>
<td>45.7%</td>
</tr>
</tbody>
</table>

local marriage market. Finally, had the US casualty rate been higher, the social structure might have been disrupted to a much greater degree, as it was in countries where the battles were fought. Then the effect of the marriage market on entrepreneurship would have been much harder to identify.

1.6 Empirical Evidence on War Impact

In this section, I run a county-level two-stage least square (2SLS) regression to formally test the hypothesis that a worse marriage market led more women to start new businesses due to the lower opportunity cost of marriage. I exploit World War II casualties as an exogenous shock to the local marriage market and use county casualty rate to instrument sex ratio in the first stage. In the first subsection, I run the first-stage regression and report World War II’s impact on the marriage markets in US counties. In the second subsection, I run the second-stage regression and report the impact of the marriage market on female entrepreneurship.
1.6.1 Impact on Sex Ratio

I start by documenting the impact of military casualties on the local marriage market, measured by the county’s sex ratio. Counties with higher military casualty rates had, on average, lower sex ratios; that is, less-favorable marriage markets for women. Women in these counties were therefore less likely to marry and faced a lower opportunity cost of marriage. The marriage-market channel would suggest that these women would instead either enter the labor market or start their own businesses. In the first-stage regression, I regress a county’s sex ratio on its World War II casualty rate, controlling for the aforementioned factors that influenced enlistment and casualty rates. Specifically, I estimate the following equation:

\[ Y_c = \alpha + \beta \kappa_c + \phi X_c + \epsilon_c \]  

(1.2)

where \( Y_c \) denotes change in sex ratio from 1940 to 1950 in county \( c \), \( \kappa_c \) denotes the casualty rate in county \( c \), and \( X_c \) denotes the set of control variables for county \( c \).

Table 1.5 reports the first-stage results. Column 1 includes the historical factors discussed in Section 5.2 and shows that they did not affect the county casualty rate. The coefficients are statistically significant. Column 2 controls for those historical factors affecting only the enlistment rate. Column 3 controls for historical factors affecting both enlistment rate and individual casualty rate. Column 4 drops counties in the southern states.\(^6\) Columns 3 and 4 show that a casualty of one man per 100 population led to a decrease of about 3.3 men per 100 women in the sex ratio for all states and 2.5 men per 100 women for northern states only. As a placebo test, I show in Column 5 that the World War II casualty rate did not predict the sex ratio in 1980. Overall, Table 1.5 suggests that there is a strong relationship between casualties from World War II and local marriage market conditions, which is robust to the use of different specifications.

Figure 1.6 represents the relationship described in the previous table at the state level. As it shows, New Mexico experienced a higher casualty rate than other states during World

---

\(^6\) Southern states: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.
Table 1.5: War Impact on Sex Ratio

Panel (a) presents coefficient estimates in the first stage regression. Columns 1, 2 and 3 regress a county’s change in sex ratio from 1940 to 1950 on its casualty rate. Column 4 is a placebo test and regresses a county’s change in sex ratio from 1970 to 1980 on its casualty rate. Casualty Rate is a county’s number of casualties divided by its population between age 20 and 40 in 1940. Economic control variables include: Pct White, the percentage of whites in a county’s population; Pct College, the percentage of college-educated people in a county’s population; Pct Farm, the percentage of agricultural workers in a county’s population. Noneconomic control variables include: Pct Young, the percentage of people between age 20 and 40 in a county’s population. Column 3 excludes southern states in the regression. Panel (b) presents impact of World War II casualties on sex ratio by age group. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

(a) Overall

<table>
<thead>
<tr>
<th></th>
<th>Sex Ratio 1950</th>
<th>Sex Ratio 1980</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Casualty Rate</td>
<td>-3.106***</td>
<td>-3.282***</td>
</tr>
<tr>
<td></td>
<td>(1.021)</td>
<td>(1.023)</td>
</tr>
<tr>
<td>Controls: Economic</td>
<td>X X X X</td>
<td></td>
</tr>
<tr>
<td>Controls: Noneconomic</td>
<td>X X X X</td>
<td></td>
</tr>
<tr>
<td>State FE</td>
<td>X X X X</td>
<td></td>
</tr>
<tr>
<td>Excl. South</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-stat</td>
<td>17.57</td>
<td>17.37</td>
</tr>
<tr>
<td>Observations</td>
<td>2,290</td>
<td>2,290</td>
</tr>
<tr>
<td>R2</td>
<td>28.2%</td>
<td>28.4%</td>
</tr>
</tbody>
</table>

(b) By Age Group

<table>
<thead>
<tr>
<th>Age Group:</th>
<th>[15,20)</th>
<th>[20,25)</th>
<th>[25,30)</th>
<th>[30,35)</th>
<th>[35,45)</th>
<th>[45,55)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casualty Rate</td>
<td>-0.439</td>
<td>-0.596</td>
<td>-1.736***</td>
<td>-1.490**</td>
<td>-1.216**</td>
<td>-0.915*</td>
</tr>
<tr>
<td></td>
<td>(0.42)</td>
<td>(0.48)</td>
<td>(0.57)</td>
<td>(0.59)</td>
<td>(0.49)</td>
<td>(0.49)</td>
</tr>
<tr>
<td>State FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Counties</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
<td>2,290</td>
</tr>
<tr>
<td>R2</td>
<td>8.1%</td>
<td>20.7%</td>
<td>26.3%</td>
<td>19.5%</td>
<td>26.0%</td>
<td>32.5%</td>
</tr>
<tr>
<td>Sex Ratio in 1930</td>
<td>99</td>
<td>97</td>
<td>98</td>
<td>100</td>
<td>105</td>
<td>110</td>
</tr>
<tr>
<td>Sex Ratio in 1940</td>
<td>100</td>
<td>97</td>
<td>97</td>
<td>98</td>
<td>100</td>
<td>105</td>
</tr>
</tbody>
</table>
War II and one may be concerned that it influences the strength of the first stage. However, omitting counties in New Mexico barely affects the regression results and the second-stage results.

I next show that the effect on the marriage market is concentrated in certain birth cohorts. Because women do not usually marry younger men, this can have a long-lasting impact on the affected cohorts. I estimate the impact of World War II casualties on sex ratio separately for each age group with the following equation:

\[ Y_{ca} = \alpha + \beta \kappa_c + \phi X_c + \epsilon_{ca} \]  

(1.3)

where \( Y_{ca} \) denotes the change in sex ratio from 1940 to 1950 for age group \( a \) in county \( c \), \( \kappa_c \) denotes the casualty rate in county \( c \), and \( X_c \) denotes the set of control variables for county \( c \).

I present the regression results in Panel (b) of Table 1.5. The cohort aged 25 to 30 in 1950 was most affected and experienced the largest decrease in the sex ratio, consistent with the national-level statistics in Figure 1.4. Men in that cohort were between age 20 and 25 in 1945 and were therefore the most likely to have fought. The coefficient of -1.74 is broadly in line with the theoretical value of -2. This negative impact on sex ratio is not obvious when looking at the nationwide average sex ratio across time. Biologically, one would expect a cohort’s sex ratio to decline with age, as women tend to live longer. However, because there was a large influx of young male immigrants at the turn of the twentieth century, they contributed to a higher sex ratio in the older age groups (30-54) than in the younger age groups (15-29).

1.6.2 Impact on Entrepreneurship

Using the business registration data, I first calculate the following measures of entrepreneurship in a county: female and male entrepreneurship participation rates (EPRs) and female share
of entrepreneurship (FSE):

$$EPR_{Female} = \frac{\text{No. of New Female Businesses}}{\text{FemaleLaborForce}}$$ \hspace{1cm} (1.4)

$$EPR_{Male} = \frac{\text{No. of New Male Businesses}}{\text{MaleLaborForce}}$$ \hspace{1cm} (1.5)

$$FSE = \frac{\text{No. of New Female Businesses}}{\text{TotalNo.ofNewBusinesses}}$$ \hspace{1cm} (1.6)

I regress these measures on the instrumented county sex ratio in a pooled county-level regression for 1940 and 1950 with the following equation:

$$y_{it} = \phi R_{it} + X'_{it}\beta_{t} + \gamma_{1950} + \delta_{s} + u_{it}$$ \hspace{1cm} (1.7)

where $y_{it}$ is one of the entrepreneurship measures for county $i$, $R_{it}$ is the sex ratio instrumented by the casualty rate, $X_{it}$ are control variables interacted with the 1950 dummy for county $i$, $\gamma_{1950}$ denotes year fixed effects, and $\delta_{s}$ denotes state fixed effects. I exclude counties that were not yet in the business registration panel in 1950. Control variables include the aforementioned historical factors that affected casualty rates and economic outcomes: agriculture share, education level, and age structure. The state fixed effects, $\delta_{s}$, allow me to control for unobservable state characteristics that were fixed over time and could generate systematic differences in entrepreneurship participation rate by gender. The year fixed effects, $\gamma_{1950}$, allow me to control for shocks to the entrepreneurship participation rate that were common to all counties. For instance, there was a pent-up demand for entrepreneurship immediately after the war because the men deployed overseas had not been able to start businesses.

Column 1 in Table L.6 shows that a decrease of one man per 100 women increases the female entrepreneurship rate by 0.57 basis points. The average female entrepreneurship rate was 0.3 basis points in 1950. Column 3 shows that it decreases the male entrepreneurship rate by 0.48 basis points. The average male entrepreneurship rate was 2.1 basis points in 1950. The decrease could be due to a decrease in the male population caused by casualties. Column 4 shows that it increases female share of new businesses by 1.42 percent overall. Impact on the entrepreneurship rate is statistically significant at the one-percent level for
women and at the five-percent level for men. The finding that the marginal propensity for entrepreneurship was higher for women than for men indicates that there was an overall increase in entrepreneurship.

**Table 1.6: War Impact on Entrepreneurship**

This table presents IV regression results using the US casualties during World War II. **Entrepreneurship Rate: Female (Male)** in Columns 1 and 2 is a county’s number of new businesses founded by women (men) divided by its total female (male) labor force. **Female Share of Entrepreneurship** in Column 3 is a county’s number of new businesses founded by women divided by its total number of new businesses. **Sex Ratio** is the fitted value for sex ratio from the first stage. Control variables include: **Pct White**, the percentage of whites in a county’s population; **Pct College**, the percentage of college-educated people in a county’s population; **Pct Farm**, the percentage of agricultural workers in a county’s population; **Pct Young**, the percentage of people between age 20 and 40 in a county’s population. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Entrepreneurship Rate: Female (basis point)</th>
<th>Entrepreneurship Rate: Male (basis point)</th>
<th>Female Share of Entrepreneurship (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrumented Sex Ratio</td>
<td>-0.570**</td>
<td>0.478*</td>
<td>-1.422</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Year FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observations</td>
<td>1,018</td>
<td>1,018</td>
<td>1,018</td>
</tr>
<tr>
<td>Mean in 1950 (%)</td>
<td>0.003</td>
<td>0.021</td>
<td>0.135</td>
</tr>
</tbody>
</table>

Migration could play a role in the estimation in both stages. People migrate away from poorer regions. If there is no gender bias in migration, it will not affect my estimates. If male migration is more sensitive to the level of local economic activity, poorer regions will have fewer men who stay relative to the women who stay. As a result, women will become more represented in entrepreneurship. As long as migration is exogenous to the World War II casualty rate, my interpretations are still plausible. However, poorer regions might have experienced more casualties. Then my estimates could be biased upward. I partially address this issue by controlling for population growth. One might also argue that there is likely to be a sorting implication between women and men in an unbalanced marriage market, as the
women who remain single are less competitive than those who find a husband. However, this claim would imply that the average woman’s propensity for entrepreneurship would be even higher than the estimate.

1.7 Mechanism

In this section, I investigate the possible mechanisms by which war casualties affected female entrepreneurship: easing of financial constraints, lower opportunity cost of marriage and less competition from men. To separate different mechanisms, I first compare the entrepreneurship decisions of single and married women facing a less-favorable marriage market. I then exploit gender complementarity in business sectors and examine whether the increase in female entrepreneurship was due to lower competition from men.

However, there could be other contemporaneous factors that influenced women’s entrepreneurship decisions. First, a war widow could have inherited the family business from her husband if they had run it together before the war. Second, many women entered the labor force during the war when men were deployed overseas. Having acquired relevant work experience and experienced the rewards of employment, they were more economically independent and were more likely to start a business. Third, a significant percentage of women who worked during the war were forced to give up their jobs after the war ended because of the government’s veteran reintegration efforts. This could have pushed them to make a living by starting their own businesses, especially when it was hard for them to find a husband.

1.7.1 Marriage Market

Because casualties affected single women who were looking for husbands as well as war widows who received lump sum payments, both the opportunity-cost and access-to-capital channels could in theory have affected female entrepreneurship after World War II. Although the two may suggest observationally equivalent predictions in this context, it is important to distinguish them, as the resulting policy implications for sustaining female entrepreneurship can be drastically different. I separate these effects by calculating the entrepreneurship rate
for single and married women separately. Table 1.7 shows the age of a series of birth cohorts in different years. The 1910–1930 cohorts were most affected by the war. During the 1950s, most women got married before age 25. Women born after 1920 were therefore largely single in 1940 and more likely to be affected through *the marriage-market channel*, whereas women born before 15 were largely married in 1940 and more likely to be affected through *the access-to-capital channel*.

Table 1.7: *Cohort Table*

This table demonstrates age of various birth cohorts (rows) in different calendar years (columns). During the 1940s and 1950s, most women got married between age 20 and 25.

<table>
<thead>
<tr>
<th>Year of Birth</th>
<th>1930</th>
<th>1935</th>
<th>1940</th>
<th>1945</th>
<th>1950</th>
<th>1955</th>
<th>1960</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>1905</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>1910</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>1915</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>1920</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>1925</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>1930</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>1935</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>1940</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By matching the business owners’ information with the individual-level 1940 census, I obtain marital status and year of birth for female entrepreneurs who started businesses from 1945 to 1960. The 1940 census should capture most of them, as it is very unlikely for a woman to start a business before age 20. With the matched demographic information, I separate the female entrepreneurship participation rate into participation by single and by married women, based on their marital status in 1940. I proxy war widows’ entrepreneurship participation rate with that of married women. The marriage market should not have had a strong impact on the entrepreneurship decisions of married women who did not lose husbands during the war. Therefore, the change in female entrepreneurship participation rate should
only be attributable to war widows, although the estimate will suffer from low-power issues.

Table 1.8 reports the impact of World War II casualties on marital outcomes and Panel (a) of Table 1.9 provides evidence in favor of the marriage-market channel as opposed to the access-to-capital channel. Columns 1 and 2 in Table 1.8 show that, after the war, counties with more casualties had more single women and war widows, respectively. Both are statistically significant. At the same time, there are fewer single men, as shown in Column 3. With fewer marriages, the birth rate—approximated by the percent of population under age five—was lower in the 1960s, as shown in Column 4. I then report the regression results on the marriage-market channel in Panel (a) of Table 1.9. Comparing Columns 2 and 4 shows that most of the new women entrepreneurs were single, supporting the marriage-market channel. The financing channel for war widows receiving startup capital seems statistically nonsignificant.

Table 1.8: Impact on Marital Outcomes

This table presents the effects of imbalanced sex ratio on marital outcomes. The dependent variable is percentage of female population who were single (Column 1), percentage of female population who were widowed or divorced (Column 2), percentage of male population who were single (Column 3), and percentage of population who were under age 5 as an approximate for birth rate (Column 4). Sex Ratio is the fitted value for a county’s change in sex ratio from 1940 to 1950. Control variables include: Pct White, the percentage of whites in a county’s population; Pct College, the percentage of college-educated people in a county’s population; Pct Farm, the percentage of agricultural workers in a county’s population; Pct Young, the percentage of people between age 20 and 40 in a county’s population. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Pct Female:</th>
<th>Pct Female:</th>
<th>Pct Male:</th>
<th>Birth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
<td>Widow/Divorce</td>
<td>Single</td>
<td></td>
</tr>
<tr>
<td>Sex Ratio</td>
<td>-1.266***</td>
<td>-0.260*</td>
<td>0.358</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.151)</td>
<td>(0.281)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observations</td>
<td>2,284</td>
<td>2,284</td>
<td>2,284</td>
<td>2,284</td>
</tr>
</tbody>
</table>

My analysis cannot completely rule out an alternative interpretation. If a woman who was single in 1940 got married after the war and started a business, she would show up as a single business owner in my analysis. Thus, it is theoretically possible that all the new
women entrepreneurs were single in 1940 and got married afterwards. Then the interpretation would be that marriage provided income support and enabled women to start their own businesses. However, this interpretation is not entirely reasonable because, during the 1950s, men expected their wives to raise children and keep house. Marriage would incur a prohibitive career cost on women and they would be less likely to start their own businesses right after marriage. Therefore, it is still reasonable to expect that business owners who were identified as single in my sample were indeed single when they founded their businesses.

Table 1.9: Mechanism: Marriage Market and Gender Substitution

This table presents results investigating the mechanisms through which World War II casualties impacted female entrepreneurship. Panel (a) compares the marriage market channel and the access-to-capital channel. Column 1 presents results for entrepreneurship rate among single women, defined as the number of new businesses founded by single women normalized by the female labor force population. Column 2 presents results for entrepreneurship rate among married women, defined as the number of new businesses founded by married women normalized by the female labor force. Panel (b) compares across business sectors. Column 3 presents results for female entrepreneurship rate in female-dominated sectors, defined as the number of new businesses founded by women in female-dominated sectors normalized by the female labor force. Column 4 presents results for female entrepreneurship rate in gender-neutral sectors, defined as the number of new businesses founded by women in gender-neutral sectors normalized by the female labor force. 

Sex Ratio is the fitted value for a county’s change in sex ratio from 1940 to 1950. Control variables include: Pct White, the percentage of whites in a county’s population; Pct College, the percentage of college-educated people in a county’s population; Pct Farm, the percentage of agricultural workers in a county’s population; Pct Young, the percentage of people between age 20 and 40 in a county’s population. Standard errors are included in parentheses below the coefficient estimates. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th>Female Entrepreneurship Rate (basis point)</th>
<th>(a) By Marital Status</th>
<th>(b) By Business Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single (1)</td>
<td>Married (2)</td>
</tr>
<tr>
<td>Sex Ratio</td>
<td>-1.152* (0.628)</td>
<td>-0.447 (0.288)</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Year FE</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observations</td>
<td>427</td>
<td>427</td>
</tr>
</tbody>
</table>
1.7.2 Substitution Effect

As noted previously, the increase in female entrepreneurship may simply have been due to less competition from men. In counties where fewer men came back, there were more business opportunities for women. To tease out this purely mechanical displacement effect, I estimate Equation 5 by business sector. I classify business sectors by gender bias: *male-dominated*, *gender-neutral*, and *female-dominated*. This substitution effect is expected to be strongest in the gender-neutral sectors, where men and women compete equally, and weakest in the female-dominated sectors. I find that new female entrepreneurs started businesses in both gender-neutral and female-dominated sectors. This indicates that the increase in female entrepreneurship is more than a purely mechanic substitution.

To infer the gender bias of a business, I apply text-mining techniques to company names. After tokenizing and stemming words in company names, I exclude common words, human names, and words with fewer than 1,000 occurrences. I calculate the percentage of businesses owned by women for each token and use it as a measure of gender bias. I then calculate the gender bias of a company name based on its bag-of-words representation from the scored tokens, weighted by their inverse frequency. Words from the female-dominated sectors include “housekeeping,” “secretarial,” “teacher,” “grooming,” “transcription,” “elementary,” ”bookkeeping,” “parent,” and “quilt.” Words from the male-dominated sectors include “carpentry,” “engineering,” “automotive,” “drilling,” “gunsmithing,” “hunting,” “aircraft,” “hydraulic,” “snowplowing,” and “inspection.” I finally sort companies into the aforementioned (equal-sized) groups according to gender bias score.

After sector classification, I separate female entrepreneurship participation rate into participation in gender-neutral sectors and participation in female-dominated sectors. I change the outcome variable in Equation 5 to the newly constructed measures:

\[
EPR_{F,FD} = \frac{\text{No. of New Female Businesses in Female-Dominated Sectors}}{FemaleLaborForce} 
\]  
\[
EPR_{F,GN} = \frac{\text{No. of New Male Businesses in Gender-Neutral Sectors}}{MaleLaborForce} 
\]
Panel (b) of Table 1.9 reports the regression results. Columns 1 and 2 present results for female entrepreneurship in the female-dominated sectors and Columns 3 and 4 for the gender-neutral sectors. The 2SLS estimates are negative, statistically significant, and economically large. New female entrepreneurs started businesses in both types of sector. The increase in entrepreneurship participation in the gender-neutral sectors could be due to the unfavorable marriage market as well as to less competition from men. However, women were 60 percent more likely to start a business in the female-dominated sectors than in the gender-neutral sectors. This provides strong evidence that the marriage-market channel itself had a considerable and independent impact on female entrepreneurship, without the mechanical substitution effect.

1.8 Long-Term Effects

Column 5 in Table 1.5 shows that World War II casualties did not have an effect on the sex ratio in 1980. However, the war’s impact on female entrepreneurship has persisted even to 2015, when the latest business registration is available. To better illustrate how military casualty rates and female entrepreneurship are related, Figure 1.7 plots separately the average percentage of new businesses founded by women in high-, medium-, and low-casualty counties, normalized with respect to the percentage in 1945 right after the war ended. It shows that counties with more casualties experienced a more positive persistent change in female entrepreneurship than counties with low casualties. Female share of entrepreneurship in high-casualty counties has consistently been about 1.5 percent higher than that of medium-casualty counties since 1945.

This difference in female entrepreneurship is not accounted for by differences in age structure, education level, agricultural share, and other control variables. My interpretation is that it was caused by greater participation by so-called “shadow spouses”—women who did not marry due to a worse marriage market. The long-term transmission mechanisms are not formally tested in this paper. It is possible, for example, that women entrepreneurs formed support groups that encouraged more women to start their own businesses. Regions
Figure 1.7: *War Impact on Female Share of Entrepreneurship*

This figure shows female share of new business founded in the US, normalized by its value in 1945 when World War II ended. The three groups of counties are formed by sorting on casualty rate: blue line for low-casualty counties, green line for medium-casualty counties and red line for high-casualty counties. Source: author’s calculation using business registration records, military records from the National Archives and data from the US decennial censuses.

with women entrepreneurs may have developed more liberal view towards women running businesses and updated their belief positively.

More importantly, Figure 1.7 shows that World War II did not have an impact on female entrepreneurship before the war, lending support to the hypothesis that the postwar shifts were caused by the war and not by differential long-run trends in female entrepreneurship. We see no such trends in female entrepreneurship across counties with varying casualty rates in the pre-war period, although the medium-casualty counties experienced slightly larger decrease.

1.9 Conclusion

In this paper, I exploit variation in the marriage market across the US to provide causal evidence on how opportunity cost influences women’s entrepreneurship decisions. I show
that marriage is an important form of opportunity cost hindering women from starting new businesses. Using novel business registration data detailing owners’ information and individual-level census data, I show that single women are more active in starting new businesses when they face worse prospects in the marriage market.

Using US casualties during World War II to isolate exogenous shocks to local marriage markets, I identify a causal link between the marriage market and female entrepreneurship. Specifically, I first show that war casualties led to a permanent localized shortage of men and had a meaningful impact on local marriage markets and marital outcomes. Counties with more casualties had more imbalanced sex ratios and worse marriage markets for women. High- and low-casualty counties appear similar to each other in terms of geographic location, employment, and so on, apart from the exogenous shocks to the sex ratios. Single women in counties with more casualties were more entrepreneurial, especially in female-dominated business sectors. This effect is robust, having persisted to our own time.

Finally, I provide evidence on the mechanism and show that the impact is not entirely due to mechanical gender substitution and to war widows' access to capital. Marginal women entrepreneurs tend to be single and to start new businesses in female-dominated sectors. My results provide suggestive evidence that it is opportunity cost rather than access to capital that hinders women from starting new businesses. Thus, merely providing financial subsidies might not be as effective in encouraging female entrepreneurship as policies that reduce opportunity cost or that change incentives in business careers, consistent with the implication in Goldin (2014).

While I focus on the marriage market and a particular period after World War II, I believe that my results can generalize to recent periods and to other forms of opportunity cost. Childrearing and housekeeping still impose significant career costs on women today. It is important to make business culture more family-friendly, by means such as longer maternity leave, more flexible work schedules, and less stigma attached to leaves taken for family reasons. Lessons on the relationship between opportunity cost and entrepreneurship apply to men as well. Reducing the opportunity cost of entrepreneurship—for example, by allowing employees
to work on experimental projects—might incubate more innovations and new businesses. Finally, my approach can be adapted to study causes and effects of entrepreneurship in other settings and to evaluate policy impacts on business formation in the twentieth and twenty-first centuries.

Findings in this paper provide motivation for future work on opportunity cost and female entrepreneurship in two directions. The first is identifying the intergenerational transmission mechanism of entrepreneurship. This will shed light on the lasting impact of World War II on female entrepreneurship. Successful women entrepreneurs can inspire future generations of women through the role model effect, both through public media and within a household. They can also shape social norms and strengthen public belief in women’s leadership ability. They might also form interest groups and provide mentorship and guidance to future women entrepreneurs. The second direction for future work is understanding the long-term real effects of female entrepreneurship. Mobilizing fully one half of the potential entrepreneurial human capital holds promise to yield large societal productivity gains that can be shared economy-wide. Moreover, women-owned businesses are likely to attract more women employees and thus reduce the gender gap in employment. It is therefore important to investigate more thoroughly the relationship between gender diversity in entrepreneurship and economic growth and to identify the mechanisms.
Chapter 2

Talking Your Book: Evidence from Stock Pitches at Investment Conferences

2.1 Introduction

Hedge funds are usually associated with secrecy and are not known for sharing their investment ideas for free. However, since 2008, a new type of industry event—investment conferences—has become popular and emerged as a hallmark event in the investment management industry. At these conferences, prominent hedge fund managers such as David Einhorn from Greenlight and David Tepper from Appaloosa pitch their investment ideas externally to the audiences. They take place throughout the year at various locations and are usually open to anyone who registers and purchases a ticket. Most of these conferences are organized by non-profit organizations, foundations and industry associations benefit charitable causes. They are well attended by a broad range of financial institutions, including activists, fundamental equity funds, investment advisors and sell-side research analysts. Together, the speakers and the

---

1Example investment conferences include Sohn New York, Sohn San Francisco, Value Investing Congress, Great Investors’ Best Ideas, and Excellence in Investing. Please see Section B.1 for a description of the Sohn Investment Conference from its website, its registration page and the price schedule.
attendees represent a sizeable portion of the capital market. Furthermore, the investment ideas presented at these conferences are closely followed in the financial media and on investment blogs. As a result, these stock pitches are market-moving events. During the first two days of the pitches, long pitches outperform the market by 1.1% whereas short pitches underperform by more than 2% and their trading volumes spike up.

In this paper, I hand-collect a novel dataset on investment ideas pitched at these investment conferences and document this new industry phenomenon. Specifically, I evaluate the performances of these investment ideas and analyze the holdings patterns of the hedge funds that pitch them. I show that pitched stocks generally have positive risk-adjusted returns both before and after the pitches. However, the funds that pitch sell their pitched stock after the conferences. In addition, I analyze how other investors react to these investment pitches and find wide heterogeneity in behavior between hedge funds and mutual funds.

There is an increasing need to understand hedge funds’ behaviors and their implications for market efficiency. Albeit still small compared to the mutual fund industry, the hedge fund industry has grown tremendously over the past two decades. The assets under management have grown from less than $50 billion in 1990 to more than $1 trillion in 2006. Their behaviors are also closely followed on the news and often make headlines because they frequently take speculative bets and play an important role in the price discovery process. One important type of behavior is disclosure behavior—how hedge funds release their book information and investment strategies to the public, either mandatorily or voluntarily. For instance, hedge funds meeting certain criteria are required by regulation to disclose their portfolios through SEC 13F and 13D filings. Hedge fund managers sometimes also voluntarily speak on TV to advocate their recent investment theses. They are likely to exhibit different disclosure behaviors at different stages of investment.

The investment conference is a unique voluntary disclosure channel compared to others channels for three important reasons. First, these conferences are highly-coordinated events and attract significant attention in the investment management industry, as described earlier. A wide variety of market participants are present at the same time. Stock pitches at the
conferences receive high levels of attention in the financial news. Second, hedge funds managers who pitch investment ideas have the attention of the audience throughout their speeches. They can take time to walk through their investment theses with control and flexibility. They can be more persuasive in person than they can be on TV or in other venues. Anecdotal evidence suggests that some audience members even begin trading the pitched stocks in the middle of the presentations. Third, it is easier for an investor to gauge general interest in pitched stocks based on soft information when interacting with other investors in person at these conferences.

Due to these unique features, investment conferences give hedge fund managers strong incentives to disclose their “best ideas” strategically and to use these conferences as a new tactic to manage their portfolio positions. Unlike mandatory regulatory filings, voluntary disclosures like stock pitches at conferences are often driven by private motivations, and the funds face complex incentives. On one hand, hedge funds have incentives to pitch bad stocks. Knowing that their disclosure can attract more investors, they may want to use investment conferences to create their own liquidity events for the stocks they want to liquidate. An increase in demand for a stock can push up its price in the short-term (Greenwood 2005). If they see that a stock in their portfolios is going to underperform, they can disclose these positions publicly to induce favorable short-term price pressure. However, doing so damages their reputation over the long run and is not a sustainable strategy in a repeated game.

On the other hand, hedge funds have incentives to pitch good stocks. Young hedge funds can use these events to build a reputation for their investment skills and attract capital from future LP investors. For instance, David Einhorn from Greenlight Capital pitched short Lehman Brothers at the Value Investing Congress in 2007 and gained significant acknowledgment. Moreover, hedge funds— especially activists and short sellers—can use these events to advocate their investment thesis and improve their returns. Due to limits to arbitrage, they often cannot correct mispricings themselves (Shleifer and Vishny 1997).

David Einhorn mentions in the book Fooling Some of the People All of the Time, A Long Short Story that some audience members left during a presentation to either sell the short pitch or tell their clients to sell.
especially short sellers. As a result, they often publicly disclose their investment ideas to “recruit” more followers to facilitate price discovery (Ljungqvist and Qian 2016). They publish detailed research reports online and talk about their investment theses in public. Activists, in particular, need shareholders’ votes to implement the strategic plans they propose. For instance, Bill Ackman from Pershing Square pitched long JC Penny at the Sohn Conference in New York in 2012 to further advocate his proposed turnaround strategy.

To understand the motives of hedge funds, I collect stock pitches at investment conferences from various online sources. I apply textual analysis techniques to extract fund names and stock tickers and merge them to returns and 13F holdings. First, I ask whether stocks pitched at investment conferences do in fact outperform. I conduct event studies around these stock pitches and calculate risk-adjusted returns using Fama-French 3 factors with momentum and DGTW stock characteristics. I find that pitched stocks outperform after the pitches. A calendar-time strategy that buys pitched stocks and sells non-pitched stocks generates an annualized return of 8.2% and has an annualized Sharpe ratio of 0.67. However, the risk-adjusted return of pitched stocks is much smaller after the pitches than before the pitches. Pitched stocks earn a cumulative risk-adjusted return of 20% over the 18 months before the pitch and 7% over the 9 months after the pitch. Furthermore, half the post-conference risk-adjusted returns revert after another 9 months.

Using the 13F holdings and Form D filings, I investigate whether the hedge funds hold onto their pitched stocks and their motives for doing so. Contrary to their claims, I find that they do not hold onto their pitched stocks any longer than to non-pitched stocks after the conferences although the pitched stocks do outperform. They tend to pitch stocks in which they have larger papers and want to take profit. These pitched stocks do not have to be bad stocks. They can simply be stocks that are not as attractive as other potential ideas in the hedge funds’ investment opportunity set. Newly bought stocks outperform the pitched stocks that the funds sell after the conferences. Moreover, I find that stock pitches at investment conferences help hedge funds raise money. Therefore, they have strong incentives not to pitch obviously bad stocks.
Finally, I turn to the behaviors of other investors. There can be multiple possible responses to these pitches. On one hand, unsophisticated investors may naively follow what sophisticated investors pitch as a good investment idea. As a result, even sophisticated investors may also follow these pitches due to rational herding. On the other hand, if other investors know that these hedge funds are not genuinely sharing their best ideas, they may not want to follow these pitches. Furthermore, the holdings of the prominent hedge funds are usually available from 13F filings albeit with a time lag. To provide evidence on possible heterogeneity in investor behaviors, I separate active 13F institution investors from passive ones and categorize them into hedge funds, mutual funds and investment advisors. I show that although they all follow the pitches in the short term, they exhibit quite different trading behaviors around these investment conferences. Specifically, hedge funds and mutual funds trade in opposite ways. Other hedge funds show trading patterns very similar to those of the hedge funds that are pitching stocks. This suggests that they either run correlated strategies or share information with each other. However, mutual funds sell before the pitches and buy afterwards. The subsequent buying pressure from mutual funds and investment advisers is a possible causes of the positive excess return after the pitches, even when the hedge funds are selling the pitched stocks. An alternative hypothesis is that mutual funds are passive liquidity providers. However, the positive return after the pitches is also harder to reconcile if there is selling pressure only from the hedge funds.

The remainder of the paper is organized as follows. Section 2 relates my paper to the current literature on fund performance and the strategic behaviors of investment management firms. Section 3 describes the data. Section 4 explains the methodology and presents empirical results on the performances of investment pitches. Section 5 analyzes the trading behaviors of different investors using 13F holdings and explores their motives. Section 6 concludes.

2.2 Related Literature

This paper first adds to the extensive literature on the performance of active money management. On one hand, current studies find insignificant excess return among professional fund
managers. Jensen (1968) documents that mutual funds do not outperform a buy-the-market-and-hold policy under CAPM. Other papers corroborate this finding with similar results (Carhart, 1997; Grinblatt and Titman, 1989, 1993). Gruber (1996) explains why investors still invest in mutual funds even if their performance has been inferior to that of index funds. On the other hand, other papers find that mutual funds exhibit some stock selection ability using stock-characteristic-based benchmarks but that their net returns underperform (Daniel and Titman, 1997; Daniel, Grinblatt, Titman, and Wermers, 1997; Wermers, 2000). More recently, numerous papers find strong evidence for stock-level selective skills among high-conviction and concentrated active holdings (Kacperczyk, Sialm, and Zheng, 2005; Alexander, Cici, and Gibson, 2007; Cremers and Petajisto, 2009; Pomorski, 2009; Cohen, Polk, and Silli, 2010; Agarwal, Fos, and Jiang, 2013a; Rhinesmith, 2014). These “best” ideas generate positive alpha, suggesting that fund managers exhibit meaningful stock-picking skills. This inconsistence can be reconciled by the fact that fund managers have incentives to add stocks with little conviction to their portfolios (Berk and Green, 2004; Cohen et al., 2010). These stocks can diversify portfolio risk, reduce the chance of lagging behind benchmarks and their peers, provide liquidity for redemptions and increase the fund’s capacity to charge management fees.

This paper also helps illuminate the strategic behaviors of investment managers and how the market processes their information content, in particular portfolio disclosure. Fund managers disclose information on their portfolios. These disclosures can have implications for other market participants. For instance, 13F filings are closely followed by investors and initial disclosures of new investments by notable hedge funds often lead to significant trading and price movements. Frank, Poterba, Shackelford, and Shoven (2004), Verbeek and Wang (2013) and Wermers, Yao, and Zhao (2010) examine the investment values of these public portfolio disclosures. Furthermore, portfolio disclosure influences how funds operate and invest (Musto, 1997, 1999; Wermers, 2000). Other strategic behaviors are also exhibited by investment managers. For instance, Gervais and Strobl (2015) analyze the optimal signal strategies based on managers’ skill level and model the dynamics of the pooling and separating equilibrium. Pension funds “window dress” their portfolios to
impress sponsors (Lakonishok, Shleifer, Thaler, and Vishny, 2004).

While the majority of previous research focuses on mutual funds because of their size and data availability, this paper specifically contributes to the emerging literature on the hedge fund industry and its strategic behaviors. Stulz (2007) gives an overview of research on the hedge fund industry. Fung and Hsieh (1997), Griffin and Xu (2009) and Bali, Gokcan, and Liang (2007) find that hedge funds’ returns are different from those of mutual funds and attempt to identify the determinants of hedge fund performance. More recently, some papers focus on strategic behaviors specific to hedge funds and the strategies they employ. Agarwal, Jiang, Tang, and Yang (2013b) find that hedge funds conceal their key holdings through confidential 13F filings. Brav, Jiang, and Kim (2015) provides a literature review on hedge fund activism. Activists usually use public disclosure as a tactics to promote their investment theses. Brunnermeier and Nagel (2004), Chen, Hanson, Hong, and Stein (2008), Griffin and Xu (2009) and Aragon and Strahan (2012) look at whether hedge funds take or provide liquidity to the market. Chen et al. (2008) and Griffin and Xu (2009) analyze the relationship between hedge funds and mutual funds in particular, and provide evidence that hedge funds profit from mutual fund distresses.

Finally, this paper is related to the literature on institutional constraint and behavioral bias among asset managers. Shleifer (1986), Greenwood (2005) and Greenwood (2008) document the downward-sloping demand curve in the capital market due to limits to arbitrage. Institutional investors can be subject to herding behavior (Scharfstein and Stein, 1990). Cohen, Gompers, and Vuolteenaho (2002) and Frazzini (2006) find that institutional investors under-react to news. Frazzini (2006) finds that mutual funds are also subject to the disposition effect as retail investors.

2.3 Data

I compile my dataset on stock pitches at investment conferences by scraping them from the Internet and linking them to traditional financial datasets including stock returns, trading volume and 13F holdings. I collect the investment pitches systematically in three steps.
First, I compile a list of investment conferences based on financial news and conversations with industry practitioners. I collect the conference schedules from their websites. I record the conference name, conference location and conference date. Then, for each conference, I hand-collect investment pitches from three online sources: conferences’ websites, investment blogs and document-sharing websites. Lastly, for each stock pitch, I apply text mining techniques to extract the stock ticker, hedge fund name, speaker name and long/short flag.

The documents from the investment blogs and the document-sharing websites are provided mostly by investment analysts, financial journalists and MBA students interested in investing. These documents are comprehensive in recording investment pitches because they are intended for information sharing in these professional investment communities.

The sample consists of 341 long stock pitches and 41 short stock pitches from 29 conferences held from 2008 to 2013. Table 2.1 reports the number of conferences, funds and pitches by year. One limitation of the short time period is that I can assess stock performance only in an upward market. Investment conferences usually happen in late spring or early fall. At each conference, about eight hedge fund managers pitch specific stocks and others talk about macro views and investment philosophies. Although a few funds pitched multiple times in the sample, they pitched different stocks at different times and most of the pitches have unique stocks.

Table 2.1: Number of Investment Conferences and Stock Pitches

<table>
<thead>
<tr>
<th>Year</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conferences</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Funds</td>
<td>16</td>
<td>24</td>
<td>19</td>
<td>34</td>
<td>36</td>
<td>62</td>
</tr>
<tr>
<td>Stock Pitches</td>
<td>28</td>
<td>51</td>
<td>44</td>
<td>76</td>
<td>70</td>
<td>116</td>
</tr>
<tr>
<td>Long Pitches</td>
<td>26</td>
<td>43</td>
<td>36</td>
<td>69</td>
<td>62</td>
<td>105</td>
</tr>
<tr>
<td>Short Pitches</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2.2 reports summary statistics on the pitched stocks and the hedge funds that pitched at investment conferences. The average size of the hedge funds that pitch stocks at conferences is $15.7 billion. On average, they hold 109 names and have existed for 12.8 years. Close to 80% of them are hedge funds and most are located in the New York and Greenwich, CT areas. Note that the size of a fund’s 13F holdings does not represent its true assets managements because its book can be levered multiple times and some holdings, such as confidential filings or non-equity positions, might not be reported. Pitched stocks in general are larger than non-pitched stocks and have higher valuation multiples. The appendix provides more detailed information on these conferences and funds. Table B.1 provides a list of the collected investment conferences. Section B.1 provides a description of one conference and Figure B.1 displays a screenshot of its registration page. Table B.2 provides detailed information for a selected list of the hedge funds that pitched most frequently at these investment conferences.

I obtain stock returns from CRSP and company fundamental data from Compustat. I focus on common stocks with share codes 10 and 11 in CRSP. I match them to the stock pitch data based on the point-in-time stock ticker and the pitch date. To calculate the risk-adjusted return, I obtain the Fama-French factors and momentum returns from Ken French’s website and the DGTW returns from Russ Wermers’ website. I also extend the DGTW returns for the post-2013 period based on the procedure in Daniel et al. (1997).

I then link the investment pitches to Thomson Reuters institution holdings database on WRDS. The stock holdings data is based on publicly available 13F forms. Investment managers with over $100 million under management are required to file 13F with the SEC on a quarterly basis. To link hedge funds at investment conferences to the stock holdings data, I apply a fuzzy string match algorithm to the fund names based on Levenshtein distance and manually validate the matches. I focus on hedge funds’ long pitches and long books because funds are only required to report long positions in 13Fs, and short pitches are relatively rare, numbering about 20 in my sample. I then filter for hedge funds running concentrated portfolios with numbers of holdings between 10 and 100.
### Table 2.2: Characteristics of Funds that Pitched and Pitched Stocks

These tables report the summary statistics on characteristics of funds that pitched at conferences and the stocks they pitched. Panel (a) reports for each fund that pitched at conferences the dollar amount of 13F holdings, number of holdings, number of years in active status, whether it is a hedge fund, and whether it is headquartered in the New York and Greenwich, CT metropolitan area. Panel (b) reports for the pitched stocks market capitalization, book-to-market ratio, debt-to-assets ratio, dividend yield, gross margin, and ROE.

(a) Funds that Pitched at Conferences

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Sd</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>13F Amt ($MM)</td>
<td>15,568.5</td>
<td>79,196.3</td>
<td>214.2</td>
<td>1,599.9</td>
<td>6,686.4</td>
</tr>
<tr>
<td>N(Holdings)</td>
<td>109.4</td>
<td>299.0</td>
<td>16.0</td>
<td>30.0</td>
<td>64.5</td>
</tr>
<tr>
<td>Life (Yr)</td>
<td>12.8</td>
<td>8.2</td>
<td>6.3</td>
<td>11.0</td>
<td>17.0</td>
</tr>
<tr>
<td>Is Hedge Fund</td>
<td>76.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is NY/Greenwich</td>
<td>71.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Pitched Stocks (Long)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Sd</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Cap ($MM) Pitched</td>
<td>$33,192</td>
<td>$71,270</td>
<td>$2,971</td>
<td>$8,827</td>
<td>$26,594</td>
</tr>
<tr>
<td></td>
<td>Not Pitched</td>
<td>$11,881</td>
<td>$30,073</td>
<td>$719</td>
<td>$2,836</td>
</tr>
<tr>
<td>Book/Market Pitched</td>
<td>0.48</td>
<td>0.36</td>
<td>0.24</td>
<td>0.38</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Not Pitched</td>
<td>0.55</td>
<td>0.54</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>Debt/Assets Pitched</td>
<td>0.66</td>
<td>0.27</td>
<td>0.50</td>
<td>0.65</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Not Pitched</td>
<td>0.59</td>
<td>0.29</td>
<td>0.41</td>
<td>0.58</td>
</tr>
<tr>
<td>Dividend Yield Pitched</td>
<td>1.79%</td>
<td>1.30%</td>
<td>0.96%</td>
<td>1.45%</td>
<td>2.18%</td>
</tr>
<tr>
<td></td>
<td>Not Pitched</td>
<td>2.16%</td>
<td>1.82%</td>
<td>1.15%</td>
<td>1.81%</td>
</tr>
<tr>
<td>Gross Margin Pitched</td>
<td>46.80%</td>
<td>23.72%</td>
<td>27.51%</td>
<td>44.73%</td>
<td>66.02%</td>
</tr>
<tr>
<td></td>
<td>Not Pitched</td>
<td>45.03%</td>
<td>24.23%</td>
<td>25.66%</td>
<td>41.47%</td>
</tr>
<tr>
<td>ROE Pitched</td>
<td>8.60%</td>
<td>81.51%</td>
<td>4.80%</td>
<td>10.72%</td>
<td>18.95%</td>
</tr>
<tr>
<td></td>
<td>Not Pitched</td>
<td>4.70%</td>
<td>139.15%</td>
<td>0.55%</td>
<td>8.48%</td>
</tr>
</tbody>
</table>

I assume that the disclosed stock holdings are representative of the long side of a hedge fund’s portfolio. The 13F holdings are not necessarily the entire long portfolio of the manager, for various reasons. Certain asset classes are not required to disclose, such as fixed income securities, cash holdings and non-US equities. Small holdings—under 10,000 shares or $200,000—need not disclose. Investment managers can also request confidential filings ([Agarwal et al., 2013b](#)). Therefore, I exclude from the sample funds whose filings are clearly not representative of their investment strategies or whose strategies are not primarily based on stock-picking. The resulting hypothetical long portfolio represents hedge funds’ beliefs and
convictions about the underlying stock positions and their investment universe.

To categorize shareholder ownership for a company, I use the FactSet Ownership dataset (LionShares) through WRDS. The dataset provides institutional ownership statistics by firm and the methodology is based on Ferreira and Matos (2008). Aggregate stock ownership is based on institutional categories such as banks, mutual funds and hedge funds. Table 2.3 reports summary statistics on stock ownership by investor category for the pitched stocks.

Table 2.3: Shareholder Composition of Pitched Stocks

This table reports the shareholder composition of the stocks pitched. Panel (a) shows stock ownership by investor category and whether it is active or passive. Panel (b) shows stock ownership by top investor types for institution investors, mutual funds and corporate insiders. Mutual funds are a subset of institution investors. Stock ownership is the percent of total shares outstanding held by a particular investor group.

(a) Stock Ownership by Investor Category

<table>
<thead>
<tr>
<th>Investor Category</th>
<th>Total</th>
<th>Active</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institution</td>
<td>79.13%</td>
<td>68.59%</td>
<td>10.99%</td>
</tr>
<tr>
<td>Institution-Mutual Fund</td>
<td>37.62%</td>
<td>30.10%</td>
<td>7.52%</td>
</tr>
<tr>
<td>Corporate Insider</td>
<td>9.94%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Stock Ownership by Top Types in Each Investor Category

<table>
<thead>
<tr>
<th>Institution Type</th>
<th>Mutual Fund Type</th>
<th>Insider Type</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment Adviser</td>
<td>42.08% Open-End Fund</td>
<td>26.72% Individual</td>
<td>3.99%</td>
</tr>
<tr>
<td>Mutual Fund Manager</td>
<td>18.70% Variable Annuity Fund</td>
<td>4.47% VC/PE</td>
<td>2.08%</td>
</tr>
<tr>
<td>Hedge Fund Manager</td>
<td>10.68% Exchange Traded Fund</td>
<td>2.64% Private Company</td>
<td>1.36%</td>
</tr>
<tr>
<td>Pension Fund Manager</td>
<td>2.47% Pension Fund</td>
<td>1.53% Public Company</td>
<td>0.64%</td>
</tr>
<tr>
<td>Broker</td>
<td>2.01% Offshore Fund</td>
<td>1.46% Subsidiary</td>
<td>0.55%</td>
</tr>
</tbody>
</table>

To measure fund raising activities, I collect SEC Form D filings. They cover fund raising through security offerings by both pooled investment funds and startups. I match funds based on phone numbers and entity names, including registered offshore funding vehicles. Hedge funds may use more than one funding vehicles to raise money and they may use different names for the funding vehicles. However, they use the same office phone number on these forms because they are the primary issuers. Some data matching issues arise if a fund uses a third-party Cayman entity to raise money. They sometimes provide services to more than
one fund client and have different phone numbers. In this case, I require name matches and
discard the rest. Figure B.2 shows the percentiles of money raised as a percentage of last
year’s assets from 2010 to 2015.

2.4 Stock Pitch Performances

In this section, I conduct event studies on investment pitches and test whether the pitched
stocks outperform market and non-pitched stocks. On one hand, funds have an incentive to
pitch good stocks. First, they want to use investment conferences to build a good reputation
that will help attract future investments. Second, they want to convince investors and correct
mispricings that they themselves cannot correct through arbitrages. On the other hand, funds
also have incentives to pitch bad stocks. They want create liquidity events and favorable
temporary price pressure for positions they want to exit. I first describe the event study
methodology and then present empirical results on the short-term market reaction to the
stock pitches and their long-term risk-adjusted performances.

2.4.1 Event Study Methodology

To analyze stock pitch performances, I conduct event studies on stock pitches using a
methodology similar to that of Binder (1998) and MacKinlay (1997). I define the pitch date
as \( t = 0 \) in event time; the time unit can be a trading day, month or quarter. For instance, if
the time unit is a day, \( t = -5 \) means 5 days before the pitch and \( t = 3 \) means 3 days after
the pitch. To calculate a stock’s cumulative return (\( CR \)), I sum up its daily returns:

\[
CR_{it} = \begin{cases} 
\sum_{\tau=0}^{t} R_{i\tau} & t \geq 0 \\
0 & t = -1 \\
-\sum_{\tau=t+1}^{-1} R_{i\tau} & t \leq -2 
\end{cases} 
\] (2.1)

where \( t \) is the event time, \( t = 0 \) is the pitch date, and \( R_{i\tau} \) is the raw return of stock \( i \) at event
time \( \tau \). This variable is normalized relative to the end of the trading day before the pitch
and is equal to 0 at \( t = -1 \) by construction. The cumulative return before the pitch has a
negative sign such that the whole return curve has a consistent direction. When analyzing
the returns of a group of stocks, I take the equal-weighted average daily return and then sum
them up using the same method.

I calculate the market-adjusted return as the excess return of a stock over the CRSP
market benchmark. I calculate the risk-adjusted abnormal return \((AR)\) using Fama-French 3
factors with momentum and DGTW stock characteristics. To calculate the 4-factor alpha, I
run the following time series factor regression:

\[
R_{it} = \alpha + \beta_1MKT_t + \beta_2SMB_t + \beta_3HML_t + \beta_4MOM_t + \epsilon_{it}
\]  

(2.2)

To calculate the DGTW-adjusted return, I assign a stock to one of the 5x5x5 DGTW portfolios
based on size, value and momentum and then subtract the portfolio’s benchmark return:

\[
AR^{DGTW}_{it} = R_{it} - BR_t(s_{it}, v_{it}, m_{it})
\]  

(2.3)

where \(s_{it}, v_{it}\) and \(m_{it}\) indicate stock \(i\)’s DGTW portfolio assignment based on size, value and
momentum, respectively, and \(BR_t(s, v, m)\) is the benchmark return for the DGTW \((s, v, m)\)
portfolio. To calculate the cumulative risk-adjusted return of a group of stocks, I first take the
equal-weighted average of their daily risk-adjusted returns, sum them up and normalize the
cumulative return relative to \(t = -1\), similar to the calculation of the cumulative raw return.

\[
CAR_{it} = \begin{cases} 
\sum_{\tau=0}^{t} AR_{i\tau} & t \geq 0 \\
0 & t = -1 \\
- \sum_{\tau=t+1}^{-1} AR_{i\tau} & t \leq -2 
\end{cases}
\]  

(2.4)

2.4.2 Short-Term Market Reaction

To understand how the market reacts to the stock pitches, I look at the returns and trading
volumes of the pitched stocks in a 10-day event window around the investment conferences. I
compare the returns of pitched stocks to the returns of non-pitched stocks and the market.
The non-pitched stocks are from the 13F holdings for the same quarter as the investment
conferences.
Figure 2.1 plots the cumulative daily returns around investment pitches. The market immediately reacts to the initial disclosure of the stock pitches. The returns of the pitched stocks diverge from the market immediately after the pitches—long pitches spike up and short pitches spike down. In contrast, the non-pitched stocks closely track the market throughout the event window. Furthermore, the return spread does not revert afterwards within the event window. Although it is not necessarily tradable, one can earn about 1% during the first two days after the pitches (including the pitch date) by buying long pitches and selling either non-pitched stocks or the market and about 2% by doing the opposite trade for short pitches. These results suggest that these investment conferences are closely followed by other investors and have high market impacts. Table 2.4 shows that there is a statistically and economically significant difference in returns between pitched stocks and non-pitched stocks.

Figure 2.1: Short-Term Market Reaction to Pitched Stocks

This figure compares the cumulative average raw returns of pitched stocks (the blue line for long pitches and the red line for short pitches) to those of non-pitched stocks (black dashed line) and the market (black solid line) in a 10-day event window around the stock pitches. Day 0 is the event date when the fund pitched the stock. The cumulative returns are normalized to day -1. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences. The market return is calculated based on the value-weighted CRSP market index.

Figure 2.2 plots the average daily trading volumes of pitched stocks—both long and short
pitches—around investment conferences. The trading volumes increase during the investment conference and stay heightened for another two days. This is in line with the return patterns. The increase in trading volume before investment conferences suggests that there might have been some information leakage about which stock might be pitched before the conferences. This is consistent with anecdotal evidence that many investors speculate on these stock pitches.

Figure 2.2: Trading Volumes of Pitched Stocks

This figure shows the average trading volumes of pitched stocks around the investment conferences. Day 0 is the event date when the funds pitched the stocks at the conferences. The daily trading volume is normalized by market capitalization and is calculated as $\frac{\text{SharesTraded}}{\text{SharesOutstanding}}$.

2.4.3 Long-Term Risk-Adjusted Returns

To analyze the risk-adjusted performances of pitched stocks in the long run, I widen the event window to two years and use various risk models to adjust the raw returns in event studies. Figure 2.3 plots the cumulative market-adjusted returns. Long pitches and non-pitched stocks outperform the market both before and after the conferences, whereas short pitches underperform the market after the conferences. The outperformance is larger for long pitches than for non-pitched stocks. Because short pitches are few and 13F holdings do not cover
short positions, I focus only on long pitches in the subsequent analyses and sections.

Figure 2.3: Market-Adjusted Returns of Pitched Stocks

This figure plots the long-term cumulative average market-adjusted returns for pitched stocks (the blue line for long pitches and the red line for short pitches) and non-pitched stocks (black line). The market-adjusted return is calculated as the excess return of a stock over the market return. Day 0 is the event date when the fund pitched the stock. The cumulative returns are normalized to day -1. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences.

Figure 2.4 plots the cumulative adjusted returns using common risk factors. Panel (a) adjusts the returns using Fama-French 3 factors plus momentum and Panel (b) the adjusts returns using DGTW stock characteristics. The results using common risk factors are broadly consistent with the results using the market-adjusted model—the pitched stocks tend to have larger outperformance than the non-pitched stocks. They further show that after controlling for common risk factors, the outperformance of pitched stocks is larger before the conferences than after the conferences. Furthermore, the outperformance of the pitched stocks becomes quite close to that of the non-pitched stocks and widens again after the conferences. However, after the conferences, the outperformance plateaus after another 9 months and even reverts slightly.

To formally test the difference in returns, I regress the cumulative return on a dummy
Figure 2.4: Risk-Adjusted Returns of Pitched Stocks

These graphs show long-term cumulative average risk-adjusted returns for pitched stocks (blue line) and non-pitched stocks (black line). Panel (a) adjusts the returns using Fama-French 3 factors plus momentum and Panel (b) adjusts the returns using DGTW stock characteristics. Day 0 is the event date when the fund pitched the stock. The cumulative returns are normalized to day -1. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences.
variable indicating whether a stock holding is pitched at a conference:

\[ R_{ij,t\rightarrow t+k} = \alpha + \delta \textit{IsPitched}_{ijt} + \eta_t + \epsilon \]  

(2.5)

where \( R_{ij,t\rightarrow t+k} \) is stock i’s cumulative return between the month of the pitch \( t \) and month \( t + k \), \( \textit{IsPitched}_{ijt} \) indicates whether stock i is pitched by fund j, and \( \eta_t \) is time fixed effects. I cluster standard errors by time to account for the correlation between stocks from the same conference. Table 2.4 reports the results for market-adjusted returns and Table 2.5 the results for risk-adjusted returns.

Table 2.6 reports the factor tilts of pitched and non-pitched stocks. Panel (a) uses time-series factor regressions with Fama-French 3 factors and momentum. Panel (b) uses DGTW stock characteristics and reports the average index of the 5x5x5 DGTW portfolio that a stock is assigned to. Pitched stocks tend to have larger market capitalizations than non-pitched stocks. However, except for size, pitched stocks have similar factor tilts as non-pitched stocks in the hedge funds’ portfolios. Factor regressions show that both have a market beta slightly above 1 and load positively on value and negatively on momentum. Moreover, pitched stocks have higher risk-adjusted returns than non-pitched stocks before the conferences.

A calendar-time portfolio that buys pitched stocks and sells non-pitched stocks produces economically significant alpha. Each position is held from 1 day after the pitch to 100 days after the pitch. When a stock is dropped or added, I rescale the remaining stocks such that they are equal-weighted and the portfolios weights sum to 100%. Figure 2.5 plots the simulated returns series of the calendar-time portfolios. During the backtest period from 2008 to 2014, the annualized average return is 14.2% for pitched stocks, 7.5% for non-pitched stocks and 6.7% for the long-short portfolio, and the annualized Sharpe ratio is 0.67. Table 2.7 reports the results of the Jensen’s alpha test for the calendar-time portfolios and shows that the pitched stocks and the long-short portfolio have statistically significant positive risk-adjusted returns.
Table 2.4: Market-Adjusted Returns of Pitched Stocks

These tables report the cumulative market-adjusted returns for pitched stocks and non-pitched stocks. The market-adjusted return is calculated as the excess return of a stock over the market return. Day 0 is the event date when the fund pitched the stock. The cumulative returns are normalized to day -1. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences. Delta is the estimated coefficient $\delta$ in Equation 2.5. Standard errors are reported in parentheses.

(a) Short-Term Return

<table>
<thead>
<tr>
<th>Days</th>
<th>Pitched (%)</th>
<th>Non-Pitched (%)</th>
<th>Delta (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>-0.03 (1.03)</td>
<td>-0.18 (0.36)</td>
<td>0.15 (0.69)</td>
</tr>
<tr>
<td>-4</td>
<td>-0.21 (0.96)</td>
<td>-0.20 (0.28)</td>
<td>0.00 (0.69)</td>
</tr>
<tr>
<td>-3</td>
<td>0.07 (0.31)</td>
<td>-0.13 (0.15)</td>
<td>0.20 (0.18)</td>
</tr>
<tr>
<td>-2</td>
<td>0.01 (0.38)</td>
<td>-0.11 (0.12)</td>
<td>0.11 (0.27)</td>
</tr>
<tr>
<td>0</td>
<td>0.36 (0.13)</td>
<td>-0.03 (0.04)</td>
<td>0.39 (0.12)</td>
</tr>
<tr>
<td>1</td>
<td>0.82 (0.12)</td>
<td>-0.09 (0.07)</td>
<td>0.91 (0.17)</td>
</tr>
<tr>
<td>2</td>
<td>1.01 (0.22)</td>
<td>-0.06 (0.04)</td>
<td>1.06 (0.22)</td>
</tr>
<tr>
<td>3</td>
<td>0.94 (0.21)</td>
<td>-0.25 (0.08)</td>
<td>1.19 (0.18)</td>
</tr>
<tr>
<td>4</td>
<td>1.39 (0.45)</td>
<td>-0.01 (0.16)</td>
<td>1.40 (0.30)</td>
</tr>
<tr>
<td>5</td>
<td>1.11 (0.31)</td>
<td>0.11 (0.14)</td>
<td>1.00 (0.20)</td>
</tr>
</tbody>
</table>

(b) Long-Term Return

<table>
<thead>
<tr>
<th>Days</th>
<th>Pitched (%)</th>
<th>Non-Pitched (%)</th>
<th>Delta (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-400</td>
<td>23.06 (3.11)</td>
<td>14.53 (1.08)</td>
<td>8.53 (2.48)</td>
</tr>
<tr>
<td>-200</td>
<td>8.78 (1.53)</td>
<td>5.49 (1.62)</td>
<td>3.30 (1.12)</td>
</tr>
<tr>
<td>-100</td>
<td>3.63 (1.82)</td>
<td>2.24 (1.51)</td>
<td>1.39 (1.36)</td>
</tr>
<tr>
<td>-75</td>
<td>2.00 (2.28)</td>
<td>1.07 (1.31)</td>
<td>0.92 (1.91)</td>
</tr>
<tr>
<td>-50</td>
<td>1.30 (1.63)</td>
<td>0.69 (1.05)</td>
<td>0.60 (1.17)</td>
</tr>
<tr>
<td>-25</td>
<td>0.50 (0.95)</td>
<td>-0.04 (0.71)</td>
<td>0.54 (0.52)</td>
</tr>
<tr>
<td>25</td>
<td>1.58 (0.88)</td>
<td>-0.35 (0.31)</td>
<td>1.93 (0.70)</td>
</tr>
<tr>
<td>50</td>
<td>2.63 (0.97)</td>
<td>-0.08 (0.31)</td>
<td>2.71 (0.97)</td>
</tr>
<tr>
<td>75</td>
<td>3.46 (1.11)</td>
<td>1.30 (0.62)</td>
<td>2.16 (1.40)</td>
</tr>
<tr>
<td>100</td>
<td>4.90 (0.89)</td>
<td>1.06 (0.51)</td>
<td>3.84 (1.16)</td>
</tr>
<tr>
<td>200</td>
<td>7.58 (2.21)</td>
<td>3.28 (1.99)</td>
<td>4.30 (1.42)</td>
</tr>
<tr>
<td>400</td>
<td>7.57 (4.12)</td>
<td>4.63 (3.34)</td>
<td>2.94 (3.83)</td>
</tr>
</tbody>
</table>

59
Table 2.5: Risk-Adjusted Returns of Pitched Stocks

These tables report the cumulative risk-adjusted returns for pitched stocks and non-pitched stocks. Panel (a) adjusts the returns using Fama-French 3 factors plus momentum and Panel (b) adjusts the returns using DGTW stock characteristics. Day 0 is the event date when the fund pitched the stock. The cumulative returns are normalized to day -1. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences. Delta is the estimated coefficient $\delta$ in Equation 2.3. Standard errors are reported in parentheses.

(a) Fama-French 3 Factors + MOM

<table>
<thead>
<tr>
<th>Days</th>
<th>Pitched (%)</th>
<th>Non-Pitched (%)</th>
<th>Delta (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-400</td>
<td>20.72 (3.61)</td>
<td>10.99 (2.11)</td>
<td>9.73 (1.70)</td>
</tr>
<tr>
<td>-200</td>
<td>8.54 (0.93)</td>
<td>3.90 (1.59)</td>
<td>4.64 (1.25)</td>
</tr>
<tr>
<td>-100</td>
<td>4.19 (1.11)</td>
<td>1.68 (1.07)</td>
<td>2.51 (1.22)</td>
</tr>
<tr>
<td>-75</td>
<td>2.98 (1.73)</td>
<td>1.19 (0.97)</td>
<td>1.79 (1.59)</td>
</tr>
<tr>
<td>-50</td>
<td>2.40 (1.20)</td>
<td>0.72 (0.79)</td>
<td>1.67 (0.77)</td>
</tr>
<tr>
<td>-25</td>
<td>1.36 (0.54)</td>
<td>0.26 (0.46)</td>
<td>1.10 (0.47)</td>
</tr>
<tr>
<td>25</td>
<td>1.78 (0.79)</td>
<td>0.03 (0.24)</td>
<td>1.75 (0.62)</td>
</tr>
<tr>
<td>50</td>
<td>2.73 (0.98)</td>
<td>0.10 (0.21)</td>
<td>2.63 (1.03)</td>
</tr>
<tr>
<td>75</td>
<td>3.30 (1.23)</td>
<td>1.22 (0.38)</td>
<td>2.08 (1.29)</td>
</tr>
<tr>
<td>100</td>
<td>5.36 (1.07)</td>
<td>1.24 (0.39)</td>
<td>4.12 (1.07)</td>
</tr>
<tr>
<td>200</td>
<td>6.75 (1.58)</td>
<td>2.56 (1.39)</td>
<td>4.19 (1.52)</td>
</tr>
<tr>
<td>400</td>
<td>6.38 (3.91)</td>
<td>3.93 (1.97)</td>
<td>2.45 (3.89)</td>
</tr>
</tbody>
</table>

(b) DGTW Stock Characteristics

<table>
<thead>
<tr>
<th>Days</th>
<th>Pitched (%)</th>
<th>Non-Pitched (%)</th>
<th>Delta (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-400</td>
<td>19.23 (5.90)</td>
<td>10.68 (2.71)</td>
<td>8.55 (4.21)</td>
</tr>
<tr>
<td>-200</td>
<td>6.76 (2.58)</td>
<td>3.55 (1.46)</td>
<td>3.21 (2.05)</td>
</tr>
<tr>
<td>-100</td>
<td>3.01 (1.29)</td>
<td>1.39 (0.86)</td>
<td>1.62 (1.46)</td>
</tr>
<tr>
<td>-75</td>
<td>2.35 (1.46)</td>
<td>0.92 (0.73)</td>
<td>1.43 (1.51)</td>
</tr>
<tr>
<td>-50</td>
<td>1.39 (0.91)</td>
<td>0.49 (0.53)</td>
<td>0.90 (0.90)</td>
</tr>
<tr>
<td>-25</td>
<td>0.86 (0.55)</td>
<td>0.22 (0.39)</td>
<td>0.64 (0.55)</td>
</tr>
<tr>
<td>25</td>
<td>1.99 (0.66)</td>
<td>-0.09 (0.32)</td>
<td>2.09 (0.54)</td>
</tr>
<tr>
<td>50</td>
<td>3.37 (0.96)</td>
<td>0.12 (0.28)</td>
<td>3.25 (0.80)</td>
</tr>
<tr>
<td>75</td>
<td>3.79 (1.09)</td>
<td>0.89 (0.72)</td>
<td>2.90 (1.02)</td>
</tr>
<tr>
<td>100</td>
<td>4.76 (1.55)</td>
<td>0.76 (0.84)</td>
<td>4.00 (1.41)</td>
</tr>
<tr>
<td>200</td>
<td>6.09 (2.63)</td>
<td>2.29 (1.63)</td>
<td>3.80 (2.11)</td>
</tr>
<tr>
<td>400</td>
<td>7.02 (3.93)</td>
<td>4.06 (1.72)</td>
<td>2.96 (3.87)</td>
</tr>
</tbody>
</table>

60
Table 2.6: Factor Tilts of Pitched and Non-Pitched Stocks

These tables report the factor tilts of pitched stocks and non-pitched stocks. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences. Panel (a) uses time-series factor regressions with Fama-French 3 factors and momentum. Panel (b) uses DGTW stock characteristics. A stock’s factor tilt is determined by the DGTW portfolio it is assigned to. For estimated factor loadings and intercepts, cross sectional standard deviations are reported under the cross-sectional average. The intercept is annualized.

(a) Fama-French Factor Regression

<table>
<thead>
<tr>
<th></th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>MOM</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitched</td>
<td>1.049</td>
<td>0.192</td>
<td>0.075</td>
<td>-0.092</td>
<td>8.39%</td>
</tr>
<tr>
<td></td>
<td>(0.325)</td>
<td>(0.408)</td>
<td>(0.609)</td>
<td>(0.282)</td>
<td>(19.11%)</td>
</tr>
<tr>
<td>Not Pitched</td>
<td>1.030</td>
<td>0.313</td>
<td>0.070</td>
<td>-0.102</td>
<td>4.79%</td>
</tr>
<tr>
<td></td>
<td>(0.129)</td>
<td>(0.198)</td>
<td>(0.244)</td>
<td>(0.111)</td>
<td>(5.65%)</td>
</tr>
<tr>
<td>Difference</td>
<td>0.019</td>
<td>-0.122</td>
<td>0.004</td>
<td>0.010</td>
<td>3.60%</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.037)</td>
<td>(0.053)</td>
<td>(0.025)</td>
<td>(1.62%)</td>
</tr>
</tbody>
</table>

(b) DGTW Risk-Adjustment

<table>
<thead>
<tr>
<th></th>
<th>SIZE</th>
<th>BM</th>
<th>MOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitched</td>
<td>3.877</td>
<td>3.005</td>
<td>2.943</td>
</tr>
<tr>
<td></td>
<td>(1.244)</td>
<td>(1.472)</td>
<td>(1.410)</td>
</tr>
<tr>
<td>Not Pitched</td>
<td>3.303</td>
<td>2.996</td>
<td>3.030</td>
</tr>
<tr>
<td></td>
<td>(1.443)</td>
<td>(1.409)</td>
<td>(1.402)</td>
</tr>
<tr>
<td>Difference</td>
<td>0.574</td>
<td>0.009</td>
<td>-0.087</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.102)</td>
<td>(0.097)</td>
</tr>
</tbody>
</table>
These figures show the cumulative average returns of calendar-time portfolios of pitched and non-pitched stocks. Panel (a) plots the returns separately, one for a portfolio of pitched stocks (blue line) and the other for a portfolio of non-pitched stocks (green line). Panel (b) plots the return of a dollar-neutral portfolio that longs pitched stocks and shorts non-pitched stocks. These portfolios enter positions the day after the stocks are pitched and hold them until 100 days after the pitches. Non-pitched stocks include stocks held by the funds in the same quarter in which they pitched the stocks at conferences.
Table 2.7: Jensen’s Alpha Test for Calendar-Time Portfolios

This table reports the results of the daily Jensen’s alpha test for the calendar-time portfolios using Fama-French 3 factors and momentum. Column 1 reports the factor tilts and annualized alpha for a standalone long portfolio of pitched stocks, Column 2 for a standalone short portfolio of non-pitched stocks and Column 3 for the long-short portfolio. Newey-West standard errors are reported in parentheses under coefficients. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Pitched</th>
<th>Non-Pitched</th>
<th>Long-Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>1.093***</td>
<td>1.034***</td>
<td>0.0583**</td>
</tr>
<tr>
<td></td>
<td>(0.0302)</td>
<td>(0.00946)</td>
<td>(0.0266)</td>
</tr>
<tr>
<td>SMB</td>
<td>0.149**</td>
<td>0.329***</td>
<td>-0.180***</td>
</tr>
<tr>
<td></td>
<td>(0.0697)</td>
<td>(0.0222)</td>
<td>(0.0659)</td>
</tr>
<tr>
<td>HML</td>
<td>-0.296***</td>
<td>-0.0720*</td>
<td>-0.224***</td>
</tr>
<tr>
<td></td>
<td>(0.0998)</td>
<td>(0.0392)</td>
<td>(0.0811)</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.274***</td>
<td>-0.156***</td>
<td>-0.118**</td>
</tr>
<tr>
<td></td>
<td>(0.0816)</td>
<td>(0.0380)</td>
<td>(0.0574)</td>
</tr>
<tr>
<td>Ann. Alpha</td>
<td>9.93%**</td>
<td>1.69%</td>
<td>8.24%*</td>
</tr>
<tr>
<td></td>
<td>(4.61%)</td>
<td>(2.42%)</td>
<td>(4.38%)</td>
</tr>
<tr>
<td>N</td>
<td>1310</td>
<td>1310</td>
<td>1310</td>
</tr>
<tr>
<td>R2</td>
<td>0.793</td>
<td>0.964</td>
<td>0.050</td>
</tr>
</tbody>
</table>

2.5 Investor Behaviors

This section presents empirical results on investors’ trading behavior around stock pitches at investment conferences. First, I investigate the trading behaviors of the funds that pitch stocks based on 13F filings. I then investigate the trade-offs the funds face when choosing which stock to pitch at a conference. Specifically, I test profit-taking, investment opportunities and fund raising incentives and find them to be consistent with the selling behavior. Finally, I investigate the trading behaviors of other investors to further understand how different investors react to pitched stocks and provide suggestive evidence on how their behaviors could contribute to the return patterns after the pitches.
2.5.1 Holdings by Funds That Pitched Stocks

If a pitched stock is a fund manager’s true “best idea,” he can be expected to hold onto it for an extended period of time after the pitch at an investment conference. If he does not, this suggests that he might face other incentives not to pitch his “best idea.” To answer the question of whether the fund manager truly believe the pitched stock is his “best idea,” I examine the funds’ 13F holdings before and after the pitches and reconstruct their hypothetical portfolios. I test whether the funds decrease their holdings of the pitched stocks after the conferences.

To track its trading behavior, I first reconstruct a fund’s long portfolios using 13F filings. I calculate the portfolio weight of the pitched stock before and after the pitch at quarterly frequency:

\[ w_{ijq} = \frac{\text{SharesHeld}_{ijq} \times P_{ijq}}{\sum_{k \in \text{Portfolio}_{jq}} \text{SharesHeld}_{kjq} \times P_{kjq}} \]  

(2.6)

where \( q \) is the event time measured in quarters after the pitch, \( w_{ijq} \) is the portfolio weight of the pitched stock \( i \) in fund \( j \)’s portfolio in quarter \( q \), \( \text{SharesHeld}_{ijq} \) is the adjusted number of shares held by fund \( j \), \( P_{ijq} \) is the adjusted price of stock \( j \) in quarter \( q \), and the denominator is the total market value of fund \( j \)’s 13F holdings in quarter \( q \). Although many equity hedge funds run market-neutral portfolios or long-biased books, their long books usually represent a large chunk of their assets allocated to the strategies except for dedicated shorts-sellers. Therefore, a stock’s portfolio weight in the long book proxies well for the level of conviction the fund has in the stock.

Figure 2.6 plots the average portfolio weight of the pitched stocks held by the funds that pitched the stocks before and after the pitches. It shows that hedge funds usually spend 2 years building up the positions of their pitched stocks before disclosing them at conferences. The position size peaks in the same quarter as the pitch. However, the portfolio weights of the pitched stocks immediately start to decrease from the first quarter after the pitch, like those of non-pitched stocks. Although pitched stocks are among the funds’ top holdings, they do not hold onto the pitched stocks any longer than they do to other stocks. This suggests that the pitched stocks were once their “best ideas” but are unlikely to be so any longer.
Figure 2.7 uses alternative measures to track hedge funds’ trading behaviors and finds similar patterns. Panel (a) normalizes the portfolio weights by the average size of a fund’s individual holding. Panel (b) uses the percentile rank of an individual holding based on size. A higher percentile rank means a larger position in the portfolio.

2.5.2 Why Funds Pitch Stocks

The previous subsections show that funds do not hold onto their stock pitches any longer than to non-pitched stocks and that they start selling them soon after pitching them at the conferences. However, at the same time, these pitched stocks do not suffer negative risk-adjusted return after the pitches and are not necessarily bad stocks. In this subsection, I analyze the trade-offs funds face when choosing which stocks to pitch and test three possible motives for such behavior. First, funds want to take profit and pitch stocks with the most paper gains. Second, funds care about their reputation and pitching stocks at conferences
Figure 2.7: Holdings of Pitched Stocks using Alternative Measures

These figures construct the quarterly holdings of pitched stocks (blue line) and non-pitched stocks (green line) using alternative measures. Quarter 0 is the quarter in which the funds pitched the stocks at the conferences. Panel (a) normalizes a stock’s portfolio weight by the average portfolio weight of a fund. By construction, the value is close to 1 for non-pitched stocks. Panel (b) uses the percentiles of a stock’s portfolio weight within a fund’s portfolio.
helps them raise money. Third, funds want to create room in their portfolios for better investment opportunities and use conferences to create market demand and help decrease their top holdings.

First, I examine how hedge funds choose which stocks in their portfolios to pitch based on holdings and prior returns. Table 2.6 shows that the pitched stocks do not seem significantly different from the non-pitched stocks in the portfolio based on their factor tiles. The speaker lineups are usually confirmed and posted on the conference websites in advance of the conferences. However, the specific pitches are not known until the actual time of the speech at the conference. Therefore, the hedge fund that is pitching has some flexibility in the choice of stock pitches.

To analyze how a hedge fund chooses stock pitches at investment conferences, I look at the levels and changes of the holdings of portfolio stocks before the pitch and test their relationships with whether a given stock is pitched at the conference. Specifically, I regress the dummy variable $IsPitched$ on variables including holdings $H_i$, change in holdings $\Delta H_i$ and the returns of various horizons $R_j$ separately:

$$
IsPitched_{ijt} = \alpha + \beta_0 H_{ij,t-\delta} + \beta_1 \Delta H_{ij,t-\delta} + \beta_2 R_{j,t-v} + \epsilon
$$

(2.7)

where $i$ and $j$ indicate the hedge fund and the stock respectively, $\delta$ refers to the number of quarters before the quarter in which the fund manager pitched at the conference; $v$ refers to the number of trading days before the pitch date and $R_{j,t-v} \rightarrow t$ is the cumulative returns from $t - v$ and $t$ for stock $j$.

Table 2.8 reports the regression results on stock pitch choices and shows that funds pitch stocks that have larger holdings and larger paper gains and stocks of which they have recently increased their holdings. Column 1 suggests that hedge funds tend to pitch stocks that have large positions in their portfolios. A 10% increase in holdings means that a stock is 9% is more likely to be pitched. A stocks with a large position will have a large market impact if the hedge fund liquidates it. Given the ex-post action of hedge funds pitching stocks, this is consistent with the hypothesis that they want to drive market demand for the stocks that
they want to liquidate. Although Column 2 reports that prior return does not predict the chance that a stock will be pitched, Column 3 suggests that funds pitch stocks to take profit. Column 3 reports the coefficient for the interaction term between the size of the holding and the past return. The positive and statistically significant estimate is consistent with the profit taking motive. A 10% larger paper gain means a stock is 2.24% more likely to be pitched. Column 4 suggests that higher portfolio holdings four quarters before the pitch positively predict that a stock within the portfolio will be pitched. Column 5 suggests that an increase in portfolio holdings from four quarters ago also positively predicts that a stock being will be pitched.

Table 2.8: Determinants of Stock Pitch Decisions

This table analyzes the factors that affect a fund’s decision about which stock to pitch. The dependent variable is a dummy variable indicating whether a stock in a portfolio is pitched. PortWgt is the portfolio weight of the pitched stocks in the quarter before the pitch and PortWgt4Q Ago is the portfolio weight 1 year before the pitch. ChgPortWgt4Q Ago is the change in the portfolio weight of the pitched stocks from 1 year before the pitch to the quarter before the pitch. PriorARet12M is the risk-adjusted return during the 1-year period before the pitch. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Pitched</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)  (2)  (3)  (4)  (5)</td>
</tr>
<tr>
<td>PortWgt</td>
<td>0.900*** 0.876*** 0.849***</td>
</tr>
<tr>
<td></td>
<td>(0.051) (0.053) (0.054)</td>
</tr>
<tr>
<td>Prior ARet 12M</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td>PortWgt x Prior ARet 12M</td>
<td>0.224**</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
</tr>
<tr>
<td>PortWgt 4Q Ago</td>
<td>0.387***</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
</tr>
<tr>
<td>ChgPortWgt 4Q Ago</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.003 0.003 0.003 0.022*** 0.022***</td>
</tr>
<tr>
<td></td>
<td>(0.003) (0.003) (0.003) (0.002) (0.002)</td>
</tr>
<tr>
<td>N</td>
<td>5,834 5,528 5,528 5,539 5,539</td>
</tr>
<tr>
<td>R2</td>
<td>0.05 0.048 0.049 0.009 0.02</td>
</tr>
</tbody>
</table>

Hedge funds also rely on their good reputation to raise LP investment. This puts constraints on whether they can pitch bad stocks that they want to liquidate, because doing so will tarnish their reputation and credibility. As a result, funds still pitch stocks that they want to
liquidate to drive market demand. However, they choose stocks with smaller upsides rather than outright bad stocks. To test this hypothesis, I collect fund raising activities from Form D filings and analyze how pitching stocks at investment conferences is related to fund raising outcomes. I regress the amount of money a fund raised on whether it pitched a stock in the prior year and how the pitched stock performed. Specifically, I run the following regression

$$PctMoneyRaised_{i,t+1} = \beta_0 Pitched_{it} + \beta_1 Pitched_{it} \times PitchRet_{it} + \delta_t + v_i + \epsilon_{it}$$  \ (2.8)

where $PctMoneyRaised_{i,t+1}$ is the money raised as a percentage of assets by fund $i$ in year $t + 1$, $Pitched_{it}$ is a dummy variable indicating whether fund $i$ pitches at an investment conference in year $t$, $PitchRet_{it}$ is the risk-adjusted returns of stocks pitched by fund $i$ in year $t$, $\delta_t$ absorbs the year fixed effects and $v_i$ absorbs the fund fixed effects. $\beta_0$ measures how much more money pitching a stock at a conference help raise. $\beta_1$ measures how the performance of pitched stocks affects fund raising.

Table 2.9 reports the regression results for fund raising activities. Column 1 shows that pitching at investment conferences helps funds raise 3.4% more of their assets. This suggests that appearing at investment conferences helps promote awareness of funds and in turn helps their future fund raising. Columns 2 to 4 suggest that while a 1-month outperformance does not help fund raising, 3-month and 6-month outperformances help funds raise more money. Columns 5 to 7 suggests that the performances of non-pitched stocks with similar sizes do not have statistically significant effects on the money raised. I control for position size because pitched stocks usually have larger portfolio weights and thus have larger impacts on portfolio returns. Column 8 is a placebo test. It shows that pitching at investment conferences does not help fund raising in the same year before the reputation effect. However, this relationship might not be causal, because hidden factors can affect a fund’s decision to pitch and its future money raising abilities. For example, a fund manager might be experiencing a strong performance streak. The placebo test in Column 8 alleviates some of these concerns.

Another possible reason that a hedge fund may want to pitch a good stock but sell it afterwards is that it has better investment opportunities. Pitched stocks are good ideas but
Table 2.9: Pitch Performances and Fund Raising Activities

This table analyzes how the performances of pitched and non-pitched stocks affect funds' money raising activities. The dependent variable is the money raised in a year as a percentage of assets in the prior year. Columns 1 to 7 regress the money raised in the year after the pitches. Column 8 is the placebo test and regresses the money raised in the same year as the pitches. Pitched is a dummy variable indicating whether a stock is pitched. The next three variables are risk-adjusted returns for the pitched stocks 1 month, 3 months and 6 months after the pitches. The last three variables are for the non-pitched stocks. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th>Pct. Chg. in Total Amount of Money Raised</th>
<th>Subsequent Yr</th>
<th>Current Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)  (2)  (3)  (4)  (5)  (6)  (7)  (8)</td>
<td></td>
</tr>
<tr>
<td>Pitched</td>
<td>0.032* 0.034* 0.027 0.042** 0.033* 0.035* 0.034* -0.0004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018) (0.018) (0.018) (0.019) (0.018) (0.018) (0.018) (0.031)</td>
<td></td>
</tr>
<tr>
<td>Pitched Stocks:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARet 1M</td>
<td>0.976</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.074)</td>
<td></td>
</tr>
<tr>
<td>ARet 3M</td>
<td>2.807**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.148)</td>
<td></td>
</tr>
<tr>
<td>ARet 6M</td>
<td>2.447**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.005)</td>
<td></td>
</tr>
<tr>
<td>Non-Pitched Stocks:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARet 1M</td>
<td>3.682</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.032)</td>
<td></td>
</tr>
<tr>
<td>ARet 3M</td>
<td>-3.959</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.303)</td>
<td></td>
</tr>
<tr>
<td>ARet 6M</td>
<td>2.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.155)</td>
<td></td>
</tr>
<tr>
<td>Time &amp; Fund FE</td>
<td>Y  Y  Y  Y  Y  Y  Y  Y</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td>0.469 0.471 0.481 0.481 0.472 0.472 0.470 0.460</td>
<td></td>
</tr>
</tbody>
</table>

not the best ideas. The fund may want to decrease some holdings to build positions for its true best ideas. In fact, the pitched stocks usually have larger positions in the portfolios. To test this possibility, I compare performances between newly-bought stocks after the conferences and pitched stocks. If newly-bought stocks perform better than the pitched stocks that the funds sell, this suggests that the funds face this security selection trade-off in their portfolios.

To test this hypothesis, I regress the post-pitch returns on dummy variables indicating whether a stock holding is increased after the pitch. To separate the doubling-down effect, I further condition on the direction and magnitude of a passive change in holdings that is
caused by stock price but not by active trading. Specifically, I run the following regression:

\[
PostPitchRet_{ijt} = IsBuy_{ijt} \times X_{ijt} + OtherNonPitched_{ijt} + \delta_t + \epsilon_{ijt}
\]  

(2.9)

where PostPitchRet_{ijt} is the post-pitch return of stock j in fund i’s portfolio when fund i pitches at a conference in time t, IsBuy is a dummy variable indicating whether fund i bought more stock j after the pitch and OtherNonPitched_{ijt} is a catch-all dummy variables for the remaining non-pitched stocks in fund i’s portfolios at time t. The base case is that a stock is pitched. X_{ijt} is a dummy variable that further separates newly-bought non-pitched stocks.

Table 2.10 reports the regression results on post-pitch returns. The first 3 months are excluded because newly-bought stocks are determined based on 13F holdings at quarter 0 and quarter 1. Quarter 0 is the quarter in which the funds pitched the stocks at the conferences. Panel (a) shows that newly-bought stocks that have positive returns during the first 3 months after the pitches have 6.7% higher returns than pitched stocks from 3 months to 6 months after the pitches and 9.3% higher returns from 3 months to 12 months after the pitches. Both estimates are statistically significant. Doubled-down stocks—newly-bought stocks that have negative returns during the first 3 months after the pitches—have 1% higher returns than pitched stocks from 3 months and 6 months after the pitches and 1% lower returns from 3 months to 12 months after the pitches. Neither is, however, statistically significant. Other non-pitched stocks in the portfolio underperform relative to pitched stocks. Panel (b) shows that newly-bought stocks that have small absolute price changes have 4.5% higher returns from 3 months to 6 months after the pitches, but that the outperformance reverts afterwards.

2.5.3 Holdings by Other Investors

I analyze the trading behaviors of other investor groups based on 13F filings and FactSet Ownership data. I first aggregate stock holdings across funds for a certain investor category and calculate the holdings as a percentage of a stocks’ shares outstanding. I separate institution investors that have active funds into hedge funds, mutual funds and other investment advisers. For hedge funds, I exclude the holdings by the funds that pitched. I then apply a similar
This table tests whether funds pitch stocks because of better investment opportunities and compares the performances of pitched stocks to those of stocks the funds bought after the pitches. Columns 1 and 2 regress risk-adjusted returns from 3 months to 6 months after the pitches. Columns 3 and 4 regress risk-adjusted returns from 3 months and 12 months after the pitches. The first 3 months are excluded because newly-bought stocks are determined based on 13F holdings at quarter 0 and quarter 1. Quarter 0 is the quarter in which the funds pitched the stocks at the conferences. To separate the doubling-down phenomenon, Panel (a) and Panel (b) are conditional on the direction and magnitude of a passive change in holdings that is caused by stock price and not by active trading. In Panel (a), \( \text{IsBuy} \times \text{IsPosPassiveChg} \) indicates stocks that are bought after the pitches and have positive returns. \( \text{IsBuy} \times \text{IsNegPassiveChg} \) indicates stocks that are bought after the pitches but have negative returns, i.e., doubling down. \( \text{OtherNonPitched} \) indicates all other non-pitched stocks in the portfolio. The coefficients measure performance relative to pitched stocks. In Panel (b), \( \text{IsBuy} \times \text{IsSmallPassiveChg} \) indicates stocks that have small price changes after the pitches. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>ARet 3M-6M</th>
<th>ARet 3M-12M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Base Case: Pitched Stocks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel (a): Cond. on Passive Pos. Chg.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\text{IsBuy} \times \text{IsPosPassiveChg}</td>
<td>0.067***</td>
<td>0.093**</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>\text{IsBuy} \times \text{IsNegPassiveChg}</td>
<td>0.010</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>\text{OtherNonPitched}</td>
<td>-0.018</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.060)</td>
</tr>
<tr>
<td><strong>Panel (b): Cond. on Size of Passive Pos. Chg.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\text{IsBuy} \times \text{IsSmallPassiveChg}</td>
<td>0.045**</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>\text{OtherNonPitched}</td>
<td>-0.009</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>\text{Time FE}</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>\text{N}</td>
<td>6,705</td>
<td>6,705</td>
</tr>
<tr>
<td>\text{R2}</td>
<td>0.026</td>
<td>0.02</td>
</tr>
</tbody>
</table>

The event study method to calculate the cumulative change in holdings before and after the stock pitches for each investor category. I normalize the quarterly holdings relative to the quarter when the stock is pitched.

Figure 2.8 plots the trading behaviors of institutional investors in pitched stocks before and after investment conferences. Different investors exhibit different behaviors, suggesting that they have different risk preferences and investment strategies. Other hedge funds exhibit
trading patterns similar to those of the fund that pitched. This herding behavior suggests that they either run correlated strategies or share information with each other. However, other hedge funds continue to increase the portfolio weight of pitched stocks for another quarter after the pitches. It is plausible that they expect stock prices to go up temporarily after the pitches due to demand from unsophisticated investors. Investment advisers do not seem to pay attention to pitched stocks before the conferences. However, they start buying the pitched stocks and hold onto them after the pitches.

More importantly, the mutual funds’ behavior is the opposite to that of the hedge funds— their holdings decrease before the pitches but increase after the pitches. Because the majority of the outperformance of pitched stocks occurs before the pitch, mutual funds leave positive alphas on the table for hedge funds. There can be multiple reasons for this behavior. First, mutual funds might be passively providing liquidity to hedge funds. They do not know that pitched stocks have positive excess returns until the investment conferences take place. This attention story might be tested using the geographic proximity between mutual funds' headquarters and conference sites. Mutual funds close to conference sites are more likely to hear about stock pitches and trade sooner than others. Second, mutual funds and hedge funds have different risk specialization. Pitched stocks might have certain idiosyncratic risks, such as financial stresses and bankruptcies, that mutual funds do not want to bear and thus offload to hedge funds. Hedge funds earn risk premia for these risks and pitch them when the risks are resolved. However, pitched stocks are not likely to be distressed stocks because they have strong outperformance.

Lastly, I analyze the trading behaviors of retail investors and corporate insiders. I approximate retail investors’ holdings as the residual stock ownership after subtracting shares owned by institutions and insiders. I further adjust for short interests by adding them back to the shares outstanding and assume that the majority of the short positions are held by institution investors. That is a reasonable assumption because short selling is usually done by sophisticated investors. Retail holdings can then be calculated as

\[ W_{Retail,it} = 100\% + SI_{it} - W_{Inst,it} - W_{Insider,it} \]  

(2.10)
Figure 2.8: Trading Behaviors of Active Institutional Investors

This figure shows the cumulative average change in holdings by other active institutional investors in the pitched stocks. Other active institutional investors include other hedge funds (blue line), investment advisers (green line) and mutual funds (red line). Quarter 0 is the quarter in which the funds pitched the stocks at the conferences. The change in holdings is calculated as the quarterly change in the number of shares of a stock held by an institution divided by the stock’s adjusted number of shares outstanding. The cumulative change in holdings is normalized to quarter 0.

Figure 2.9 plots the trading behaviors of retail investors and corporate insiders. First, retail investors are likely to be uninformed and to pay less attention to these industry events. They are likely to provide liquidity to institutions, both hedge funds and mutual funds. They miss outperformances both before and after the conferences. However, this result might be affected by measurement errors on retail holdings. Second, corporate insiders show similar trading behaviors and consistently decrease their positions. However, they are less likely to be liquidity providers. They have private information about companies and are more likely to engage in strategic actions to time stock sales.
This figure shows the cumulative average change in holdings by retail investors (red line), corporate insiders (green line) and institutions (blue line) in the pitched stocks. Quarter 0 is the quarter in which the funds pitched the stocks at the conferences. The retail holdings are approximated as the residual stock ownership after controlling for institutional ownership, corporate insiders and short interests. The change in holdings is calculated as the quarter-over-quarter change in the number of shares of a stock held by an investor group divided by the stock’s adjusted number of shares outstanding. The cumulative change in holdings is normalized to quarter 0.

**2.6 Conclusion**

Using novel data on investment conferences, I examine the motives of hedge funds that pitch investment ideas at conferences and the market reaction to pitched stocks. Through event studies, I find that pitched stocks exhibit positive risk-adjusted returns both before and after the pitches. However, the majority of the outperformance occurs before the pitches. Outperformance after the pitches, moreover, is likely driven by inflows from other investors that follow these investment conferences. In spite of outperformance after the pitches, I find that hedge funds do not hold on their pitched stocks any longer than to non-pitched stocks. They instead start to decrease the portfolio weight of the pitched stocks after the investment conferences and after they have earned most of the positive alphas of the pitched stocks.

In addition, I examine the trading behaviors of other investors around investment pitches. I separate stock ownership by investor categories including hedge funds, mutual funds and
other investment advisors. One common pattern is that they all buy into pitched stocks after the conferences in the short term. However, their holdings patterns are fairly different at other time horizons. Other hedge funds behave similarly to the funds that pitch stocks—they buy before the pitches and sell afterwards. This suggests that hedge funds either run similar strategies or share investment ideas regularly. Mutual funds, however, trade the pitched stocks in a pattern opposite to that of hedge funds. They sell pitched stocks to hedge funds before the conferences when the pitched stocks earn significant positive alphas. After the conferences, the outperformance is smaller and they buy the pitched stock. The possible explanations for this opposite trading behaviors include passive liquidity provision and risk specialization.

Due to intensive competitions in the investment management industry, hedge funds are rapidly adopting new tactics to improve performance and generate alphas. I document an important new industry phenomenon and examine how hedge funds might strategically use it to their own advantage. There are two potential directions for future research on the implications of these new tactics, which include using investment conferences, for market efficiency. First, as the sample size of the investment conferences expands, it is important to understand how other investors learn about the various new techniques employed by hedge funds. Do investors, over time, learn which conferences and hedge funds genuinely produce good pitches and which produce bad pitches? How do their reactions to these investment conferences change? The second direction for future work is better understanding the interaction between hedge funds and mutual funds. Does the opposite trading behavior of mutual funds occur because they are less attentive than hedge funds? Or does it occur because there are certain idiosyncratic risks that mutual funds do not want to bear? Mutual funds represent the bulk of the assets in the industry whereas hedge funds are more active in information production and price discovery. It is therefore important to investigate how the two players interact.
Chapter 3

Detecting Anomalies: The Relevance and Power of Asset Pricing Tests

3.1 Introduction

Capital market anomalies fall into two categories. The first is when a security characteristic or signal, such as the ratio of a firm’s book to market value or its recent change in share price, predicts future returns but is otherwise not obviously related to risk. The second is when a signal, such as market beta, is theoretically and empirically connected to portfolio risk but nonetheless does not predict returns.

Anomalies attract both academic and practical interest. The focus of academic asset pricing is rationalizing seemingly anomalous predictability, by redefining or expanding what an investor considers to be risk. The focus of academic behavioral finance is uncovering a parsimonious set of psychological or sociological biases and institutional frictions that break the standard link between risk and return. And, the focus of practical investment management is delivering products to investors with the aim of delivering positive risk-adjusted returns.

---

1Co-authored with Malcolm Baker and Ryan Taliaferro
In this paper, our two-part goal is to examine the practical relevance and statistical power of standard asset pricing tests in identifying anomalies. The protagonist we have in mind is an investor who is seeking to make sensible security selection decisions using historical data as a guide. This is in the spirit of Brennan, Schwartz, and Lagnado (1997) and Campbell and Viceira (1999), who examine the consequences of return predictability for portfolio choice in partial equilibrium. While those papers focus on asset allocation across stocks and bonds, we focus on security selection, much like the classic analysis of Markowitz (1952) or more recently Garleanu and Pedersen (2013). We consider a security characteristic or signal to be a relevant anomaly if it has non-zero weight in our protagonist’s selection decision.

The sheer number of potential anomalies accumulated over decades of research demands a certain degree of simplification. Some recent attempts at grander simplification include Fama and French (2008, 2015, 2016) and Stambaugh, Yu, and Yuan (2012). Broadly speaking, there are two standard asset pricing tests. The first identifies candidate anomalies in the first category, looking at return predictability without considering risk. Fama and French (1992) is the canonical citation. The main empirical tool is cross-sectional return prediction using security-level signals and a pooled estimation, typically with the procedure of Fama and MacBeth (1973). We call a candidate that passes this first test a “score anomaly.” The second test determines whether a candidate anomaly adds to a set of existing return factors. Fama and French (1993) is the canonical citation. The main empirical tool is an intercept, or alpha, test in a time series return prediction using contemporaneous factor returns, typically with the procedure of Jensen (1968). We call a candidate that passes this second test a “factor anomaly.”

We begin with the question of relevance, and a simple approach, where our protagonist

\[2\] A given signal can be both a score anomaly and a factor anomaly. For example, profitability is a score anomaly in Table IV of Fama and French (2008) and a factor anomaly in Table 6 of Fama and French (2015). A signal can pass the first test as a score anomaly but fail the second. For example, the ratio of book to market value is a score anomaly in Table IV of Fama and French (2008) but it is not an independent factor anomaly in Table 6 of Fama and French (2015). And, a signal that fails the first test can in principle still be a factor anomaly. Members of this group do not predict stock-level returns but do hedge contemporaneous returns on other factors. Market beta is a leading example. It does not predict returns in the cross section, but it qualifies as a factor anomaly in the sense that it has a non-zero intercept in factor regressions.
investor is a Markowitz-style mean-variance optimizer. When our investor is risk averse and
faces no transaction costs or other frictions, only time-series tests and the factor anomalies
that emerge from these tests are relevant. In that sense, we can think of Fama and French
(1993) as catering to the needs of this type of investor. When our investor is risk neutral and
faces a simple form of transaction costs, constant across securities, then only cross-section
test and the score anomalies that emerge are relevant. Fama and French (1992) are catering
to this type of investor. For a risk-averse investor facing simple transaction costs, both sets of
anomalies are relevant, in the sense that our protagonist investor will not be satisfied using
only those anomalies that emerge from Fama and French (1993) time series tests. Those
anomalies left in the editing room of Fama and French (1992) are also relevant. The upshot
is that academic research might consider either test to be sufficient to establish a new and
relevant asset pricing anomaly.

A caveat is that this framework of relevance ignores practical differences across anomalies.
Investors face transaction costs that differ across securities, differ with capacity constraints,
and differ in multi-period portfolio choice. Unfortunately, the standard asset pricing tests
in their simple form are no longer relevant, except in special cases. In principle, they can
be replaced with intuitive modifications. Essentially, the simple returns and alphas from
cross-section and time series tests can be replaced with returns that are adjusted for execution
costs and our investor's specific level of assets under management. More ephemeral anomalies
whose conditional score variance is higher among securities with high transaction costs are,
all else equal, less relevant.

Having established relevance, we then turn the power of the two tests. The power of
both tests rises with the number of securities and the number of time periods and falls
with idiosyncratic security variance and factor variance in predictable ways. In addition
and importantly, factor tests have two additional terms. On the one hand, the power of a
factor test increases when other factors are useful in reducing the residual variance of the
test factor's time series returns. On the other hand, the power a factor falls as the in-sample
Sharpe ratio of the incumbent factors rises, as in Shanken (1992). In this sense, there can be
a higher chance of a false negative, or Type II error. Moreover, this tendency rises with the size of the incumbent model, because the in-sample Sharpe ratio is strictly increasing in the number of incumbent factors. Through this second channel, there is a lower natural limit to the number of new factor anomalies that can be identified. Intuitively, this is accentuated in small samples, where degrees of freedom are consumed in parameter estimation.

A final note is that the power of time series tests can be resurrected by shortening the return horizon. This is immediately obvious from the power formula that we derive. Power falls with the in-sample Sharpe ratio. The Sharpe ratio is itself mechanically increasing in the return horizon, because returns (the numerator of the ratio) rise linearly with horizon, while standard errors (the denominator) rise with the square root of horizon. This means asset pricing tests that rely on quarterly returns have much lower power than tests that rely on monthly horizons, and should be avoided if possible. And, asset pricing tests with daily return horizons resolve the problem of relative power: The daily Sharpe ratio is sufficiently small that the difference in power between time series and cross section tests becomes negligible. However, we stop short of recommending that the standard in asset pricing tests move from monthly to daily. Scholes and Williams (1977) and Liu and Strong (2008) suggest reasons why inferences might be biased in a daily analysis. The optimal approach trades off bias and power. In 30 portfolios from Ken French’s data library, analyzing 10-day return horizon essentially solves the problem of bias, suggesting that tests that use longer return horizons for these anomalies are needlessly sacrificing power.

We make connections to several related papers along the way. Garleanu and Pedersen (2013) consider the sort of partial equilibrium analysis that we do here, but they do not consider the econometric relevance of their optimal portfolios. Instead, they take the return generating process as given. Moreover, they use one simplifying assumption—that transaction costs are proportional to risk—to come to an elegant closed-form solution, while we consider a range of less elegant assumptions about transaction costs. Like us, Hoberg and Welch (2009) consider Fama and French (1992) style tests. Their focus, unlike ours, is the use in time series tests of optimized portfolios, whose returns are derived from cross-sectional regressions, versus
sorted portfolios, which are favored by Fama and French (1993). A large number of papers, including Fama (1998), Mclean and Pontiff (2016), Harvey, Liu, and Zhu (2015), Bailey and López de Prado (2014), and Novy-marx (2016) consider the issue of data mining and Type I error, in identifying anomalies that do not really exist. While this is a serious problem, our focus is instead on power and Type II error, in failing to identify legitimate anomalies, especially in time series tests. Just as we do, Loughran and Ritter (2000) and Ang, Liu, and Schwarz (2008) consider the power of asset pricing tests. Loughran and Ritter (2000) focus on the effects of weighting on the power of tests, both in aggregating firms at a point in time as well as in aggregating test statistics by market capitalization. Ang et al. (2008) focus on the use of aggregation in estimating factor loadings, while we focus on a comparison of cross-section and time series tests. We aggregate firm-level returns into factor returns, and we abstract from weighting schemes and the estimation of right-hand-side variables—including factor loadings—in the cross section, and instead focus on the lost power that comes from estimating covariances in time series tests. Consistent with our logic, Lewellen (2015) finds strong predictive power in a model that uses coefficients from more powerful cross-sectional estimation, while Simin (2008) and others find much less predictive power using less powerful time series estimation.

The paper proceeds as follows. Section 2 develops the investor’s security selection problem, and considers the relevance of the factor and score anomalies that emerge from standard asset pricing tests for portfolio choice. Section 3 derives the asymptotic and small-sample power of score and factor anomaly tests. Section 4 concludes.

3.2 The Relevance of Standard Asset Pricing Tests

There are three potential audiences for asset pricing tests. The first, rational asset pricing, considers anomalies to be a misspecification of the risks that are relevant to the representative investor. If a characteristic reliably predicts stock returns, it must be compensation for risk. The factor returns covary with some underlying state variable that drives investor utility. New anomalies, if they are deemed to be robust, are added to the set of known risk factors. With
the presumption that risk covariances are at the root of all seeming anomalies, rational asset
pricing has necessarily focused on time-series intercept tests. The second audience, behavioral
finance, considers anomalies to be examples of mispricing, driven by some combination of
less-than-fully- rational preferences and limits to arbitrage. The third audience, practitioners
in investment management, considers anomalies to be potential sources of risk-adjusted return
that can improve the welfare of their clients in partial equilibrium.

While our focus is on relevance to the third audience, it is worth saying a few words
qualitatively about the second. A candidate anomaly that passes the cross-section test but
not the time series test is arguably of academic relevance. In particular, the limits to arbitrage
and so-called intermediary asset pricing stipulates that shocks to arbitrageur or intermediary
capital can make the returns to seemingly unrelated anomalies correlated in the time series. If
the goal is to understand investor preferences or beliefs, then a characteristic that is uniquely
useful in explaining the cross-section of returns, but is spanned by other stronger anomalies
in the times series is nonetheless relevant for behavioral finance. A full exposition of this
argument is beyond the scope of this paper.

We apply the classic portfolio choice model of Markowitz (1952) to the problem facing the
third audience—an investor who cares about single period portfolio returns and variances.
Rather than attempting to characterize the general equilibrium in the spirit of Tobin (1958)
or Sharpe (1964) or Lintner (1965) that arises if all investors were rational and had these
preferences, we stay in partial equilibrium. We are interested in the case where active
portfolio management can deliver superior investment decisions for our non-representative
investor. This happens when our investor has a different view of risk and return from the
representative investor, either because of differences in preferences or beliefs. It is worth
noting that mean-variance portfolio choice is commonly used by practitioners. For example,
the portfolio construction software developed by MSCI, Axioma, and Northfield all use some
form of myopic mean-variance optimization, with constraints and non-linear transaction costs.

In this context, our definition of an anomaly is simple: It is a set of scores, for each security
in the opportunity set, that is relevant for our investor’s portfolio choice. If our investor can
safely ignore a set of scores, there is no anomaly. If our investor chooses to use this set of scores in his decision making, then there is an anomaly. We build intuition in three steps. The first is the classic case where our investor is risk averse and faces no trading frictions. The second is where our investor becomes risk neutral but faces a simple form of transaction costs that are constant across securities. And, the final step combines both risk aversion and transaction costs. These help establish the applicability of two standard asset pricing tests: the cross section Fama and MacBeth (1973) test popularized the Fama and French (1992) assessment of anomalies; and the time series Jensen (1968) alpha test popularized with the introduction of the Fama and French (1993) three factor portfolio.

We also consider three extensions to the basic model in Appendix C.1 that allow: for transaction costs that vary across securities; for varying levels of assets under management; and, for dynamic trading of the sort in Garleanu and Pedersen (2013). These extensions drive a wedge between gross and net returns that varies across anomalies and thereby suggest straightforward modifications to the two standard asset pricing tests that have the effect of netting out execution costs. In some cases, these adjustments are dependent on the level of assets under management, making the relevance of a particular anomaly context dependent, which is why we analyze them as extensions.

3.2.1 The Return Generating Process: Scores and Factor Returns

We suppose that returns for \( N \) securities follow a linear factor structure at discrete times \( t \in \{0, 1, \ldots, T\} \).

\[
\mathbf{r}_t = \mathbf{\Gamma}_t \mathbf{f}_t + \mathbf{\varepsilon}_t \\
\mathbf{f}_t \sim N(\mathbf{\mu}, \mathbf{\Sigma}) \\
\mathbf{\varepsilon}_t \sim N(0, \sigma^2 \mathbf{I}) \tag{3.1}
\]

The vector of individual security returns \( \mathbf{r} \), measured in excess of a risk-free rate of return, is governed by a matrix \( \mathbf{\Gamma} \) consisting of \( K < N \) row vectors of scores \( \gamma' \) that vary across securities \( i \) at a time \( t \), and a vector of normally distributed factor returns \( \mathbf{f} \) that vary over time but not across securities. The return of any security is equal to the sum product of its scores and corresponding factor returns plus a residual idiosyncratic return. The \( K \) factor returns can
be thought of as returns to portfolios of stocks that can be estimated with scores and observed returns, \( \hat{r}_t = (\Gamma'_t \Gamma_t)^{-1} \Gamma'_t r_t \). What we refer to as scores are sometimes called characteristics in the academic literature. The canonical Fama and French (2015) characteristics now include the ratio of book to market value, the firm’s market capitalization, the annual rate of growth in assets, and operating profitability scaled by assets, each transformed into buckets with common scores to limit the effect of extreme scores. To this, many researchers add stock price momentum, typically measured as the most recent annual return excluding the most recent month.

The assumption of normality in Equation 3.1 and a constant investment opportunity set, with no time subscripts on \( \mu \) or \( \Sigma \) aligns with the single-period mean-variance portfolio choice problem that we evaluate in the next subsection. Unlike essentially all of our other assumptions about the return generating process, these two come at the expense of generality. No doubt, the investment opportunity set changes over time and some factor returns are not normally distributed, and partial equilibrium investors likely care about timing factor returns and about the higher moments of their portfolio returns.

Below, we will use examples with two factors at a time to build intuition, and will often leave off the subscript \( t \) to simplify the notation:

\[
 r_i = \gamma_{a,i} f_a + \gamma_{b,i} f_b + \varepsilon_i
\]

We can map this into Equation 3.1:

\[
 r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \gamma_{a,1} & \gamma_{b,1} \\ \gamma_{a,2} & \gamma_{b,2} \\ \vdots & \vdots \\ \gamma_{a,N} & \gamma_{b,N} \end{bmatrix}, \quad f = \begin{bmatrix} f_a \\ f_b \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}
\]

To simplify the analysis, we often define the scores in a particular way, roughly in the spirit of Fama and French, to make them analogous to portfolio strategies. The first column of scores \( \gamma_a \) is equal to 1 for all firms. The rest are defined so as to sum to zero. This means that the first factor return \( f_a \) will be the average or market return on all securities. Under
the capital asset pricing model (CAPM), for example, where returns are governed by a single factor, the second column of scores \( \gamma_b \) is the demeaned CAPM beta \( f_i \), equal to the standard CAPM beta less 1, and the second factor return is equal to the first, \( f_a = f_b = r_m \), so that \( r = \beta r_m + \varepsilon \).

We are not otherwise specifying the distribution of the scores \( \Gamma \), so in principle this setup could accommodate sorted portfolios of the type in [Fama and French (1993)] or continuous variables of the type in [Fama and French (1992)]. There is a large literature on the relative merits of sorted portfolios versus more continuous or optimized weights. [Hoberg and Welch (2009)] argue that test portfolios and factor portfolios are better constructed via optimization than via sorting. [Daniel and Titman (1997)] and [Davis, Fama, and French (2000)] further consider whether sorting by characteristics and covariances helps to resolve the age-old question of whether a firm characteristic that is correlated with future returns is a risk factor or mispricing. We side-step both of these issues, but our analysis is closer in spirit to [Hoberg and Welch (2009)]. In addition to possible improvements in explanatory power that they document, we find linearly orthogonalized portfolios easier to work with analytically.

### 3.2.1.1 Exposures to Raw and Unit Factor Portfolios

A portfolio is defined by its vector of weights \( w \) on the available securities. It will be useful for us to characterize any portfolio’s exposure to two sets of factor portfolios, denoted by the matrices \( Q_{raw} \) and \( Q_{unit} \). We refer to the matrix \( Q_{raw} = \Gamma \) as the set of raw factor portfolios that convert the \( K \) sets of scores directly into portfolio weights. In general, as we have defined the scores above, the first raw factor portfolio is the market portfolio, and the subsequent raw factor portfolios are dollar neutral portfolios that tilt toward firm characteristics, such as \( f_i \). The weights in these raw portfolios can in principle be correlated in the cross section. (The canonical Fama and French firm characteristics are correlated, to some extent, in their final formulation.) We will also refer to the set of portfolios that are orthogonal, or cross-sectionally uncorrelated with all but one of the raw factor portfolios, with the remaining covariance designed to be exactly one. This is the set of unit factor portfolios that can be
obtained by cross sectional regression of returns on scores, \( Q_{\text{unit}} = \Gamma (\Gamma' \Gamma)^{-1} \).

Any portfolio \( w \) can be expressed as a linear combination of either raw or unit factor portfolios plus an orthogonal residual \( \eta \) using a multivariate regression of the portfolio weights on the matrix \( Q \). This results in a vector \( e(w) \) of \( K \) multivariate exposures to the factor portfolios \( Q \).

\[
w = Qe + \eta \Rightarrow e(w) = (Q'Q)^{-1} Q'w
\]

The function \( e_{\text{raw}} \) takes any set of portfolio weights \( w \) as an input and uses the matrix of raw factor portfolios \( Q = Q_{\text{raw}} \), while the function \( e_{\text{unit}} \) uses the matrix of unit portfolio weights \( Q = Q_{\text{unit}} \).

\[
e_{\text{raw}}(w) = (\Gamma'\Gamma)^{-1} \Gamma'w \quad e_{\text{unit}}(w) = \Gamma'w \tag{3.2}
\]

Note that the full set of unit factor exposures in the matrix \( E_{\text{unit}} \) of the set of unit factor portfolios \( Q_{\text{unit}} \) is, as designed, equal to the identity matrix. Each unit factor portfolio has exactly unit exposure to a single factor and zero to the rest. Meanwhile, the unit factor exposure of the raw factor portfolios has off diagonal elements that are not zero. Raw factor portfolios can in principle have different gross exposure and also have incidental correlations among each other.

\[
E_{\text{unit}}(Q_{\text{unit}}) = \Gamma'Q_{\text{unit}} = I \quad E_{\text{unit}}(Q_{\text{raw}}) = \Gamma'Q_{\text{raw}} = \Gamma'\Gamma
\]

### 3.2.1.2 Computing Portfolio Expected Return and Variance

A portfolio’s realized return can be characterized as the product of its unit factor exposures and the realized factor returns plus a residual return. Its expected return is the product of its unit factor exposures and the expected factor returns. Portfolio variance can be computed analogously.

\[
E(r'w) = E(f'\Gamma'w + \varepsilon'w) = E(f'e_{\text{unit}}(w) + \varepsilon'w) = \mu'e_{\text{unit}}(w)
\]

\[
\text{var}(r'w) = \text{var}(f'\Gamma'w + \varepsilon'w) = \text{var}(f'e_{\text{unit}}(w) + \varepsilon'w) = e_{\text{unit}}(w)' \Sigma e_{\text{unit}}(w) + \sigma^2 w'w
\]

With the return generating process in Equation 3.1, both expected return and variance can
be computed parsimoniously with the knowledge of factor exposures and the distributional properties of factor returns. This is because the residual variance will often be small for large and diversified portfolios as \( N \) becomes large, but the risk from the factor covariance matrix remains.

\[
\lim_{w'w \to 0} e_{\text{unit}}(w)' \Sigma e_{\text{unit}}(w) + \sigma^2 w'w = e_{\text{unit}}(w)' \Sigma e_{\text{unit}}(w)
\]

It is important to note that the factor returns themselves are not necessarily uncorrelated, even though they are returns to unit factor portfolios. They have unique exposure to a single set of scores. But, it is quite possible, and often true in US data, that two unit factor portfolio returns will be correlated with each other in the factor covariance matrix \( \Sigma \).

### 3.2.2 The Investor’s Security Selection Problem

We consider a single-period investor, without limits on leverage or short-selling constraints, who cares about mean and variance and knows the return generating process in Equation 3.1. In selecting a portfolio, our investor faces a simple form of quadratic transaction costs, which act as limits on position size:

\[
\max_w E (r'w) - \frac{\lambda}{2} \text{var} (r'w) - \frac{\theta}{2} w'w
\]

or, substituting the return generating process as simplified in Equation 3.3:

\[
\max_w -e_{\text{unit}}(w)' - \frac{\lambda}{2} (e_{\text{unit}}(w)' \Sigma e_{\text{unit}}(w) + \sigma^2 w'w) - \frac{\theta}{2} w'w
\]

and, substituting the definition of the unit exposures of a given portfolio shown in Equation 3.2:

\[
\max_w (\Gamma^-)'w - \frac{1}{2} w'(\lambda \Gamma \Sigma \Gamma' + I \left( \sigma^2 + \theta \right)) w
\]  

(3.4)

We acknowledge that all of these modeling assumptions come at the expense of generality. Most investors care about more than just mean and variance, they face sundry portfolio constraints, they have the ability to trade dynamically, and dynamic trade leads to more complicated effects of transaction costs, changing scores \( \Gamma_t \), and changes in the investment opportunity set \( \mu \) and \( \Sigma \). We analyze some of these as extensions below.
We consider two special cases of this objective function. The first is where the aversion to risk \( \lambda \) is equal to zero. In other words, the investor is risk neutral, but the non-zero transaction costs that he faces cause his optimization problem to remain convex. The second is where transaction costs \( \theta \) are zero, but there is aversion to risk \( \lambda \). This situation, where our investor can trade at no cost, is the classic problem in the academic literature on mean-variance optimization. Practically speaking, its outputs might apply approximately to an investor with low assets under management. For investors with higher levels of assets under management, variable costs of trade limit position sizes. Realistically, investors care about both execution costs and risk, but it is easier to build intuition for the two separate cases before we consider the general case.

In extensions in Appendix C.1 we first replace \( \theta \) with a vector of trading costs \( \theta \) that vary across securities. We then consider limits on leverage and the resulting effects of assets under management on investors facing the vector of cost parameters \( \theta \). And we finally consider simple dynamics in the spirit of Garleanu and Pedersen (2013) to capture the extra costs of factors whose value decays over time. These three extensions highlight several intuitive notions of execution costs at the level of factor portfolios, which can then be neatly characterized in closed form solutions and examples. Rather than using the standard asset pricing tests applied to gross returns, the two asset pricing tests must be applied to net of execution cost returns that depend on our investor’s specific circumstances.

3.2.3 Optimal Weights

The solution to the investor’s problem in Equation 3.4 involves a tradeoff between risk, return, and execution costs. The optimal weights on each security are a function of security scores \( \Gamma \), expected factor returns \( \mu \), the covariance of factor returns \( \Sigma \), transaction costs \( \theta \), the investor’s risk aversion \( \lambda \), and assets under management \( A \). At optimal portfolio weights, the marginal benefit of incremental weight in each security is equal its marginal cost in the
optimal portfolio:

\[ \Gamma \mu = \left( \lambda \Gamma \Sigma \Gamma' + I \left( \lambda \sigma^2 + \theta \right) \right) w^* \]

\[ \Rightarrow w^* = \left( \lambda \Gamma \Sigma \Gamma' + I \left( \lambda \sigma^2 + \theta \right) \right)^{-1} \Gamma \mu \]  

(3.5)

We analyze two special cases, when transaction costs are zero and when risk aversion is zero, and then proceed to the general case, over the next three subsections, before considering the case of transaction costs that vary across securities in Appendix C.1.

3.2.3.1 Risk Neutral, Constant Transaction Costs

First, we consider the simplest case, where \( \lambda \) is equal to zero and risk considerations are unimportant. Our investor is interested in maximizing returns net of transaction costs. Then, the optimal weights from Equation (3.5) simplify to:

\[ w_{tc}^* = \frac{1}{\theta} \Gamma \mu \]

Intuitively, the optimal weight for an individual security is high when it scores well on factors that have high expected returns. To get more visibility into the optimal weights, we can compute the exposure \( e \) of this portfolio \( w_{tc}^* \) to the raw and unit factor portfolios using Equation (3.2):

\[ e_{raw}(w_{tc}^*) = \frac{1}{\theta} \mu \]

\[ e_{unit}(w_{tc}^*) = \frac{1}{\theta} \Gamma \Gamma \mu \]  

(3.6)

The upshot is that our risk neutral investor's problem can be reduced constructing the raw factor portfolios \( Q_{raw} = \Gamma \), learning the magnitude of the expected factor returns \( \mu \), and using this vector as weights on the \( K \) raw factor portfolios, which are the columns of \( Q_{raw} \). If the second factor were to have a zero expected return \( \mu_b = 0 \), it can be ignored in the optimization problem, regardless of its risk properties. What is trivially absent is \( \Sigma \). If a factor has a positive expected return \( \mu_b > 0 \), but it has a zero alpha with respect to an existing factor \( c \), the standard logic of Fama and French (1993) says this is not a distinct anomaly. If factor \( b \) has a zero expected return \( \mu_b = 0 \), but it has a non-zero alpha, hedging an
existing factor \( c \) with positive expected return \( \mu_c > 0 \), the standard logic of Fama and French (1993) says this is a distinct anomaly. But, for an investor who cares only about minimizing transaction costs, factors are anomalies if and only if they have a non-zero expected return in the sense of \( \mu_b \).

So, what to make of the exposure to unit portfolios in the second half of Equation 3.6? The unit exposures depend not only on expected factor returns but also on the correlation structure of scores. These are the unintended common risks of the raw portfolio exposures in the first half of Equation 3.6. These exposures to unit portfolios are interesting, but not relevant to the investor’s optimization problem. For a risk neutral investor, the optimal portfolio inherits unintended but irrelevant risk exposures. Hedging these unintended risks requires transaction costs and is therefore suboptimal.

**Portfolio Choice With Transaction Costs:** For a risk neutral investor, identifying anomalies relies only on procedures like Fama and MacBeth (1973), using the results of papers that are in the spirit of Fama and French (1992) to test the significance of \( \mu_a \):

Cross Section Test: \( \hat{\mu} = \bar{f} = \frac{1}{T} \sum_t f_t = \frac{1}{T} \sum_t (\Gamma_t' \Gamma_t)^{-1} \Gamma_t' r_t = 0 \)  \( (3.7) \)

If a given factor \( \gamma_a \) passes this test of \( \mu_a \neq 0 \), we call it a score anomaly. Performing factor regressions in the spirit of Fama and French (1993) on the resulting factor portfolios will lead to mistakes in factor selection and the search for anomalies, excluding valuable anomalies and including apparent anomalies that are valuable only for their risk properties \( \Sigma \) and not their expected returns. For the risk neutral investor facing transaction costs, the search for anomalies starts and ends with score anomalies.

**Example 1**

Consider an example of two factors, call them CAPM beta and operating profitability. (We are ignoring the market portfolio for the moment to keep \( K = 2 \).) Suppose that these two factors have the following structure, so that operating profitability and beta have a negative
correlation from a common component $c$, with the residual components $a$ and $b$ uncorrelated with each other.

$$\gamma_{a,i} = \frac{OP}{A_i} = c_i + a_i$$

$$\gamma_{b,i} = \beta_i - 1 = -c_i + b_i$$

Suppose the expected returns of the two factor portfolios are roughly equal to their average returns in US data, so that $\mu_a > 0$ and $\mu_b = 0$. The optimal stock weight is proportional to its own mean return:

$$w^*_{t,c,i} = \frac{1}{\theta} (\gamma_{a,i}\mu_a + \gamma_{b,i}\mu_b) = \frac{1}{\theta} \gamma_{a,i}\mu_a$$

The optimal portfolio is a scaled version of the raw portfolio $a$ that tilts towards firms with high operating profits and away from firms with low operating profits. The absolute magnitude of the weights depends on the expected factor return $\mu_a$ net of transaction costs $\theta$. When the ratio of return to cost is high, the weights are correspondingly large. When the ratio of returns to cost is low, the weights are correspondingly small. This is in loose terms a reflection of the capacity of the strategy in light of execution costs.

No information about CAPM beta is needed to form the optimal portfolio. To this investor, CAPM beta is dead, and can be safely ignored, because it does not pass a significance test that rejects $\mu_b = 0$ in the sense of Fama and French (1992). But, the exposure of the optimal portfolio to the unit CAPM beta portfolio $b$ reveals that the optimal weights have an incidental exposure, which here comes from the common component $c$. For the risk neutral investor, this incidental exposure is irrelevant in setting weights. Our investor could neutralize the exposure to CAPM beta, but he chooses not to. The intuitive rationale is that hedging this zero return exposure raises transaction costs without increasing the investor’s utility. The results of this choice can be seen in the raw and unit exposures of the optimal portfolio, using
Equation 3.2 and using notation $\text{var}(a) = s_a^2$:

\[
e_{\text{raw}} (w_{tc}^*) = \frac{1}{\theta} \mu = \frac{1}{\theta} \begin{bmatrix} \frac{1}{\theta} \mu_a \\ 0 \end{bmatrix}
\]

\[
e_{\text{unit}} (w_{tc}^*) = \frac{1}{\theta} \Gamma' \Gamma \mu = \frac{1}{\theta} \begin{bmatrix} s_c^2 + s_a^2 & -s_c^2 \\ -s_c^2 & s_c^2 + s_b^2 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \mu_a \\ 0 \end{bmatrix} = \frac{1}{\theta} \begin{bmatrix} (s_c^2 + s_a^2) \mu_a \\ 0 \\ -1/s_c^2 \mu_a \end{bmatrix}
\]

The second entry in the unit portfolio exposure vector shows the negative beta tilt that comes incidentally from exploiting operating profitability. Meanwhile, the second entry in the raw portfolio exposure vector shows that market beta is not a score anomaly in this two factor example for a risk neutral investor.

3.2.3.2 Risk Averse, No Transaction Costs

Next, we consider the case where transaction costs $\theta$ are equal to zero but our investor is risk averse, so that $\lambda > 0$. Then, the optimal weights from Equation 3.5 simplify to:

\[
w_{ra}^* = \frac{1}{\lambda} \left( \Gamma \Sigma \Gamma' + \sigma^2 I \right)^{-1} \Gamma \mu
\]

This is the classic solution to mean-variance optimization, when the return generating process is expressed with a linear factor structure. Weights are increasing in the individual stock expected returns, which here can be expressed as the linear combination of factor scores and expected factor returns. and weights are decreasing in the individual stock contributions to risk, which here can be expressed as the product of factor scores and the covariance of factor returns $\Gamma \Sigma \Gamma'$ plus idiosyncratic risk $\sigma^2$. To get more visibility into the optimal weights, we can compute the exposure $e$ of this portfolio $w_{ra}^*$ to the unit factor portfolio. This is easiest to do by rearranging the first order condition in Equation 3.8 and substituting the definition of exposure to the unit portfolio from Equation 3.2:

\[
\Gamma \mu = \lambda \left( \Gamma \Sigma \Gamma' w_{ra}^* + \sigma^2 w_{ra}^* \right) = \lambda \Gamma \Sigma e_{\text{unit}} (w_{ra}^*) + \lambda \sigma^2 w_{ra}^*
\]

We can further rearrange, and take limits as the number of stocks $N$ grows large. As this happens the weight on any one security becomes small, and both sides of the equation go
towards zero, allowing us to derive a simple and intuitive expression for the exposure of the optimal portfolio \( w_{ra}^* \), to the unit factor portfolio:

\[
\Gamma (\mu - \lambda \Sigma e_{\text{unit}} (w_{ra}^*)) = \lambda \sigma^2 w_{ra}^* \rightarrow 0
\]

\[
\Rightarrow e_{\text{unit}} (w_{ra}^*) = \frac{1}{\lambda} \Sigma^{-1} \mu
\]

The upshot is that our risk averse investor’s problem reduces to a mean-variance optimization of factor portfolios, when he can trade frictionlessly at \( \theta = 0 \) with the number of available stocks \( N \) large relative to the number of factors \( K \). It no longer suffices to learn \( \mu \). Now, the covariance properties of the factor portfolios \( \Sigma \) are also relevant. If a factor has a positive expected return \( \mu_a > 0 \), but it has a zero alpha with respect to an existing factor \( b \), the standard logic of Fama and French (1993) says this is not a distinct anomaly, and indeed it is not for a risk-averse investor who faces no transaction costs. If factor \( a \) has a zero expected return \( \mu_a = 0 \), but it has a non-zero alpha, hedging an existing factor \( b \) with positive expected return \( \mu_b > 0 \), the standard logic of Fama and French (1993) says this is a distinct anomaly. And, for a risk averse investor who does not care about minimizing transaction costs, this is indeed a useful hedge.

**Portfolio Choice With Risk Aversion:** For a risk averse investor facing no transaction costs, the search for anomalies occurs in two steps. The first step is to use a procedure like Fama and MacBeth (1973) to estimate both \( \mu \) and \( \Sigma \) as:

\[
\hat{\mu} = \bar{f} = \frac{1}{T} \sum_t f_t = \frac{1}{T} \sum_t (\Gamma' \Gamma^{-1}) \Gamma' r_t
\]

\[
\hat{\Sigma} = \text{vár} (f_t) = \frac{1}{T} \sum_t (f_t - \bar{f}) (f_t - \bar{f})'
\]

The second step is to perform factor regressions in the spirit of Fama and French (1993) on the resulting factor portfolio returns, including those with zero means \( \mu_a = 0 \). This will lead to the elimination of factor portfolios with positive means but zero alphas, and lead to the resurrection of factor portfolios with zero means but non-zero alphas that come from their useful hedging properties. Alpha here and above refers to a Jensen (1968) alpha test. Sample averages are inserted into
Equation 3.9, $\hat{\sigma}$ refers to elements of $\hat{\Sigma}$, and the null is that this first factor $b$ is irrelevant in the choice of portfolio weights. If under the null, $e_{\text{unit},b}(w_{ra}^*) = 0$, this implies:
\[
\sum_{k \neq b} \hat{\sigma}_{bk} e_{\text{unit},k}(w_{ra}^*) = \hat{\mu}_b
\] (3.10)

Equation 3.10 is equivalent to regressing the time series of $f_b$ on the time series of the remaining factors excluding $b$, estimating the multivariate factor loadings $\beta$ and testing the significance of the intercept:

Time Series Test: $\hat{\mu}_b - \sum_{k \neq b} \hat{\beta}_{bk} \hat{\mu}_k = 0$ (3.11)

This derivation is in Appendix C.2. This means that some factors with $\mu_b = 0$ can nonetheless be factor anomalies because of their covariance properties mean that $\mu_b - \beta_{ba} \mu_a - \beta_{bc} \mu_c + \cdots \neq 0$. Some factors with $\mu_b \neq 0$ may nonetheless not be factor anomalies, because of their covariance properties $\mu_b - \beta_{ba} \mu_a - \beta_{bc} \mu_c + \cdots = 0$. For the risk averse investor facing no transaction costs, the search for anomalies starts and ends with factor anomalies. In our nomenclature, these remain score anomalies, but they are not factor anomalies.

Example 2

Consider again two factors, as in Example 1, but the first is now the standard market portfolio, and the second is still the CAPM beta, as before, with the mean return on the market factor $\mu_a > 0$ and the mean return on the beta factor $\mu_b = 0$. Further, assume that the payoffs to the two factors are positively correlated, so that the off-diagonal elements of $\Sigma$ are positive, so that $\sigma_{ab} > 0$, meaning that a portfolio that is long firms with high betas has relatively higher returns when the market also has relatively higher returns, as is true in US data. Plugging in to Equation 3.9

\[
e_{\text{unit}}(w_{ra}^*) = \begin{bmatrix}
e_{\text{unit},r_m} \\
e_{\text{unit},\beta}
\end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix}
\frac{\sigma_{ba}^2}{\sigma_{aa}^2 - (\sigma_{ab})^2} \mu_a \\
- \frac{\sigma_{ba}^2}{\sigma_{aa}^2 - (\sigma_{ab})^2} \mu_a
\end{bmatrix}
\]
The optimal strategy involves exposures to the market portfolio and the unit beta portfolios. The absolute magnitude of the exposure to the market portfolio is increasing in its expected factor return $\mu_a$ and decreasing in its expected favor risk $\sigma_{aa}^2$, so roughly speaking increasing in the Sharpe ratio of the market, and this exposure is further increased because its risk can be mitigated with a short position in the unit beta portfolio. Unlike the case of a risk neutral investor facing transaction costs, this exposure is worth hedging despite its zero mean return, because it lowers risk and is costless to execute. So, in that sense the early reports of the death of market beta are exaggerated. It is very much alive as a factor anomaly, relevant for a risk averse investor facing no transaction costs, in this two factor example.

**Example 3**

Consider again an example of two factors in Example 1, now call them the ratio of book to market equity and low asset growth. Suppose that the mean returns are positive $\mu_a, \mu_b > 0$ as they are in US data. Further, assume that the payoffs to the two factors are positively correlated, so that the off-diagonal elements of $\Sigma$ are positive, so that $\sigma_{ab} > 0$, meaning that a portfolio that is long firms with high ratios of book to market equity has relatively higher returns when a portfolio that is long firms with low asset growth also has relatively higher returns. Now we have the more general version of the unit exposures from Equation 3.9:

$$
e_{\text{unit}}(w^*_{ra}) = \begin{pmatrix} e_{\text{unit}, \frac{\mu_a}{\lambda}} \\ e_{\text{unit}, \frac{\Delta \lambda}{\lambda}} \end{pmatrix} = \frac{1}{\lambda} \begin{pmatrix} \sigma_{bb}^2 (\sigma_{aa}^2 - (\sigma_{ab})^2) \mu_a - \sigma_{ab} \sigma_{bb}^2 (\sigma_{aa}^2 - (\sigma_{ab})^2) \mu_b \\ \sigma_{ab}^2 (\sigma_{aa}^2 - (\sigma_{ab})^2) \mu_b - \sigma_{ab} \sigma_{bb}^2 (\sigma_{aa}^2 - (\sigma_{ab})^2) \mu_b \end{pmatrix}
$$

It turns out that empirically, the first entry is indistinguishable from zero, because the alpha of the unit book to market equity portfolio when low asset growth is included as a reference portfolio is approximately zero: $\mu_a - \frac{\sigma_{ab}}{\sigma_{bb}^2} \mu_b \approx 0$. In that sense, we could substitute out information on book to market equity:

$$
e_{\text{unit}}(w^*_{ra}) = \frac{1}{\lambda} \begin{pmatrix} 0 \\ \frac{1}{\sigma_{bb}^2} \mu_b \end{pmatrix} \tag{3.12}
$$

So, the ratio of book to market equity is not a factor anomaly, and not relevant for a risk
averse investor facing no transaction costs, in this two factor example. The only information that is needed is the unit low asset growth portfolio and its Sharpe ratio. Interestingly though, the ratio of book to market equity remains a score anomaly, relevant to an investor facing realistic transaction costs, as we see next.

### 3.2.3.3 Risk Averse, Constant Transaction Costs

The more general case involves both risk aversion and transaction costs, bringing us back to Equation 3.5. Intuitively, when transaction costs are small, the conclusions of the simple risk aversion case apply. Transaction costs act much like idiosyncratic risk, which is an issue when the investment opportunity set is small, but not when the number of securities \( N \) is very large. However, there are reasons to believe that the effect of \( \theta \) might be meaningful even when the effect of idiosyncratic \( \sigma^2 \) is small. When assets under management are large, even a small weight in a given stock might be large in comparison to its trading volume. Recall that Fama and MacBeth tests as in Equation 3.7 are necessary and sufficient for identifying anomalies in the presence of transaction costs alone. And Jensen’s alpha tests as in Equation 3.11 are necessary and sufficient for identifying anomalies in the presence of risk aversion alone. When our investor is risk averse and faces transaction costs, anomalies of both types are relevant.

We can start with a variant of Equation 3.9 where we assume that the considerations of idiosyncratic risk are second order but transaction costs remain first order, so that \( \theta \gg \lambda \sigma^2 \):

\[
\begin{align*}
\mathbf{\Gamma} \left( \mathbf{\mu} - \lambda \Sigma \mathbf{e}_{\text{unit}} \left( \mathbf{w}^\ast \right) \right) &= (\lambda \sigma^2 + \theta) \mathbf{w}^\ast \rightarrow \theta \mathbf{w}^\ast \\
\Rightarrow \mathbf{e}_{\text{unit}} \left( \mathbf{w}^\ast \right) &= \frac{1}{\lambda \Sigma^{-1}} \left( \mathbf{\mu} - 2 \theta \mathbf{e}_{\text{raw}} \left( \mathbf{w}^\ast \right) \right) 
\end{align*}
\]  

(3.13)

While it is obviously unappealing to have optimal weights and exposures on both sides of the equation, Equation 3.13 is a useful intermediate relationship. Note that the optimal unit exposures are driven by the standard risk and return considerations, but the return is haircut by the transaction costs associated with the raw exposures of the optimal weights. The intuitive appeal is that we can think of the risk averse investor optimizing over risk and net of transaction cost return.
We can eliminate the dependence using the definition of raw exposures from Equation 3.2:

$$e_{\text{unit}}(\mathbf{w}^*) = \left[ \lambda \Sigma + \theta (\Gamma' \Gamma)^{-1} \right]^{-1} \mu$$  (3.14)

We examine the intuition in Equation 3.14 in the example below, but it is apparent that the Jensen’s alpha test will no longer be necessary for a factor to be a relevant anomaly. And, in general, factors that pass either test will be worthy of consideration.

**Portfolio Choice With Risk Aversion and Transaction Costs:** For a risk averse investor facing transaction costs, the search for anomalies means finding factors that pass either the test in the spirit of Fama and MacBeth (1973) in Equation 3.7 or the test in the spirit of Jensen’s alpha test in Equation 3.11.

**Example 4**

Consider the same setup as Example 3, with the ratio of book to market equity and low asset growth. Recall we found that the ratio of book to market was a score anomaly, because it has a mean return greater than zero $\mu_a > 0$, but not a factor anomaly, at least controlling for low asset growth, because its Jensen’s alpha is zero, with $\mu_a - \frac{\sigma_{ab}}{\sigma_{bb}} \mu_b \approx 0$. Further, assume that book to market equity and low asset growth scores, $\gamma_a$ and $\gamma_b$ are constructed so as to be orthogonal to one another, with unit standard deviation, which has the notation benefit of making $\Gamma' \Gamma = NI$. Note that the correlation of factor returns $\Sigma$ can still be high, even if the scores are uncorrelated. Now, we start from the result in Example 2, Equation 3.12 and generalize the unit exposures to include transaction cost effects:

$$e_{\text{unit}}(\mathbf{w}^*) = \frac{1}{\lambda C} \left[ \begin{array}{cc} \sigma_{aa}^2 + \frac{N \theta}{\lambda} & \sigma_{ab} \\ \sigma_{ab} & \sigma_{bb}^2 + \frac{N \theta}{\lambda} \end{array} \right]^{-1} \left[ \begin{array}{c} \mu_a \\ \mu_b \end{array} \right]$$

We substitute $C \equiv \left( \left( \sigma_{aa}^2 + \frac{N \theta}{\lambda} \right) \left( \sigma_{bb}^2 + \frac{N \theta}{\lambda} \right) - (\sigma_{ab})^2 \right)$ to save space. The optimal exposure
to the unit portfolios is a balance of two concerns. If the diagonal of factor risk $\Sigma$ is high and risk aversion is high, then the solution looks like Equation 3.9 while if the transaction costs $\theta$ are high, then the solution looks like Equation 3.6. And, in general, the optimal solution is a blend of these two concerns: managing portfolio risk-adjusted return while keeping execution costs low. Even substituting in the implications of a Jensen’s alpha of zero for the book to market factor portfolio, there is still positive portfolio exposure to this factor, because of its raw factor return $\mu_a$:

$$e_{unit}(w^*) = \frac{1}{\lambda C} \left[ \frac{N\theta}{\lambda} \mu_a \left( \frac{\sigma_{a\alpha}^2 - (\sigma_{ab})^2}{\sigma_{b\beta}^2} + \frac{N\theta}{\lambda} \right) \mu_b \right]$$

So, the ratio of book to market equity is not an anomaly for a risk averse investor facing no transaction costs in this two factor example, but it is resurrected in the case involving transaction costs. Our investor tilts somewhat toward the lower alpha unit portfolio to access raw returns at lower cost. The general conclusion is that the investor will put weight on both factor and score anomalies.

### 3.2.4 Summary

Suppose a candidate factor appears: A firm characteristic that has the potential to predict its stock returns. Is it a relevant anomaly? The literature contains two standard asset pricing tests. The first uses Fama and MacBeth regressions of the sort in Equation 3.7 testing whether the mean of the new factor returns is equal to zero using a series of multivariate cross-sectional regressions that contain other factors known to predict the cross section of stock returns. We show that this test is always applicable for an investor facing a simple form of constant transaction costs, whether he is risk averse or not. The second uses Jensen’s alpha tests of the sort in Equation 3.11. These test whether the intercept in a regression of the new factor portfolio returns, from the aforementioned cross-sectional regressions, on the portfolio returns of existing factors. We show that this test is always applicable for a a risk averse investor, whether he faces transaction costs or not. As it turns out, both are applicable for the general case of a risk-averse investor facing a simple form of transaction costs.
Realistic transaction costs complicate these conclusions, but in an intuitive way. When transaction costs vary across stocks, when assets under management are substantial, and when dynamic trading considerations appear, each of tests should in principle be performed on factor returns that are net of execution costs, and gross of future, persistence in returns. Tests of statistical significance in the cross section test are still valid in a number of special cases. Appendix C.1 provides some illustrations of how this might be done in practice.

We now turn to the power of the two tests.

### 3.3 The Power of Standard Asset Pricing Tests

The previous section shows the relevance of both score and factor anomalies for portfolio choice. In this section, we turn to the econometrics of identifying anomalies. There are two standard tests in empirical asset pricing. Tests in the spirit of Fama and MacBeth (1973) and Fama and French (1992) are cross-sectional and do not consider the covariance properties of factor portfolio returns. Tests in the spirit of Jensen (1968) and Fama and French (1993) use time-series data and focus on the covariance properties of the factor portfolios as well their means.

We take two approaches to estimating the power of these two tests. In the first, we compute the asymptotic power curve analytically. The score test is more powerful by a multiplier that is increasing in the Sharpe ratio of the factor portfolios. In the second, we simulate the data generating process in small samples, and compute a simulated power curve. The effect of the Sharpe ratio is magnified by a small sample estimation of the factor portfolio return covariances.

#### 3.3.1 Asymptotic Power Curves

Section 2 considers the problem facing an investor who understands the return generating process in Equation 3.1. Realistically, the investor does not know the expected payoffs to factor portfolios $\mu$ or their covariances $\Sigma$. He must use data on scores and stock returns to estimate these. And, in practice, our investor also needs to consider whether market forces
might make estimates irrelevant for future returns. For now, we assume that our investor differs from the representative investor in either preferences or beliefs, so that the history of stock returns can be used to produce reliable forecasts of the parameters in Equation 3.1.

3.3.1.1 Data Mining

This raises the issue of data mining, which can take two forms: selection and overfitting. Problems of selection stem from starting with $n$ candidate anomalies and choosing $k < n$ that work in the historical data. Problems of overfitting come from optimally weighting $n$ candidate anomalies using their in-sample performance into one aggregate blended superscore. In both cases, the likelihood of Type I error, where a firm characteristic is deemed spuriously to be an anomaly, is high. These issues are discussed in Mclean and Pontiff (2016), Harvey et al. (2015), and Bailey and López de Prado (2014). For example, Mclean and Pontiff (2016) examine the efficacy of anomalies using realistic rolling estimations of factor average returns and the incremental effect of publicizing the finding. Simin (2008) and Levi and Welch (2014) find that rolling estimates of expected returns from the Fama-French three factor model do not work well at forecasting return realizations. Their findings point to researcher data mining. Novy-marx (2016) computes adjusted t-statistics as a function of $n$, $k$, and the nature of the overfitting procedure employed. Given the number of researchers focused on asset pricing and the academic and commercial incentives for documenting anomalies, the issue of Type I errors is paramount.

However, Lewellen (2015) finds that rolling estimates using characteristics as in our return generating process in Equation 3.1 produce a t-statistic greater than 10. Moreover, it is worth considering the power to detect anomalies in the first place, and the possibility of Type II errors, which is the focus of this paper. For example, it is possible to turn the argument in Novy-marx (2016) around. Suppose the objective was not to establish an anomaly but rather to overturn an existing anomaly. A similar logic then applies. By starting with $n$ potential controls or established anomalies and risk factors and choosing $k < n$ in the historical data, it is possible to lower the power a test of the risk anomaly. This is particularly true in time
series tests, as we argue below, and Novy-Marx (2012), who argues that the risk anomaly is subsumed by profitability and value, is an example of this approach.

In this paper, we sidestep the issue of data mining and imagine that there is an established set of anomalies to which a single new candidate may be added. We start with cross sectional tests, which are relevant for an investor facing material transaction costs.

### 3.3.1.2 The Cross Section Test

The investor, or the econometrician, estimates Equation 3.1 where we assume that the first factor is the market portfolio, so that $\gamma_a = 1$, with a series of cross sectional regressions:

$$ r_t \sim \Gamma_t f_t + \varepsilon_t \Rightarrow \hat{f}_t = (\Gamma'_t \Gamma_t)^{-1} \Gamma'_t r_t $$

The time series mean of the factor payoffs is the Fama and MacBeth estimator of the mean payoff $\mu$:

$$ \hat{\mu} = \frac{1}{T} \sum_t \hat{f}_t $$

In a Bayesian sense, all factors will be relevant in at least a small way, in the sense that the point estimate for the return $\mu_b$ on a factor $b$ will never be exactly zero. But, in a frequentist sense, there are some factors that cannot be statistically deemed relevant. This is the usual notion of a score anomaly: We must be able to say that the hypothesis that $\mu_b \neq 0$ is true with sufficient probability. The standard error of the estimate of $\mu_b$ is equal to:

$$ se(\hat{\mu}_b) = \frac{1}{\sqrt{T}} \sqrt{\frac{\sigma^2_{bb}}{N^2} + \frac{1}{\sigma^2} \sum_i (\delta'_b \gamma_i)^2 } $$

(3.16)

Where $\sigma^2_{bb}$ is the variance of factor $b$ from the factor covariance matrix $\Sigma$, $N$ is the number of firms in the cross section, and $\sigma^2$ is the idiosyncratic variance that is assumed for simplicity to be constant across firms. And, we define the vector $\delta_b$ from the inverse of the matrix $\Gamma' \Gamma$ as follows:

$$ N (\Gamma' \Gamma)^{-1} = \begin{bmatrix} \delta_a & \delta_b & \delta_c & \cdots \end{bmatrix} $$

We derive the expression for standard error in Equation (3.16) in Appendix C.2 Because we estimate the time series mean of the regression coefficients, we lose 1 degree of freedom.
and the estimate of the asymptotic variance is

$$\hat{se}(\hat{\mu}_b) = \frac{1}{\sqrt{T-1}} \sqrt{\hat{\sigma}_{bb}^2 + \frac{1}{N^2} \sigma^2 \sum_i (\delta'_i \gamma_i)^2}$$

For a large cross section of stocks $N$, the effect of residual risk $\varepsilon$ becomes small in the second part of the expression under the radical. The estimate of the factor payoff in each period using only returns $r_t$ and scores $\Gamma_t$ at any given time $t$ becomes very precise. The error in the estimate of the mean then depends only on the number of time periods $T$. This is the asymptotic standard error in the sense that $\sigma_{bb}^2$ and $\sigma^2$ are not known to the investor or the econometrician in small samples. We consider the small sample properties with Monte Carlo simulations below. For now, we plot the power curve using Equation 3.16 in Figure 3.1, using

$$N = [50; 100; 250; 1,000; 10,000]$$
$$T = [50; 100; 200; 500; 1,000]$$
$$\sigma_\varepsilon = [0; 3.27; 5; 6; 7]$$
$$\sigma_{bb} = [0.5; 1; 2; 3; 4]$$

Underlined values are the base case and represents by the blue line in the graph. Monthly idiosyncratic risk $\sigma_\varepsilon$ and factor risk $\sigma_{bb}$ are in percentage. We vary monthly factor return $\mu_b$ so that its annual Sharpe ratio ranges from 0 to 2. Annual Sharpe ratio is calculated as

$$SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{12}.$$  

The power curves illustrate the probability that the null of zero is rejected given a variety of inputs for the true mean of factor return $\mu_b$. It shows the power our investor has to detect relevant score anomalies. The $y$-intercept is the size of the test. There is a 5% chance of rejecting the null of a zero factor return when it is truly zero. In these situations, the investor concludes that there is a score anomaly when in truth there is not one. The shape of the power curve is otherwise not terribly interesting, as it depends on the assumed distribution of factor scores, the number of separate periods $T$, the assumed factor variance $\sigma_{bb}^2$, and the assumed number of firms $N$, and idiosyncratic risk $\sigma_2$. These shift the power curve in intuitive ways. The power rises more steeply with more firms in Panel (a) and lower idiosyncratic variance in Panel (b). A larger number of firms in the cross-section helps to eliminate the
Figure 3.1: Asymptotic Power Curve of The Cross Section Test

Note: The annual Sharpe ratio of the test factor is calculated as $SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{12}$.

effect of idiosyncratic risk on factor returns. A smaller amount of idiosyncratic risk has a similar effect, in that even a small number of firms deliver a pure factor return. In both cases though, the improvements in power are limited. Even an infinite number of firms does not lead to an extremely powerful test, capable of detecting small anomalies. The power also rises more steeply with more time periods in Panel (c) and lower factor variance in Panel (d). A larger number of time periods means that the mean factor return per period can be estimated with greater and greater accuracy, assuming there are no changes in the underlying return
generating process. Similarly, power rises more quickly if the factor payoffs are very reliable, falling very close to the mean in every period. These are situations where a score anomaly can be reliably detected even when the true average return is quite small. All four panels show that economically large score anomalies are always detected even in modest time series, but these will be rare in competitive markets, so power is important.

We are more concerned with the relative power of the cross section and time series tests, which we turn to next, than we are about the other comparative statics, which will improve the power of both proportionally.

### 3.3.1.3 The Time Series Test

We next move to the time series test, which is relevant to an investor who is risk averse. In this case, there is a second step after the estimation of factor payoffs in Equation 3.15. Practically, our investor is interested in whether a particular factor will have zero effect on his portfolio choice in Equation 3.9. And, as we argue above, this is equivalent to the factor passing a Jensen (1968) alpha test in Equation 3.11. We leave off the hats on the factor returns in the regression of the returns of a given factor $b$ on the remaining factors other than $b$:

$$f_{b,t} \sim \alpha_b + \beta'_b f_{-b,t} + \epsilon_t \Rightarrow \hat{\alpha}_b = \bar{f}_b - \bar{\beta}'_{-b} = \hat{\mu}_b - \hat{\beta}'_b \hat{\mu}_{-b}$$

We use the subscript to indicate the full set of factor returns $f_{-b}$ or means of the factor returns $\mu_{-b}$ to indicate the vectors that exclude the factor $b$. The factor loadings are the vector $\beta_{-b}$, and the variable of interest is the factor-risk-adjusted return $\alpha_b$. This is an intercept test to see whether the factor return $\mu_b$ is large enough given its covariances with the other factors in the set of anomalies. Again, in a Bayesian sense, all factors will be relevant in some small way. In a frequentist sense, we are interested in ruling out factors that are statistically irrelevant. This is the usual sense of a factor anomaly, that we must be able to say that the hypothesis that $\alpha_b \neq 0$ is true with sufficient probability. The standard error of
the estimate of $\alpha_b$ is equal to:

$$
\text{se} (\hat{\alpha}_b) = \frac{1}{\sqrt{T}} \sqrt{\sigma^2 \left(1 + \mu'_{-b} \Sigma^{-1}_{-b} \mu_{-b}\right)}
$$

$$
= \frac{1}{\sqrt{T}} \sqrt{\left(\sigma^2_{bb} + \frac{1}{N^2} \sigma^2 \sum \gamma_i^2 \right)(1 - R^2) (1 + SR^2)}
$$

We derive the expression for standard error in Equation (3.17) in Appendix C.2. Because in the time series regression we estimate the $K + 1$ coefficients, we lose $K + 1$ degrees of freedom and the estimate of asymptotic variance is

$$
\text{se} (\hat{\epsilon}_b) = \frac{1}{\sqrt{T - K - 1}} \sqrt{\sigma^2 \left(1 + \mu'_{-b} \Sigma^{-1}_{-b} \mu_{-b}\right)}
$$

$$
= \frac{1}{\sqrt{T - K - 1}} \sqrt{\left(\sigma^2_{bb} + \frac{1}{N^2} \sigma^2 \sum \gamma_i^2 \right)(1 - R^2) (1 + SR^2)}
$$

For a large cross section of stocks $N$, the effect of residual risk $\epsilon$ becomes small and the variance of the residual factor return risk $\epsilon$ is simply $\sigma^2_{bb}$. The estimate of the factor payoff in each period using only returns $r_t$ and scores $\Gamma_t$ at any given time $t$ becomes very precise. When the cross-section of firms $N$ is not so large, then the variance is greater, so that $\sigma^2 > \sigma^2_{bb}$, and equals the quantity under the radical in Equation (3.16) discussed above. In addition to these comparative statics, there are two additional drivers of the standard error of the time series test. The standard error now depends on the means and covariances of the other factor returns too. The first term in parentheses is an increase in power that comes from the fact that the other factors can in principle reduce the residual variation in the regression equation. The residual in the time series test is smaller than the residual in the cross section test by an amount equal to one minus the time series R-squared. The second term in parentheses is decrease in power, equal to one plus the in-sample maximum squared Sharpe Ratio (SR) of the other factor returns. This is the mean-variance optimal combination of the existing factors. When the existing factors are very powerful predictors of return, then the standard error in Equation (3.17) rises, as in Shanken (1992). It is harder to reject the null. This is the asymptotic standard error again in the sense that $\sigma^2_{bb}$, $\sigma^2$, $\mu_{-b}$, and $\Sigma_{-b}$ are not known to the investor or the econometrician in small samples. We consider the small sample properties with Monte Carlo simulations below. For now, we plot the power curve using Equation (3.17) in Figure 3.2 using the same parameters as in Figure 3.1.
Figure 3.2: Asymptotic Power Curve of The Time Series Test

Note: The annual Sharpe ratio of the test factor is calculated as $SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{T}$.

The power curve exactly mirrors the results in Figure 3.1 but with the power shifted down by the in-sample Sharpe ratio. As before power rises more steeply with more firms and lower idiosyncratic variance. The gains from these two parameters are bounded. Power rises more quickly with more periods and lower factor variance. With a large number of periods and low factor variance, the time series test can detect small anomalies.

To these comparative statics, we now add the Sharpe ratio of the existing factors in the next section. While it is not immediately apparent in the comparison of Figure 3.1 and Figure
3.2 The time series tests are all shifted down somewhat.

3.3.1.4 A Comparison of the Cross Section and Time Series Tests

It is immediately apparent that the standard error in Equation 3.17 contains a potential loss in power, when compared to the standard error in 3.16. If a new factor has no true connection to the time series payoffs of the incumbent set, so that the R-squared in 3.17 is zero, then the standard errors are related by the following formula:

$$\text{se}(\hat{\alpha}_b) = \text{se}(\hat{\mu}_b) \sqrt{1 + \mu_{-b}^\prime \Sigma_{-b}^{-1} \mu_{-b}}$$  \hspace{1cm} (3.18)

Correspondingly, the relation between estimates of asymptotic standard error is

$$\hat{\text{se}}(\hat{\alpha}_b) = \hat{\text{se}}(\hat{\mu}_b) \sqrt{\frac{T - 1}{T - K - 1}} \sqrt{1 + \mu_{-b}^\prime \Sigma_{-b}^{-1} \mu_{-b}}$$

In other words, it is harder to find a new factor anomaly than it is to find a new score anomaly. We overlay the power curves in Figure 3.1 onto Figure 3.3 at various Sharpe ratios:

$$SR = [0.25, 0.42, 0.50, 0.75, 1]$$

This has a nice intuition. When the predictive power of existing factors is large relative to the portfolio variance—the Sharpe ratio is high—the estimation of a covariances $\beta_{-b}$ between the new factor $b$ and existing factors becomes a source of error that drives power down relative to the cross-section test. It is possible to connect the new factor to existing ones in a way that is spurious. For example, both momentum and CAPM beta suffered very poor returns in the market reversal of the spring of 2009, but are otherwise essentially uncorrelated. Similarly much of the overlap between CAPM beta and the ratio of price to book occurs in the late 1990s and early 2000s. If we consider the Sharpe ratio of momentum and the price-to-book to be high—as they are in US data—and for these to be a spurious correlations—which may be true—then it is easier to reject CAPM beta as a factor anomaly. It is this possibility which lowers the relative power of the test. In this sense, our argument is related to Ang et al. (2008). They focus on the loss of power that can come from aggregation. Forming portfolios
Figure 3.3: Asymptotic Power Curve of the CS and TS Test at Various Sharpe Ratios of the Incumbent Factor Model

Note: The annual Sharpe ratio of the test factor is calculated as \( SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{12} \).

improves the estimation of covariances but throws away information. In our context, it is the extra need—for risk averse investors—to compute covariances that diminishes their ability to find relevant anomalies.

3.3.1.5 The Sharpe Ratio in US Data

This begs the question: Which line in Figure 3.3 in the comparison of cross section and time series tests is the relevant one? This depends on the size of the in-sample Sharpe ratio for some set of existing factors, like the Fama-French five factors. Power falls monotonically in two ways: as the factor set increases, and as the horizon rises, from daily to weekly to monthly to annual returns.

First, a larger factor set by definition means a higher in-sample Sharpe ratio. The Fama and French (2015) five-factor model, for example, will mechanically reject more potential factor anomalies than the Fama and French (1993) three-factor model. So, there is a practical limit on the number of factor anomalies that can be discovered. By contrast, the number of
score anomalies that can be discovered is only limited by the number of stocks \( N \) and the number of periods \( T \), but it is not otherwise constrained by the predictive power of existing cross-sectionally orthogonal anomalies.

Second, a property of the Sharpe ratio is that it rises as the return horizon increases. This is because it is the ratio of mean to standard deviation. While the mean increases linearly in \( T \), the standard deviation increases linearly in \( \sqrt{T} \).

The qualitative impact of these two effects is clear. We use US data to convert qualitative to quantitative effects on econometric power. The Sharpe ratio of the set of standard, existing factors is reasonably high, when we use monthly returns, as is common practice in the academic literature. The bold line in Figure 3.3 shows the loss in power of a test that uses a standard set of the Fama-French five factors, momentum, and short-term reversal and a monthly return horizon. The empirical moments of these portfolios over the period from 1963:07 to 2016:03 are shown in Table 3.1. In Table 3.1 we also include a Fama-French style portfolio using market beta, using the beta estimation approach of Frazzini and Pedersen (2014). This portfolio divides the CRSP universe into small and large, using the median size among NYSE stocks as the breakpoints and further divides small and large stocks into three terciles according to market beta, again using NYSE breakpoints.

The individual annual Sharpe ratios range from 0.29 to 0.56, and the in-sample optimal annual Sharpe ratio of the seven factor returns, excluding market beta, is 1.45. The corresponding quarterly Sharpe is 0.73, the corresponding monthly Sharpe is 0.42 which is what we show in bold in Figure 3.3, and the corresponding daily Sharpe is 0.09.

### 3.3.1.6 Choosing Return Horizon

Plugging these Sharpe ratios into Equation 3.18, we can compute the loss in power. To fix ideas, we use \( T = 200, N = 250, \sigma_{bb} = 3, \sigma_{\varepsilon} = 3.27 \). For monthly tests, which is the standard in the literature and which we indicate in bold in Figure 3.3, the maximum loss in power is 22.4 percent for a new factor that has a true annual Sharpe ratio of 0.67—in other words, a very strong anomaly. Annual return horizons are rarely used, likely because they involve an
Table 3.1: Empirical Moments for Factor Returns (1963:07 to 2016:03)

(a) Mean and variance

<table>
<thead>
<tr>
<th></th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
<th>MOM</th>
<th>STREV</th>
<th>RISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>0.50</td>
<td>0.25</td>
<td>0.34</td>
<td>0.25</td>
<td>0.31</td>
<td>0.69</td>
<td>0.48</td>
<td>-0.22</td>
</tr>
<tr>
<td>Annualized</td>
<td>5.98</td>
<td>3.05</td>
<td>4.11</td>
<td>3.06</td>
<td>3.71</td>
<td>8.29</td>
<td>5.81</td>
<td>-2.63</td>
</tr>
<tr>
<td>SD</td>
<td>4.45</td>
<td>3.06</td>
<td>2.85</td>
<td>2.12</td>
<td>2.01</td>
<td>4.24</td>
<td>3.13</td>
<td>4.65</td>
</tr>
<tr>
<td>Annualized</td>
<td>15.41</td>
<td>10.62</td>
<td>9.89</td>
<td>7.35</td>
<td>6.96</td>
<td>14.67</td>
<td>10.84</td>
<td>16.10</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.11</td>
<td>0.08</td>
<td>0.12</td>
<td>0.12</td>
<td>0.15</td>
<td>0.16</td>
<td>0.15</td>
<td>-0.05</td>
</tr>
<tr>
<td>Annualized</td>
<td>0.39</td>
<td>0.29</td>
<td>0.42</td>
<td>0.42</td>
<td>0.53</td>
<td>0.56</td>
<td>0.54</td>
<td>-0.16</td>
</tr>
</tbody>
</table>

(b) Covariance

<table>
<thead>
<tr>
<th></th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
<th>MOM</th>
<th>STREV</th>
<th>RISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>19.75</td>
<td>3.82</td>
<td>-3.77</td>
<td>-1.93</td>
<td>-3.46</td>
<td>-2.44</td>
<td>3.96</td>
<td>14.79</td>
</tr>
<tr>
<td>SMB</td>
<td>3.82</td>
<td>9.38</td>
<td>-1.01</td>
<td>-2.35</td>
<td>-0.70</td>
<td>-0.30</td>
<td>1.51</td>
<td>3.60</td>
</tr>
<tr>
<td>HML</td>
<td>-3.77</td>
<td>-1.01</td>
<td>8.13</td>
<td>0.53</td>
<td>4.02</td>
<td>-2.00</td>
<td>-0.10</td>
<td>-4.68</td>
</tr>
<tr>
<td>RMW</td>
<td>-1.93</td>
<td>-2.35</td>
<td>0.53</td>
<td>4.49</td>
<td>-0.36</td>
<td>0.82</td>
<td>-0.41</td>
<td>-1.97</td>
</tr>
<tr>
<td>CMA</td>
<td>-3.46</td>
<td>-0.70</td>
<td>4.02</td>
<td>-0.36</td>
<td>-4.03</td>
<td>-0.08</td>
<td>-0.81</td>
<td>-4.19</td>
</tr>
<tr>
<td>MOM</td>
<td>-2.44</td>
<td>-0.30</td>
<td>-2.00</td>
<td>0.82</td>
<td>-0.08</td>
<td>17.92</td>
<td>-3.87</td>
<td>-4.16</td>
</tr>
<tr>
<td>STREV</td>
<td>3.96</td>
<td>1.51</td>
<td>-0.10</td>
<td>-0.41</td>
<td>-0.81</td>
<td>-3.87</td>
<td>9.77</td>
<td>3.22</td>
</tr>
<tr>
<td>RISK</td>
<td>14.79</td>
<td>3.60</td>
<td>-4.68</td>
<td>-1.97</td>
<td>-4.19</td>
<td>-4.16</td>
<td>3.22</td>
<td>21.57</td>
</tr>
</tbody>
</table>

Intuitively dramatic reduction in power, with the maximum loss at 73.0 percent. Quarterly returns are occasionally used. For example, see Cederburg and O’Doherty (2016). Quarterly returns lie in between, with a substantial maximum loss at 45.1 percent. This is an effective way to stack the deck against rejecting the null. Meanwhile, daily returns largely solve the power problem with a loss in power of only 1.4 percent at maximum.

To provide an illustration, we compute alphas in Table 3.2 for six of the Fama-French style factor portfolios summarized in Table 3.1. We leave out reversal, which has a payoff that is very short-lived. We use monthly, weekly, and daily returns and the Fama-French five-factor model in Panels A, B, and C, respectively. Our prior from Equation 3.18 and the properties of the Sharpe ratio as horizon falls is that the daily tests will be the most powerful, delivering lower standard errors and higher t-statistics on average. Table 3.2 bears this out. We are focused on the first two columns, which display coefficient estimates, standard errors, and t-statistics for the intercept, or alpha of each portfolio In the monthly tests in Panel A, the
average standard error of the alpha estimate for these six portfolios, at 1.25 percent annualized, is 25 percent higher than in the daily tests, shown in Panel C. The average t-statistic, at 4.38, is 44 percent higher. Notably, the conclusion from Fama and French (2015) that the price-to-book ratio is subsumed by the other factors is reversed in daily data. HML has a t-statistic of 0 in monthly data and a t-statistic of 2.19 in daily data. It retains a statistically significant alpha of 1.96 percent annualized.

### Table 3.2: Time Series Tests: Monthly, Weekly, Daily

#### (a) Monthly

<table>
<thead>
<tr>
<th>Intercept</th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>coef</strong></td>
<td><strong>se</strong></td>
<td><strong>coef</strong></td>
<td><strong>se</strong></td>
<td><strong>coef</strong></td>
<td><strong>se</strong></td>
</tr>
<tr>
<td>SMB</td>
<td>3.65 (1.36)</td>
<td>0.13 (0.03)</td>
<td>0.06 (0.05)</td>
<td>-0.42 (0.05)</td>
<td>-0.12 (0.08)</td>
</tr>
<tr>
<td>HML</td>
<td>0.00 (0.99)</td>
<td>0.02 (0.02)</td>
<td>0.03 (0.03)</td>
<td>0.14 (0.04)</td>
<td>1.01 (0.04)</td>
</tr>
<tr>
<td>RMW</td>
<td>4.72 (0.98)</td>
<td>-0.10 (0.02)</td>
<td>-0.22 (0.03)</td>
<td>0.14 (0.04)</td>
<td>-0.29 (0.06)</td>
</tr>
<tr>
<td>CMA</td>
<td>2.74 (0.66)</td>
<td>-0.11 (0.01)</td>
<td>-0.03 (0.02)</td>
<td>0.46 (0.02)</td>
<td>-0.13 (0.03)</td>
</tr>
<tr>
<td>MOM</td>
<td>8.69 (1.98)</td>
<td>-0.13 (0.04)</td>
<td>0.07 (0.06)</td>
<td>-0.52 (0.08)</td>
<td>0.24 (0.08)</td>
</tr>
<tr>
<td>RISK</td>
<td>-3.40 (1.54)</td>
<td>0.63 (0.03)</td>
<td>0.04 (0.04)</td>
<td>0.05 (0.06)</td>
<td>-0.21 (0.06)</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>3.87 (1.25)</td>
<td>0.19 (0.03)</td>
<td>0.06 (0.03)</td>
<td>0.20 (0.04)</td>
<td>0.19 (0.04)</td>
</tr>
</tbody>
</table>

#### (b) Weekly

<table>
<thead>
<tr>
<th>Intercept</th>
<th>Mkt</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>coef</strong></td>
<td><strong>se</strong></td>
<td><strong>coef</strong></td>
<td><strong>se</strong></td>
<td><strong>coef</strong></td>
<td><strong>se</strong></td>
</tr>
<tr>
<td>SMB</td>
<td>3.63 (1.13)</td>
<td>0.00 (0.01)</td>
<td>0.01 (0.02)</td>
<td>-0.41 (0.02)</td>
<td>-0.07 (0.03)</td>
</tr>
<tr>
<td>HML</td>
<td>1.54 (0.97)</td>
<td>-0.00 (0.01)</td>
<td>0.00 (0.02)</td>
<td>-0.08 (0.02)</td>
<td>0.84 (0.02)</td>
</tr>
<tr>
<td>RMW</td>
<td>4.61 (0.84)</td>
<td>-0.09 (0.01)</td>
<td>-0.23 (0.01)</td>
<td>-0.06 (0.02)</td>
<td>-0.06 (0.02)</td>
</tr>
<tr>
<td>CMA</td>
<td>2.80 (0.65)</td>
<td>-0.12 (0.01)</td>
<td>-0.02 (0.01)</td>
<td>0.38 (0.01)</td>
<td>-0.03 (0.01)</td>
</tr>
<tr>
<td>MOM</td>
<td>8.23 (1.69)</td>
<td>-0.07 (0.02)</td>
<td>0.05 (0.03)</td>
<td>-0.43 (0.03)</td>
<td>0.04 (0.04)</td>
</tr>
<tr>
<td>RISK</td>
<td>-4.19 (1.39)</td>
<td>0.72 (0.01)</td>
<td>0.01 (0.02)</td>
<td>0.17 (0.03)</td>
<td>-0.07 (0.03)</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>4.17 (1.11)</td>
<td>0.17 (0.01)</td>
<td>0.05 (0.02)</td>
<td>0.21 (0.02)</td>
<td>0.14 (0.02)</td>
</tr>
</tbody>
</table>
To illustrate the difference in power further, we repeat this exercise for the five-by-five portfolios from Ken French’s data library that double sort on size and book-to-market, profitability, and investment. We focus on these three, because daily returns are available, and the sorting variables are updated monthly, making comparisons among the horizons valid. We focus on the top and bottom sets of five portfolios, where we expect to find alphas different from zero, or 30 in all. For each of these 30 portfolios, we conduct a time-series test using the Fama-French five-factor model as in Table 3.2, excluding the factor of interest, recording the absolute annualized alpha coefficient, the annualized standard error, the absolute t-statistic, and the p-value for the resulting alpha, using first daily returns, then weekly, and then monthly returns. We plot the cumulative distribution of these four values in the four panels of Figure 3.4. The absolute annualized coefficient ranges from 0.01 percent to 22.8 percent, the annualized standard error of the coefficient ranges from 0.09 percent to 0.51 percent, the absolute t-statistic ranges from zero to 25.3, and the p-value ranges from zero to 1.0. Daily alphas reject the null of zero alpha relative to the Fama-French 5-factor model 63.3 percent of the time, relative to the 5% size of the test. Meanwhile, weekly alphas reject 56.7 percent. And, monthly alphas reject only 40.0 percent. There are far more novel asset pricing
anomalies to be discovered in daily data.

Figure 3.4: Monthly, Weekly, and Daily TS Test Statistics

Why not then simply use the more powerful daily, or better yet intraday, returns? There are two reasons why the literature has tended to use monthly returns, beyond the ease of computation. One is that higher frequency covariances may understate the true, tradeable covariances, because of asynchronous correlation. For example, if some stocks simply do not trade every day, or trade in a way that is slow to incorporate aggregate information, then the estimated betas to the Fama-French five factors will be biased downward in daily regressions. Rather than use monthly returns, Scholes and Williams (1977) recommend aggregating lagged and leading covariances, effectively moving toward weekly regressions. The other is that daily returns may understate or overstate the tradeable annual returns of a given
anomaly, as emphasized by Liu and Strong (2008). This is a closely related point. Instead of asynchronous correlation between a given anomaly and the Fama-French portfolios, this is the autocorrelation of anomaly returns. Positive autocorrelation, or anomaly momentum, means that annualized monthly returns are higher than annualized daily returns. Negative anomaly autocorrelation, or anomaly reversal, means that annualized monthly returns are lower than annualized daily returns. Both of these effects can cause biased inference in identifying tradeable anomalies, moving from monthly to more powerful daily regressions. Both of these are reasons that the analysis of shorter horizons might bias inference, but it is important to note that shifting to much longer horizons comes at the expense of power.

Fortunately, it is possible to split these effects apart and make a modest suggestion for best practice—at least if the anomalies portfolio from Ken French’s data library that we analyzed in Figure 3.4 are a useful guide. The p-value results in panel D of Figure 3.4 come from two separate effects. The p-value is higher because of higher power—that is the reduction in average annualized standard errors in Panel B. In addition, the p-value is higher because of potential bias from asynchronous correlations and factor momentum and reversal—that is reflected in the increase the average annualized absolute coefficient in Panel A. What is the right tradeoff between power and bias? To answer this question, we plot the coefficients and standard errors in Figure 3.5. We scale the annualized coefficient estimate for each horizon by the 50-day annualized coefficient estimate, on the argument that the bias arising from trading effects is negligible at that point. We scale the annualized standard error by the 1-day estimate, where power is maximized. Panel A of Figure 3.5 shows the averages of these scaled values across our 30 portfolios as a function of return horizon, from daily returns to overlapping 50-trading-day returns. We repeat the exercise using medians instead of averages in Panel B.
Figure 3.5: Bias and Power, Annualized Alphas By Return Horizon

Note that the average (and median) coefficient decreases as we move from daily to ten days and then levels off. This in principle might reflect biased inference, though it could also come from a more accurate estimation of covariances. Meanwhile, the average (and median) standard error continues to rise from daily horizons to 50-day horizons and beyond. If the data from Ken French’s data library are suggestive of typical anomalies, our analysis suggests that the problem of bias for these factors appears to be largely solved at 10 trading days, or
two weeks, and completely solved by 20 days. At ten days, the maximum loss in power is 11.6 percent, which still represents a considerable loss of power in time series tests relative to cross-sectional tests, but it is more modest than in the standard practice of analyzing monthly returns, with a loss of 22.4 percent. In what follows, we continue to use monthly returns, as the standard in the literature, but an important conclusion is that, by shifting from monthly to two-week returns, there appears to be a free increase in the power of standard asset pricing tests.

### 3.3.2 Small Sample Power Curves

The power curves in Figure 3.3 assume that all of the distributional parameters in the return generating process in Equation 3.1 are known by the investor. In practice, they are not. Sample estimates must replace their respective population values. We estimate the small sample properties of Equations 3.16 and 3.17 by running Monte Carlo simulations.

![Small Sample and Simulated Power Curve of the CS and TS Test](image)

**Figure 3.6:** Small Sample and Simulated Power Curve of the CS and TS Test

**Note:** The annual Sharpe ratio of the test factor is calculated as $SR = \frac{\mu_b}{\sigma_b} \cdot \sqrt{12}$.

In the base case, we consider $N = 250$ securities and $T = 50$ periods. We use the seven factor portfolios discussed in last section. We first draw factor coefficients $\Gamma_i$ from normal
distribution $N(0, 1)$ for each stock $i$. Then we draw factor returns from a multivariate normal distribution $N(\hat{\mu}, \hat{\Sigma})$ with mean and variance estimated from US market data. Finally, we draw the idiosyncratic return from a normal distribution $N(0, \hat{\sigma}_\varepsilon^2)$ with variance estimated from US market data $\hat{\sigma}_\varepsilon = 3.27\%$ and calculate return of each stock. We repeat this Monte-Carlo simulation for 2000 times. We plot the simulated power curves against their theoretical values in Figure 3.6 and confirm they are consistent with each other. Then we plot the theoretical small sample distribution in Figures 3.7, 3.8 and 3.9.

The patterns we see in the asymptotic power curves also hold in the small sample power curves. Power of both tests get stronger as there are more securities, longer time periods, smaller variances of factor and idiosyncratic returns. The number of time periods and the test factor’s variance have a more significant effect on the power of both tests. The Sharpe ratio drives the wedge between power of the cross section and time series tests. The difference in power for small sample is 23.7\% at $SR = 0.69$, which is larger than that for asymptotic result, 22.4\%, because of additional loss of degrees of freedom in the time series test.
Figure 3.7: Small Sample Power Curve of the CS Test

Note: The annual Sharpe ratio of the test factor is calculated as \( SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{12} \).
Figure 3.8: Small Sample Power Curve of the TS Test

Note: The annual Sharpe ratio of the test factor is calculated as $SR = \frac{\mu_b}{\sigma_{bs}} \cdot \sqrt{12}$. 

119
Figure 3.9: Small Sample Power Curve of the CS and TS Test at Various Sharpe Ratios of the Incumbent Factor Model

Note: The annual Sharpe ratio of the test factor is calculated as $SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{T}$. 
### Table 3.3: Power Difference between the CS and TS Test

**Note:** The annual Sharpe ratio of the test factor is calculated as $SR = \frac{\mu_b}{\sigma_{bb}} \cdot \sqrt{T}$.

<table>
<thead>
<tr>
<th>$SR = 0.25$</th>
<th>Asymptotic</th>
<th>Small Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>TS</td>
</tr>
<tr>
<td>Base</td>
<td>17.3%</td>
<td>12.2%</td>
</tr>
<tr>
<td>$N = 50$</td>
<td>14.8%</td>
<td>10.7%</td>
</tr>
<tr>
<td>$N = 10,000$</td>
<td>17.9%</td>
<td>12.5%</td>
</tr>
<tr>
<td>$\sigma_x = 7$</td>
<td>17.9%</td>
<td>13.0%</td>
</tr>
<tr>
<td>$\sigma_x = 0$</td>
<td>11.2%</td>
<td>8.8%</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>8.0%</td>
<td>6.8%</td>
</tr>
<tr>
<td>$T = 1,000$</td>
<td>61.8%</td>
<td>41.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$SR = 0.50$</th>
<th>Asymptotic</th>
<th>Small Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>TS</td>
</tr>
<tr>
<td>Base</td>
<td>52.5%</td>
<td>34.3%</td>
</tr>
<tr>
<td>$N = 50$</td>
<td>44.2%</td>
<td>28.6%</td>
</tr>
<tr>
<td>$N = 10,000$</td>
<td>54.5%</td>
<td>35.7%</td>
</tr>
<tr>
<td>$\sigma_x = 7$</td>
<td>54.6%</td>
<td>37.5%</td>
</tr>
<tr>
<td>$\sigma_x = 0$</td>
<td>30.4%</td>
<td>20.8%</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>17.3%</td>
<td>12.2%</td>
</tr>
<tr>
<td>$T = 1,000$</td>
<td>99.5%</td>
<td>93.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$SR = 1.00$</th>
<th>Asymptotic</th>
<th>Small Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>TS</td>
</tr>
<tr>
<td>Base</td>
<td>97.7%</td>
<td>86.0%</td>
</tr>
<tr>
<td>$N = 50$</td>
<td>94.4%</td>
<td>77.8%</td>
</tr>
<tr>
<td>$N = 10,000$</td>
<td>98.2%</td>
<td>87.6%</td>
</tr>
<tr>
<td>$\sigma_x = 7$</td>
<td>98.2%</td>
<td>89.4%</td>
</tr>
<tr>
<td>$\sigma_x = 0$</td>
<td>80.6%</td>
<td>60.8%</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>50.6%</td>
<td>33.0%</td>
</tr>
<tr>
<td>$T = 1,000$</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$SR = 1.50$</th>
<th>Asymptotic</th>
<th>Small Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>TS</td>
</tr>
<tr>
<td>Base</td>
<td>100.0%</td>
<td>99.6%</td>
</tr>
<tr>
<td>$N = 50$</td>
<td>100.0%</td>
<td>98.5%</td>
</tr>
<tr>
<td>$N = 10,000$</td>
<td>100.0%</td>
<td>99.7%</td>
</tr>
<tr>
<td>$\sigma_x = 7$</td>
<td>100.0%</td>
<td>99.8%</td>
</tr>
<tr>
<td>$\sigma_x = 0$</td>
<td>99.0%</td>
<td>92.2%</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>84.8%</td>
<td>63.2%</td>
</tr>
<tr>
<td>$T = 1,000$</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
3.4 Conclusion

What is an anomaly? Empirical asset pricing papers that aim to establish an anomaly often rely on two types of tests. One is the cross-sectional test in the spirit of Fama and MacBeth (1973) and famously used in Fama and French (1992) and the other is the times series test in the spirit of Jensen (1968) and popularized by Fama and French (1993). We consider these tests from two different points of view: relevance and power. The cross-section test is relevant to a risk neutral investor facing a simple form of transaction costs. The time series test is relevant to a risk-averse investor facing no transaction costs. Meanwhile, both are relevant in the more general case of risk aversion and transaction costs. Next, we show that the time series test can be inherently lower powered. Given that most professional investors face meaningful transaction costs and most commercial portfolio optimizers target a mean-variance objective, we believe that a test that passes either of the two tests can be considered an anomaly—in the sense that it is practically relevant for a large class of investors. Viewed in this light, the literature on empirical asset pricing has the potential to identify a richer set of interesting anomalies than are contained in the Fama- French 3-Factor model or even the newer 5-factor model. Unlike the time series analysis in Fama and French (1993), the cross-sectional framework in Fama and French (1992) has a higher upper limit on the number of relevant and statistically significant factors.
References


123


—— (2017b): “Diversity in Innovation,”


Appendix A

Appendix to Chapter 1

A.1 Supplementary Tables and Figures

Figure A.1: Cross-County Variation in Casualty Rate

This figure shows cross-county standard deviation of casualty rate by county population. Green line is the theoretical value, assuming Bernoulli distribution for probability of soldier mortality. Red line is the empirical value for counties of similar population. Source: author’s calculation using military records from the National Archives and data from the US decennial censuses.
Figure A.2: Examples of Historical Documents
Appendix B

Appendix to Chapter 2

B.1 Description of the Sohn Investment Conference

From the conference’s website:

Since 1995, the world-renowned Sohn Investment Conference, held annually in New York, has been the premier investment event, bringing the world’s savviest investors together to share fresh insights and money-making ideas to benefit The Sohn Conference Foundation’s work to end childhood cancer.

Wall Street’s best and brightest investors participate in this unique, “must attend” event to share their expertise with an audience of more than 3,000 people, comprised of portfolio managers, asset allocators and private investors. Most speakers manage large proprietary investment portfolios that have outperformed the market for many years and do not share their insights in any public forum, but they volunteer their time to the Conference for the benefit of the Foundation. All contributions support the Foundation’s mission to support pediatric research and care.

Source: http://www.sohnconference.org/new-york/
Figure B.1: Registration Page of Sohn Conference
### B.2 Supplementary Tables and Figures

**Table B.1: Conference List**

<table>
<thead>
<tr>
<th>Conference</th>
<th>Location</th>
<th>Date</th>
<th>Num Stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sohn</td>
<td>New York, NY</td>
<td>2008-05-21</td>
<td>10</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>New York, NY</td>
<td>2008-10-07</td>
<td>18</td>
</tr>
<tr>
<td>Sohn</td>
<td>New York, NY</td>
<td>2009-05-27</td>
<td>13</td>
</tr>
<tr>
<td>Great Investors’ Best Ideas</td>
<td>Dallas, TX</td>
<td>2009-10-06</td>
<td>14</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>New York, NY</td>
<td>2009-10-19</td>
<td>21</td>
</tr>
<tr>
<td>Invest For Kids</td>
<td>New York, NY</td>
<td>2009-11-12</td>
<td>3</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>Pasadena, CA</td>
<td>2010-05-05</td>
<td>11</td>
</tr>
<tr>
<td>Sohn</td>
<td>New York, NY</td>
<td>2010-05-26</td>
<td>21</td>
</tr>
<tr>
<td>Sohn</td>
<td>San Francisco, CA</td>
<td>2010-10-06</td>
<td>3</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>New York, NY</td>
<td>2010-10-12</td>
<td>9</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>Pasadena, CA</td>
<td>2011-05-04</td>
<td>17</td>
</tr>
<tr>
<td>Sohn</td>
<td>New York, NY</td>
<td>2011-05-25</td>
<td>24</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>New York, NY</td>
<td>2011-10-17</td>
<td>20</td>
</tr>
<tr>
<td>Invest For Kids</td>
<td>Chicago, IL</td>
<td>2011-11-09</td>
<td>15</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>Omaha, NE</td>
<td>2012-05-08</td>
<td>8</td>
</tr>
<tr>
<td>Sohn</td>
<td>New York, NY</td>
<td>2012-05-16</td>
<td>20</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>New York, NY</td>
<td>2012-10-02</td>
<td>26</td>
</tr>
<tr>
<td>Great Investors’ Best Ideas</td>
<td>Dallas, TX</td>
<td>2012-11-01</td>
<td>8</td>
</tr>
<tr>
<td>Invest For Kids</td>
<td>Chicago, IL</td>
<td>2012-11-07</td>
<td>8</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>Las Vegas, NV</td>
<td>2013-05-06</td>
<td>21</td>
</tr>
<tr>
<td>Sohn</td>
<td>New York, NY</td>
<td>2013-05-08</td>
<td>23</td>
</tr>
<tr>
<td>London Value Investor</td>
<td>London, UK</td>
<td>2013-05-09</td>
<td>7</td>
</tr>
<tr>
<td>Virginia Investment Symposium</td>
<td>Charlottesville, VA</td>
<td>2013-06-04</td>
<td>1</td>
</tr>
<tr>
<td>Alpha Hedge West</td>
<td>San Francisco, CA</td>
<td>2013-09-15</td>
<td>5</td>
</tr>
<tr>
<td>Value Investing Congress</td>
<td>New York, NY</td>
<td>2013-09-17</td>
<td>20</td>
</tr>
<tr>
<td>Excellence in Investing</td>
<td>San Francisco, CA</td>
<td>2013-10-23</td>
<td>11</td>
</tr>
<tr>
<td>Great Investors’ Best Ideas</td>
<td>Dallas, TX</td>
<td>2013-10-29</td>
<td>11</td>
</tr>
<tr>
<td>Invest For Kids</td>
<td>Chicago, IL</td>
<td>2013-10-29</td>
<td>12</td>
</tr>
<tr>
<td>Sohn</td>
<td>London, UK</td>
<td>2013-11-01</td>
<td>5</td>
</tr>
<tr>
<td>Fund</td>
<td>Type</td>
<td>Location</td>
<td>Person</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>GREENLIGHT CAPITAL, INC.</td>
<td>Activist</td>
<td>New York, NY</td>
<td>David Einhorn</td>
</tr>
<tr>
<td>PERSHING SQUARE CAP MGMT, L.P.</td>
<td>Activist</td>
<td>New York, NY</td>
<td>Bill Ackman</td>
</tr>
<tr>
<td>MFP INVESTORS, LLC</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Michael Price</td>
</tr>
<tr>
<td>T2 PTNR MANAGEMENT, L.P.</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Whitney Tilson</td>
</tr>
<tr>
<td>NIERENBERG INVESTMENT MGMT CO</td>
<td>Hedge Fund</td>
<td>Camas, WA</td>
<td>David Nierenberg</td>
</tr>
<tr>
<td>VALUEACT CAPITAL MGMT, L.P.</td>
<td>Activist</td>
<td>San Francisco, CA</td>
<td>Jeff Ubben</td>
</tr>
<tr>
<td>ATLANTIC INVESTMENT MGMT, INC.</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Alexander Roepers</td>
</tr>
<tr>
<td>OMEGA ADVISORS, INC.</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Lee Cooperman</td>
</tr>
<tr>
<td>GLENVIEW CAPITAL MGMT, LLC</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Larry Robbins</td>
</tr>
<tr>
<td>KYNIKOS ASSOCIATES LP</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Jim Chanos</td>
</tr>
<tr>
<td>EMRYS PARTNERS, LP</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Steve Eisman</td>
</tr>
<tr>
<td>FRONTPOINT PTNR LLC</td>
<td>Hedge Fund</td>
<td>Greenwich, CT</td>
<td>Steve Eisman</td>
</tr>
<tr>
<td>MARCATO CAPITAL MANAGEMENT LP</td>
<td>Hedge Fund</td>
<td>San Francisco, CA</td>
<td>Mick McGuire</td>
</tr>
<tr>
<td>CENTAUR CAPITAL PARTNERS, L.P.</td>
<td>Hedge Fund</td>
<td>Southlake, TX</td>
<td>Zeke Ashton</td>
</tr>
<tr>
<td>GOTHAM ASSET MANAGEMENT, LLC</td>
<td>Mutual Fund</td>
<td>New York, NY</td>
<td>Joel Greenblatt</td>
</tr>
<tr>
<td>KEELEY ASSET MANAGEMENT CORP.</td>
<td>Investment Adviser</td>
<td>Chicago, IL</td>
<td>John Keeley</td>
</tr>
<tr>
<td>LONE PINE CAPITAL, L.L.C.</td>
<td>Hedge Fund</td>
<td>Greenwich, CT</td>
<td>Steve Mandel</td>
</tr>
<tr>
<td>PACIFICA CAPITAL INVTS LLC</td>
<td>Investment Adviser</td>
<td>Leander, TX</td>
<td>Steve Leonard</td>
</tr>
<tr>
<td>BP CAPITAL MANAGEMENT L P</td>
<td>Hedge Fund</td>
<td>Dallas, TX</td>
<td>T. Boone Pickens</td>
</tr>
<tr>
<td>EAGLE CAPITAL MGMT L.L.C. (NY)</td>
<td>Investment Adviser</td>
<td>New York, NY</td>
<td>Boykin Curry</td>
</tr>
<tr>
<td>TIGER MANAGEMENT, LLC</td>
<td>Family Office</td>
<td>New York, NY</td>
<td>Julian Robertson</td>
</tr>
<tr>
<td>GABELLI FUNDS, LLC</td>
<td>Mutual Fund</td>
<td>Rye, NY</td>
<td>Daniel Miller</td>
</tr>
<tr>
<td>PERELLA WEINBERG PTNR CAP MGMT</td>
<td>Investment Adviser</td>
<td>New York, NY</td>
<td>Dan Arbess</td>
</tr>
<tr>
<td>HEALTHCOR MANAGEMENT, L.P.</td>
<td>Hedge Fund</td>
<td>New York, NY</td>
<td>Joseph Healey</td>
</tr>
<tr>
<td>FAIRHOLME CAPITAL MGMT, L.L.C.</td>
<td>Investment Adviser</td>
<td>Miami, FL</td>
<td>Bruce Berkowitz</td>
</tr>
<tr>
<td>BOYAR ASSET MANAGEMENT INC</td>
<td>Investment Adviser</td>
<td>New York, NY</td>
<td>Mark Boyar</td>
</tr>
<tr>
<td>JANA PTNR LLC</td>
<td>Activist</td>
<td>New York, NY</td>
<td>Barry Rosenberg</td>
</tr>
<tr>
<td>STARWOOD CAPITAL GROUP LLC</td>
<td>PE/VC</td>
<td>Greenwich, CT</td>
<td>Berry Sternlicht</td>
</tr>
<tr>
<td>BLUE HARBOUR GROUP, L.P.</td>
<td>Hedge Fund</td>
<td>Greenwich, CT</td>
<td>Clifton Robbins</td>
</tr>
<tr>
<td>WEST COAST ASSET MGMT, INC.</td>
<td>Investment Adviser</td>
<td>Montecito, CA</td>
<td>Lance Helfert</td>
</tr>
<tr>
<td>Fund</td>
<td>N(Pitches)</td>
<td>13F Amt (§MM)</td>
<td>N(Holdings)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GREENLIGHT CAPITAL, INC.</td>
<td>16</td>
<td>7,228.1</td>
<td>39</td>
</tr>
<tr>
<td>PERSHING SQUARE CAP MGMT, L.P.</td>
<td>15</td>
<td>16,045.9</td>
<td>7</td>
</tr>
<tr>
<td>MFP INVESTORS, LLC</td>
<td>14</td>
<td>812.4</td>
<td>88</td>
</tr>
<tr>
<td>T2 PTNR MANAGEMENT, L.P.</td>
<td>13</td>
<td>53.8</td>
<td>14</td>
</tr>
<tr>
<td>NIERENBERG INVESTMENT MGMT CO</td>
<td>13</td>
<td>75.1</td>
<td>6</td>
</tr>
<tr>
<td>VALUEACT CAPITAL MGMT, L.P.</td>
<td>11</td>
<td>15,133.7</td>
<td>13</td>
</tr>
<tr>
<td>ATLANTIC INVESTMENT MGMT, INC.</td>
<td>10</td>
<td>1,580.9</td>
<td>16</td>
</tr>
<tr>
<td>OMEGA ADVISORS, INC.</td>
<td>10</td>
<td>5,378.8</td>
<td>85</td>
</tr>
<tr>
<td>GLENVIEW CAPITAL MGMT, LLC</td>
<td>10</td>
<td>18,500.3</td>
<td>68</td>
</tr>
<tr>
<td>KYNIKOS ASSOCIATES LP</td>
<td>9</td>
<td>214.2</td>
<td>25</td>
</tr>
<tr>
<td>EMRYS PARTNERS, LP</td>
<td>9</td>
<td>108.9</td>
<td>26</td>
</tr>
<tr>
<td>FRONTPOINT PTNR LLC</td>
<td>9</td>
<td>8.9</td>
<td>6</td>
</tr>
<tr>
<td>MARCATO CAPITAL MANAGEMENT LP</td>
<td>8</td>
<td>3,058.9</td>
<td>10</td>
</tr>
<tr>
<td>CENTAUR CAPITAL PARTNERS, L.P.</td>
<td>7</td>
<td>52.0</td>
<td>21</td>
</tr>
<tr>
<td>GOTHAM ASSET MANAGEMENT, LLC</td>
<td>7</td>
<td>12,059.1</td>
<td>908</td>
</tr>
<tr>
<td>KEELEY ASSET MANAGEMENT CORP.</td>
<td>6</td>
<td>4,726.6</td>
<td>334</td>
</tr>
<tr>
<td>LONE PINE CAPITAL, L.L.C.</td>
<td>6</td>
<td>20,539.5</td>
<td>38</td>
</tr>
<tr>
<td>PACIFICA CAPITAL INVTS LLC</td>
<td>6</td>
<td>151.9</td>
<td>14</td>
</tr>
<tr>
<td>BP CAPITAL MANAGEMENT L P</td>
<td>6</td>
<td>59.9</td>
<td>16</td>
</tr>
<tr>
<td>EAGLE CAPITAL MGMT L.L.C. (NY)</td>
<td>5</td>
<td>23,415.1</td>
<td>56</td>
</tr>
<tr>
<td>TIGER MANAGEMENT, LLC</td>
<td>5</td>
<td>422.0</td>
<td>51</td>
</tr>
<tr>
<td>GABELLI FUNDS, LLC</td>
<td>5</td>
<td>19,592.7</td>
<td>1045</td>
</tr>
<tr>
<td>PERELLA WEINBERG PTNR CAP MGMT</td>
<td>4</td>
<td>377.6</td>
<td>65</td>
</tr>
<tr>
<td>HEALTHCOR MANAGEMENT, L.P.</td>
<td>3</td>
<td>1,501.1</td>
<td>44</td>
</tr>
<tr>
<td>FAIRHOLME CAPITAL MGMT, L.L.C.</td>
<td>3</td>
<td>6,420.6</td>
<td>11</td>
</tr>
<tr>
<td>BOYAR ASSET MANAGEMENT INC</td>
<td>3</td>
<td>127.9</td>
<td>63</td>
</tr>
<tr>
<td>JANA PTNR LLC</td>
<td>3</td>
<td>6,713.1</td>
<td>18</td>
</tr>
<tr>
<td>STARWOOD CAPITAL GROUP LLC</td>
<td>3</td>
<td>374.3</td>
<td>4</td>
</tr>
<tr>
<td>BLUE HARBOUR GROUP, L.P.</td>
<td>3</td>
<td>2,751.5</td>
<td>19</td>
</tr>
<tr>
<td>WEST COAST ASSET MGMT, INC.</td>
<td>3</td>
<td>53.4</td>
<td>29</td>
</tr>
</tbody>
</table>
Figure B.2: Distribution of Money Raised
C.1 Extensions to the Basic Model

C.1.1 Variable Transaction Costs

So far, we have considered the case where every stock has the same level of transaction costs $\theta$. In reality, producing a factor portfolio with a given exposure to a particular set of scores varies in terms of its execution cost. This is in part because the extreme scores may tilt towards smaller, more volatile, or less liquid securities, which have a higher cost $\theta_i$ than those in the middle. The change in our investor’s weights is slight, as we replace $\theta$ with a vector of transaction costs $\theta$.

\[ \mathbf{w}^{*}_{vtc} = \text{diag}(\mathbf{\theta})^{-1} \mathbf{\Gamma} \mathbf{\mu} \]

The stock level weights are simple, and intuitive. Instead of scaling the expected return by a common transaction cost $\theta$, each stock has its own stock-specific transaction cost scalar $\theta_i$. In the case of simple transaction costs in Equation 3.6, the stock-level logic extends to factors. The exposure to all factor portfolios scales up or down with a single transaction cost parameter. The factor exposures with variable transaction costs are more complicated. To simplify notation, we assume again that the factors are defined to be orthogonal so that $\mathbf{\Gamma}^T = N \mathbf{I}$ and the raw and unit exposures are the same. Also, we suppose for simplicity that all of the factors are structured so that the mean return $\mathbf{\mu}$ is positive. This is not necessary,
but it serves to simplify the notation and develop intuition.

\[
\mathbf{e}_{\text{raw}}(\mathbf{w}^*_\text{vtc}) = \mathbf{e}_{\text{unit}}(\mathbf{w}^*_\text{vtc}) = \Gamma' \text{diag}(\theta)^{-1} \mathbf{\Gamma} \mu \\
= \frac{1}{N} \begin{bmatrix}
\sum_i \frac{\gamma^2_{ai}}{\theta_i} & \sum_i \frac{\gamma_{ai} \cdot \gamma_{bi}}{\theta_i} & \cdots \\
\sum_i \frac{\gamma_{ai} \cdot \gamma_{ai}}{\theta_i} & \sum_i \frac{\gamma^2_{bi}}{\theta_i} & \cdots \\
\vdots & \vdots & \ddots
\end{bmatrix} \mu
\]

Our investor chooses higher exposure to a factor when its return \(\mu_a\) is high relative to its execution cost. Execution costs are low when extreme scores are negatively correlated with transaction costs. In other words, exposure is higher on a factor portfolio if the securities with the highest absolute scores, either positive or negative, so that \(\gamma_{ai}^2\) is large, also happen to be easier to trade, with low \(\theta_i\). There is a second, subtler force in the off-diagonal terms. Exposure is higher when the factor of interest has scores that are correlated with other high return \(\mu_b\) factors, so that \(\gamma_{ai} \cdot \gamma_{bi}\) is large, among stocks that are easier to trade, with low \(\theta_i\). This is an echo of the unit portfolio exposures in the case of a risk neutral investor with constant transaction costs and correlated scores. Here the scores are by assumption uncorrelated, but incidental positive or negative exposures can nonetheless arise if there is a conditional positive or negative score correlation in the set of stocks that are cheaper to trade. This is one reason to prefer a single multi-factor portfolio optimization to a portfolio of factor portfolios, because it takes into account execution savings across factor portfolios.

The factor exposures can be further simplified in the case where the scores are also uncorrelated conditional on transaction costs, so that \(E(\gamma_{ai} \cdot \gamma_{bi}|\theta) = 0\). That eliminates any spillover from one factor exposure to another:

\[
\mathbf{e}_{\text{raw}}(\mathbf{w}^*_\text{vtc}) = \mathbf{e}_{\text{unit}}(\mathbf{w}^*_\text{vtc}) = \begin{bmatrix}
\frac{1}{N} \sum_i \frac{\gamma^2_{ai}}{\theta_i} \mu_a \\
\frac{1}{N} \sum_i \frac{\gamma^2_{ai}}{\theta_i} \mu_b \\
\vdots
\end{bmatrix}
\]

(C.1)

The conclusion is that our investor uses a simple formula to adjust expected returns for execution costs, which here are increasing in the correlation of trading costs with extreme scores. This will shift the tests in Equation 3.7 and 3.11, replacing the gross return with the
net return above that adjusts for transaction costs. For the simple case of uncorrelated scores, the cross section test of statistical significance is unchanged.

**Example 5**

Consider an example of two factors, using low asset growth and operating profitability once again. Suppose that these two factors are constructed so that they have zero correlation, even conditional on transaction costs. Also assume that the extreme asset growth stocks, those that are growing very fast or contracting significantly, are less liquid and more costly to trade. Moreover, assume that firms with high operating profits are more liquid with lower cost to trade, so that $\sum \frac{\gamma^2_i}{\theta_i} > \sum \frac{\gamma^2_i}{\theta_i}$, which is roughly consistent with US data. The optimal exposures to the unit or raw factor portfolios, given that they are unconditionally uncorrelated, are:

$$e_{raw}(w_{vtc}^*) = e_{unit}(w_{vtc}^*) = \begin{bmatrix} e_{unit, -A} \\ e_{unit, OP} \end{bmatrix} = \begin{bmatrix} \frac{1}{N} \sum \frac{\gamma^2_i}{\theta_i} \mu_a \\ \frac{1}{N} \sum \frac{\gamma^2_i}{\theta_i} \mu_b \end{bmatrix}$$

The optimal portfolio has positive exposure to both the unit low asset growth portfolio and the unit operating profits portfolio. If they have the same gross factor portfolio return $\mu_a = \mu_b$, then unit exposure to the operating profits portfolio is higher because its execution costs are lower, making its net factor portfolio returns higher.

**C.1.2 Capacity Constraints**

So far, we have sidestepped the issue of the capacity of the factor portfolios. Our investor has constant absolute risk aversion and no constraints on leverage, so there are no natural effects of assets under management, which we label $A$. A practical way of making capacity relevant is to add a few changes to the basic set up in Equations 3.1 and 3.4. We first imagine that returns in Equation 3.1 are defined relative to a benchmark, so that all of the factors have zero mean. Second, we suppose that our investor delegates his portfolio decision, while insisting on some fixed level of gross exposure, a fixed tracking error, or a minimum level of benchmark-adjusted return. For example, our investor might ask for a dollar neutral portfolio, where, for each dollar of equity, one dollar must be invested long and one dollar must be
invested short. Or, as we do in Appendix B, we solve a typical case of fixed tracking error \( \text{var}(r) \mathbf{w} = \sigma_r^2 \). We proceed here with a further simplified, risk neutral case where the exposure constraint is expressed as \( \mathbf{w}' \mathbf{w} = 1 \).

\[
\text{max} \ (\mathbf{\mu} \mathbf{\Gamma})' \mathbf{w} - \frac{1}{2} \mathbf{w}' \text{diag}(\mathbf{\Gamma}) \mathbf{w} \text{ s.t. } \mathbf{w}' \mathbf{w} = 1
\]

(C.2)

Our investor’s new optimal weights a slight variation of the unconstrained version:

\[
\mathbf{w}_{\text{aum}}^* = (A \cdot \text{diag}(\mathbf{\theta}) + 2C(\mathbf{A}) \mathbf{I})^{-1} \mathbf{\Gamma} \mathbf{\mu}
\]

(C.3)

The Lagrange multiplier is \( C(\mathbf{A}) \), which satisfies \( (\mathbf{\Gamma} \mathbf{\mu})' (A \cdot \text{diag}(\mathbf{\theta}) + 2C(\mathbf{A}) \mathbf{I})^{-2} (\mathbf{\Gamma} \mathbf{\mu}) = 1 \) and \( \frac{dC}{dA} < 0 \) as we show in Appendix B. Again, the weight on any given security is limited by its security specific transaction costs in the first term in parenthesis. Now, there is also a second consideration. We explore these effects of assets under management on the factor portfolio exposures. To simplify notation, we assume again that the factors are defined to be orthogonal so that \( \mathbf{\Gamma}' \mathbf{\Gamma} = N \mathbf{I} \).

\[
\mathbf{e}_{\text{raw}}(\mathbf{w}_{\text{aum}}^*) = \mathbf{e}_{\text{unit}}(\mathbf{w}_{\text{aum}}^*) = \frac{1}{N} \mathbf{\Gamma}' (A \cdot \text{diag}(\mathbf{\theta}) + C(\mathbf{A}) \mathbf{I})^{-1} \mathbf{\Gamma} \mathbf{\mu}
\]

\[
= \frac{1}{N} \begin{pmatrix}
\sum_i \frac{\gamma_a^2 a_i}{\mathbf{A} \mathbf{\theta}_i + C(\mathbf{A})} & \sum_i \frac{\gamma_a \gamma_b b_i}{\mathbf{A} \mathbf{\theta}_i + C(\mathbf{A})} & \cdots \\
\sum_i \frac{\gamma_b \gamma_a a_i}{\mathbf{A} \mathbf{\theta}_i + C(\mathbf{A})} & \sum_i \frac{\gamma_b^2 b_i}{\mathbf{A} \mathbf{\theta}_i + C(\mathbf{A})} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix} \mathbf{\mu}
\]

Our investor again chooses higher exposure to a factor when its return \( \mu_a \) is high relative to its execution cost. Execution costs are low when extreme scores \( \gamma_a^2 \) are negatively correlated with execution costs. But execution costs themselves are now a function of assets under management, with \( \theta_i \) replaced by \( A \theta_i + C(\mathbf{A}) \). As before, the factor exposures can be further simplified in the case where the scores are also uncorrelated conditional on transaction costs, so that \( E(\gamma_a \cdot \gamma_b | \mathbf{\theta}) = 0 \):

\[
\mathbf{e}_{\text{raw}}(\mathbf{w}_{\text{aum}}^*) = \mathbf{e}_{\text{unit}}(\mathbf{w}_{\text{aum}}^*) = \begin{pmatrix}
\frac{1}{N} \sum_i \frac{\gamma_a^2 a_i}{A \mathbf{\theta}_i + C(\mathbf{A})} \mu_a \\
\frac{1}{N} \sum_i \frac{\gamma_b^2 b_i}{A \mathbf{\theta}_i + C(\mathbf{A})} \mu_b \\
\vdots
\end{pmatrix}
\]
The key insight here is that $\frac{dC}{dA} < 0$ and $C$ does not vary across stocks $i$. So, for very small assets under management $A$, the factor allocation is directly proportional to return $\mu$ as in Equation 3.6. The gross exposure constraint binds so quickly that transaction costs have no effect. As assets rise, the importance of transaction costs $\theta_i$ increases to the point that the allocations are proportional to those in Equation C.1. The conclusion is that our investor uses a formula to adjust expected returns for execution costs, but this formula is highly dependent on assets under management, so the exact adjustment of the tests in Equation 3.7 and 3.11, which replace gross returns with net returns are context specific. An anomaly for one investor may not be an economically meaningful anomaly for another, because its execution costs are too high at the relevant level of assets $A$. Again, for the simple case of uncorrelated scores, the cross section test of statistical significance is unchanged.

C.1.3 Dynamic Trading

So far, we have considered a static trading decision. A fully dynamic optimization with factor scores $\Gamma$ that vary through time is beyond the scope of this paper. We analyze a very simple case where our risk neutral investor trades over a finite number of periods $T$. The assumption of risk neutrality is somewhat awkward, because our investor will simply accumulate positions over time, with the per period accumulation limited by transaction costs. But, this situation still provides the intuition that the investor in the first period will need to consider the returns not just in the next period but in subsequent periods, with a discount rate $\delta$. More persistent factors deserve greater weight, because they will generate returns over multiple future holding periods.

$$\max_{\Gamma_t} \sum_{t=0}^{T} \delta^t \left( (\mu \Gamma_t)' w_t - \frac{1}{2} (w_t - w_{t-1})' \text{diag}(\rho) (w_t - w_{t-1}) \right)$$

For simplicity, we imagine that the factor payoffs and transaction costs are constant through time, but the factor scores vary according to an autoregressive process:

$$\Gamma_t = \Gamma_{t-1} \text{diag}(\rho_{K \times 1})_{K \times K} + \Lambda_t$$

This is straightforward to solve by backward induction, with the caveat that without
a budget constraint or risk aversion the weights accumulate through time. Garleanu and Pedersen (2013) also arrive at a similarly interpretable solution with per period risk aversion by assuming that transaction costs are proportional to the stock level covariance matrix.

$$w_0 = \text{diag} (\theta)^{-1} \Gamma_0 \left( \sum_{s=0}^{T} (\delta \text{diag} (\rho))^s \right) \mu$$

Again, we are interested in the resulting factor exposures of our investor’s initial portfolio, and we make the same simplifying assumption that $$\Gamma'\Gamma = NI$$ so that the raw and unit exposures are the same.

$$e_{\text{raw}}(w_{\text{it}}^*) = e_{\text{unit}}(w_{\text{it}}^*) = \Gamma' \text{diag} (\theta)^{-1} \Gamma \left( \sum_{s=0}^{T} (\delta \text{diag} (\rho))^s \right) \mu$$

$$= \frac{1}{N} \begin{bmatrix}
\sum_{i} \frac{\gamma_i^2}{\theta_i} & \sum_{i} \frac{\gamma_{ai} \gamma_{bi}}{\theta_i} & \cdots \\
\sum_{i} \frac{\gamma_{ai} \gamma_{ai}}{\theta_i} & \sum_{i} \frac{\gamma_{ai}^2}{\theta_i} & \cdots \\
\vdots & \vdots & \ddots
\end{bmatrix} \begin{bmatrix}
\frac{1}{1-\delta \rho a} \mu_a \\
\frac{1}{1-\delta \rho b} \mu_b \\
\vdots
\end{bmatrix}$$

Our investor chooses higher exposure to a factor when the full present value of its future returns $$\frac{1}{1-\delta \rho a} \mu_a$$ is high relative to its execution cost. Execution costs are low as before, but now the benefits of trade extend beyond one period. The factor exposures can again be further simplified in the case where the scores are also uncorrelated conditional on transaction costs, so that $$E (\gamma_{ai} \cdot \gamma_{bi} | \theta) = 0$$. That eliminates any spillover from one factor exposure to another:

$$e_{\text{raw}}(w_{\text{itc}}^*) = e_{\text{unit}}(w_{\text{itc}}^*) = \frac{1}{N} \begin{bmatrix}
\sum_{i} \frac{\gamma_i^2}{\theta_i} \frac{1}{1-\delta \rho a} \mu_a \\
\sum_{i} \frac{\gamma_{ai}^2}{\theta_i} \frac{1}{1-\delta \rho b} \mu_b \\
\vdots
\end{bmatrix}$$

(C.4)

As an aside, the exposure of persistent factors increases through time, because the portfolio inherits the effects of past decisions. Nonetheless, the relevant information for our investor on a forward-looking is the execution-cost-adjusted and persistence-adjusted mean returns in Equation [C.4]. Layering on capacity constraints does not produce nicely interpretable exposures, but the effect of combining capacity constraints with dynamic trading is to lower the exposure of costly-to-trade factors in particular when assets under management are high, and when these factors are less persistent. As in the previous two subsections, for the simple
case of uncorrelated scores, the cross section test of statistical significance is unchanged.

Example 6

Consider an example of two factors, using operating profitability as before and using high frequency reversal instead of low asset growth. Jegadeesh and Titman (1993), among others have observed that the firms with the highest returns in the previous month have lower average returns in the month that follows. So, a stock with a high trailing one-month return might then have a low high frequency reversal score $\gamma_{a,i}$. Again, suppose that these two factors are constructed so that they have zero correlation with one another. For this particular example, we assume that the sets of stocks with high and low operating profitability are very persistent, so that there is a monthly mean persistence of $\rho_a = 0.98$ in factor returns on some initial set of scores $\gamma_b$. And, we assume that high frequency reversal scores are essentially uncorrelated through time, with monthly persistence of $\rho_b = 0$. This is roughly consistent with US data. Finally, we use a monthly discount rate $\delta = 0.9$, meaning that the investor’s portfolio might fully turnover about once per year. The dynamic target, optimal exposures to the unit or raw factor portfolios are:

$$
e_{raw}(w^*_{vtc}) = e_{unit}(w^*_{vtc}) = 
\begin{bmatrix}
e_{unit,r_{t-1}} \\
e_{unit,OP}
\end{bmatrix} = 
\begin{bmatrix}
\frac{1}{N} \sum_i \frac{\gamma^2_{a,i}}{\theta_i} \left(\frac{1}{1 - 0.9 \cdot 0.98} \mu_a \right) \\
\frac{1}{N} \sum_i \frac{\gamma^2_{a,i}}{\theta_i} \mu_b
\end{bmatrix}$$

The optimal portfolio has positive exposure to both the unit high frequency reversal portfolio and the unit operating profits portfolio. If they have the same gross, per period factor portfolio return $\mu_a = \mu_b$, then exposure to operating profits is approximately 8.5 times higher because of its much higher persistence.
C.2 Derivations

C.2.1 Derivation of Equation 3.11

Here, we would like to prove that Equation 3.11 is equivalent to Equation 3.10:

$$\sum_{k \neq a} \hat{\sigma}_{ak} e_{unit,k} (w^*_{ra}) = \hat{\mu}_a$$

First, we substitute in the optimal unit exposure from Equation 3.9 and rearrange the terms, putting $\hat{\mu}_a$ on the left-hand side.

$$\hat{\mu}_a = 2\lambda (\hat{\sigma}_{ab} e_b + \hat{\sigma}_{ac} e_c + \ldots)$$

Second, we note that Equation 3.11 can be rewritten as follows:

$$\hat{\mu}_a = \sum_{k \neq a} \hat{\beta}_{ak} \hat{\mu}_k \Rightarrow \hat{\mu}_a = \sum_{k \neq a} \hat{\beta}_{ak} 2\lambda \left( \sum_{l=b} \hat{\sigma}_{kl} e_l \right)$$

Now, it remains to show that the right hand side of these two previous equations are the same. This will be true if the following holds for all factors $l$ not equal to the test factor $a$:

$$\forall l, \hat{\sigma}_{al} = \sum_{k \neq a} \hat{\beta}_{ak} \hat{\sigma}_{kl}$$

This identity is simply the identity that the covariance of the sum is equal to the sum of covariances:

$$\text{cov}(f_a, f_l) = \text{cov} \left( \sum_{k=b} \hat{\beta}_{ak} f_k, f_l \right) = \sum_{k=b} \hat{\beta}_{ak} \text{cov}(f_k, f_l)$$

C.2.2 Solution for Equation C.2 with a Budget Constraint

Here, we would like to derive the optimal weights for an investor who faces variable transaction costs with variable assets under management. The general optimization problem is repeated from Equation C.2 above:

$$\max (\mu^\prime) w - \frac{A}{2} w^\prime \text{diag}(\theta) w \text{ s.t. } w^\prime w = 1$$

We solve the Lagrangian, where we use the parameter C as the Lagrange multiplier, noting
along the way that $C$ is a function of assets under management $A$,

$$L = (\mu \Gamma)' w - \frac{A}{2} w' \text{diag} (\theta) w - C (A) (w'w - 1) \quad (C.5)$$

by taking the first order condition with respect to portfolio weights:

$$\frac{\partial L}{\partial w} = 0 \quad (C.6)$$

$$\Rightarrow \mu \Gamma - (A \cdot \text{diag} (\theta) + 2C (A) I) w = 0 \quad (C.7)$$

This gives us a solution for $w$:

$$w = (A \cdot \text{diag} (\theta) + 2C (A) I)^{-1} \mu \Gamma \quad (C.8)$$

### C.2.3 Derivation of $\frac{\partial C(A)}{\partial A} < 0$ in Equation C.3

Here, we would like to show that the Lagrange multiplier, $C(A)$, from the previous section, and in Equation C.3 is decreasing in assets under management, $A$. The FOC can rewritten as follows:

$$(\Gamma \mu)' (A \cdot \text{diag} (\theta) + 2C (A) I)^{-2} (\Gamma \mu) = 1 \quad (C.9)$$

$$\Rightarrow \mu_R' (A \cdot \text{diag} (\theta) + 2C (A) I)^{-2} \mu_R = 1 \quad (C.10)$$

Because $A \cdot \text{diag} (\theta) + 2C (A) I$ is an diagonal matrix, its inverse is still an diagonal matrix with inverses of corresponding diagonal elements. We can then express the squared inverse as follows:

$$(A \cdot \text{diag} (\theta) + 2C (A) I)^{-2} = \begin{bmatrix}
\frac{1}{(A\theta_1 + 2C(A))^2} \\
\frac{1}{(A\theta_2 + 2C(A))^2} \\
\vdots
\end{bmatrix}$$

Substituting for the squared inverse, simplifying, performing matrix multiplication, gives
us the following first order condition:

\[
\begin{bmatrix}
\mu_{R,1} \\
\mu_{R,2} \\
\vdots
\end{bmatrix}'
\begin{bmatrix}
\frac{1}{(A\theta_{1} + 2C(A))^2} \\
\frac{1}{(A\theta_{2} + 2C(A))^2} \\
\ddots
\end{bmatrix}
\begin{bmatrix}
\mu_{R,1} \\
\mu_{R,2} \\
\vdots
\end{bmatrix}' = 1 \quad (C.11)
\]

\[
\Rightarrow \begin{bmatrix}
\frac{\mu_{R,1}}{(A\theta_{1} + 2C(A))^2} \\
\frac{\mu_{R,2}}{(A\theta_{2} + 2C(A))^2} \\
\vdots
\end{bmatrix} = 1 \quad (C.12)
\]

\[
\Rightarrow \sum_{i=1}^{N} \frac{\mu_{R,i}^2}{(A\theta_{i} + 2C(A))^2} = 1 \quad (C.13)
\]

Moving all terms to left hand side of the equation, we can now label the new FOC as $F(A)$:

\[
F(A) = \sum_{i=1}^{N} \frac{\mu_{R,i}^2}{(A\theta_{i} + 2C(A))^2} - 1 = 0
\]

Finally, we apply the implicit function theorem and express total derivative in terms of partial derivatives,

\[
\frac{dF}{dA} = 0 \quad (C.14)
\]

\[
\Rightarrow \frac{\partial F}{\partial A} + \frac{\partial F}{\partial C(A)} \frac{\partial C(A)}{\partial A} = 0 \quad (C.15)
\]

and we rearrange terms, solving for $\frac{\partial C(A)}{\partial A}$:

\[
\frac{\partial C(A)}{\partial A} = -\frac{\partial F}{\partial A}/\frac{\partial F}{\partial C(A)} \quad (C.16)
\]

Because the numerator and denominator are both negative,

\[
\frac{\partial F}{\partial A} = -2 \sum_{i} \frac{\theta_i \mu_{R,i}^2}{(A\theta_{i} + 2C(A))^3} < 0 \quad (C.17)
\]

\[
\frac{\partial F}{\partial C(A)} = -2 \sum_{i} \frac{\mu_{R,i}^2}{(A\theta_{i} + 2C(A))^3} < 0 \quad (C.18)
\]

we have proven that $\frac{\partial C(A)}{\partial A} < 0$. 

150
C.2.4 Derivation of Equation 3.16 for the Cross Section Test

Here, we derive the standard error formula, Equation 3.16, for the cross section test. For the cross section test, we start by running a cross-sectional regression at each time \( t \):

\[ r_t \sim \Gamma_t f_t + \varepsilon_t \]  
(C.19)

where

\[ \Gamma' = \begin{bmatrix} 1 & \cdots & 1 \\ \gamma_{a1} & \cdots & \gamma_{aN} \\ \gamma_{b1} & \cdots & \gamma_{bN} \\ \vdots & \cdots & \vdots \end{bmatrix} = [\gamma_1, \cdots, \gamma_N] \]  
(C.20)

We obtain the regression coefficient at time \( t \):

\[ \hat{f}_t = (\Gamma' \Gamma)^{-1} \Gamma' r_t \]  
(C.21)

and we introduce the notation \( \delta \) to refer to the columns of \((\Gamma' \Gamma)^{-1}):(\Gamma' \Gamma)^{-1} = \frac{1}{N} \begin{bmatrix} \delta_1 & \delta_a & \delta_b & \cdots \end{bmatrix} \]  
(C.23)

We first rewrite the regression coefficient \( \hat{f}_t \) by substituting for \( r_t \) with its definition:

\[ \hat{f}_t = f_t + (\Gamma' \Gamma)^{-1} \Gamma' \varepsilon_t \]  
(C.24)
The term $\Gamma'\varepsilon_t$ can be rewritten as:

$$
\Gamma'\varepsilon_t = \begin{bmatrix}
\sum_i \varepsilon_{it} \\
\sum_i \gamma_{ai}\varepsilon_{it} \\
\sum_i \gamma_{bi}\varepsilon_{it} \\
\vdots
\end{bmatrix} = \sum_i \begin{bmatrix}
\varepsilon_{it} \\
\gamma_{ai}\varepsilon_{it} \\
\gamma_{bi}\varepsilon_{it} \\
\vdots
\end{bmatrix} = \sum_i \gamma_i\varepsilon_{it} \quad (C.25)
$$

We then substitute this into the formula for $\hat{f}_t$ and obtain the regression coefficient for a test factor which we label arbitrarily as $b$ as follows:

$$
\hat{f}_{b,t} = f_{b,t} + \frac{1}{N} \delta_b' \sum_i \gamma_i\varepsilon_{it} \quad (C.26)
$$

Now, we note that the test statistics for the cross section test is the time-series means of the regression coefficients is:

$$
\hat{\mu} = \frac{1}{T} \sum_t \hat{f}_t \quad (C.27)
$$

This means that the standard error of the test statistic for test factor $b$ is:

$$
se(\hat{\mu}_b) = \frac{1}{\sqrt{T}} \sqrt{\sigma_{bb}^2 + \frac{1}{N^2} \sum_i (\delta_b' \gamma_i)^2}
$$

### C.2.5 GMM Derivation of Equation [3.17] for the Time Series Test

Here, we derive the standard error formula, Equation [3.17] for the time series test. We start by noting that the return generating process is as follows:

$$
r_t = \Gamma_t f_t + \varepsilon_t
$$

Next, we run a Jensen’s alpha test using an arbitrary test factor $b$ by regressing its return on existing factors:

$$
f_{b,t} \sim \alpha_{b} + \beta_{b}' f_{-,b,t} + \varepsilon_t \quad (C.28)
$$

If the test factor’s returns are orthogonal to all other factor returns, then $\alpha_{b} = \mu_{b}$, the mean return of the test factor portfolio from the return generating process. If the test factor’s
returns are not orthogonal and are instead fully spanned by the other factor returns, then \( \alpha_b = 0 \). Now, we note that the test statistic for the time series test is the intercept \( \hat{\alpha}_b \) from this regression, where we substitute in population estimates for the parameters in the return generating process:

\[
\hat{\alpha}_b = \bar{f}_b - \beta_b \bar{\epsilon}_b = \bar{\mu}_b - \beta_b \hat{\mu} - b
\]  
(C.29)

We then run a GMM estimation using the OLS moments to estimate coefficients, following the standard approach in Cochrane (2009, Ch 12):

\[
g_T(b) = E_T \left( \begin{bmatrix} \epsilon_t \\ f_{a,t} \epsilon_t \\ f_{c,t} \epsilon_t \\ \vdots \end{bmatrix} \right) = 0
\]

where we define \( b \) as the vector of parameters to be estimated:

\[
\hat{b} = [\hat{\alpha}, \hat{\beta}_{ba}, \hat{\beta}_{bc} \cdots]'
\]  
(C.30)

so that \( \hat{\alpha} \) can be written as follows:

\[
\hat{\alpha} = \bar{f}_b - \beta_{ba} \cdot \bar{f}_a
\]  
(C.31)

Note that the GMM estimates here are the same as the OLS regression coefficients, and the full asymptotic joint distribution of the GMM estimates is as follows:

\[
\sqrt{T}(\hat{b} - b) \rightarrow N(0, (AD)^{-1}ASA'(AD)^{-1}^t)
\]

where the matrices \( A, D, \) and \( S \) are defined next. Because \( A_{gT}(\hat{b}) = 0, A = I \) and it drops
Matrix $D$, according to the GMM formula, is

$$D \equiv \frac{\partial g_T (b)}{\partial b} = -\Phi = -\begin{bmatrix} 1 & \bar{f}_b \\ \bar{f}_b & \bar{\Sigma}_{-b} + \bar{f}_b \cdot \bar{f}_b \end{bmatrix}$$

where we define $\bar{f}_b$ and $\bar{\Sigma}_{-b}$ as follows:

$$\bar{f}_b = (f_a, f_c, \cdots)' \quad \quad \text{(C.32)}$$

$$\bar{\Sigma}_{-b} = \frac{1}{T} \sum_t (f_{-b,t} - \bar{f}_b)(f_{-b,t} - \bar{f}_b)' \quad \quad \text{(C.33)}$$

The third matrix $S$ is defined as:

$$S \equiv \sum_{j=-\infty}^{\infty} E \begin{bmatrix} \epsilon_t \epsilon_{t-j} & \epsilon_t \epsilon_{t-j} f_{a,t-j} & \epsilon_t \epsilon_{t-j} f_{c,t-j} & \cdots \\ \epsilon_t f_{a,t} \epsilon_{t-j} & \epsilon_t f_{a,t} \epsilon_{t-j} f_{a,t-j} & \epsilon_t f_{a,t} \epsilon_{t-j} f_{c,t-j} & \cdots \\ \epsilon_t f_{c,t} \epsilon_{t-j} & \epsilon_t f_{c,t} \epsilon_{t-j} f_{a,t} & \epsilon_t f_{c,t} \epsilon_{t-j} f_{c,t-j} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Additionally, we assume that the residual terms are uncorrelated over time, are homoskedastic, and the factors other than the test factor $b$ are orthogonal to the residuals in the return generating process for factor $b$. This means that the matrix $S$ can simplified to:

$$S \equiv E \begin{bmatrix} \epsilon_t^2 & \epsilon_t^2 f_{a,t} & \epsilon_t^2 f_{c,t} & \cdots \\ \epsilon_t^2 f_{a,t} & \epsilon_t^2 f_{a,t} f_{c,t} & \cdots \\ \epsilon_t^2 f_{c,t} & \epsilon_t^2 f_{c,t} f_{c,t} & \cdots \\ \epsilon_t^2 & \epsilon_t^2 f_{c,t} & \cdots \\ \epsilon_t^2 f_{c,t} & \cdots \\ \epsilon_t^2 f_{c,t} & \cdots \\ \epsilon_t^2 f_{c,t} & \cdots \\ \epsilon_t^2 f_{c,t} & \cdots \\ \epsilon_t^2 f_{c,t} & \cdots \end{bmatrix} = \Phi E[\epsilon_t^2]$$

Note when the test factor portfolio return $f_b$ are orthogonal to the other factor returns, $E[\epsilon_t^2] = \sigma_{\epsilon b}^2$. Now, we can substitute the matrices $A$, $D$ and $S$ into the asymptotic distribution and obtain the variance of the GMM estimate $\hat{b}$:

$$Var(\hat{b}) = \frac{1}{T} D^{-1} S D^{-1} = \frac{1}{T} \Phi^{-1} \Phi \sigma_{\epsilon}^2 \Phi^{-1} = \frac{1}{T} \Phi^{-1} \sigma_{\epsilon}^2$$

To obtain the variance of the test statistic $Var(\hat{\alpha})$, we need to calculate the top left corner of the matrix $Var(\hat{b})$. First, we calculate the top left corner of $\Phi^{-1}$. To do so, we perform a

\[\text{\cite{Cochrane2012} precisely, } d \text{ is defined as the population moment in the first equality, which we estimate in sample by the second equality.}\]
matrix inversion, which takes the following form:

$$
\Phi^{-1} \equiv \begin{bmatrix} C_1^{-1} & \cdots \\ \cdots & C_2^{-1} \end{bmatrix}
$$  \hfill (C.34)

where, in our case, the upper left block is simply a scalar, though we continue to refer to it as the matrix $C_1$:

$$
C_1 = 1 - \Gamma_b' \left( \hat{\Sigma}_{-b} + \Gamma_b' \Gamma_b \right)^{-1} \Gamma_b
$$  \hfill (C.35)

The inverse of $\left( \hat{\Sigma}_{-b} + \Gamma_b' \Gamma_b \right)^{-1}$ can be rewritten as:

$$
\hat{\Sigma}_{-b}^{-1} - \hat{\Sigma}_{-b}^{-1} \Gamma_b \left( 1 + \Gamma_b' \hat{\Sigma}_{-b}^{-1} \Gamma_b \right)^{-1} \Gamma_b' \hat{\Sigma}_{-b}^{-1}
$$

which we can then substitute into the formula for $C_1$, simplify, and invert:

$$
C_1 = 1 - \frac{\Gamma_b' \hat{\Sigma}_{-b}^{-1} \Gamma_b}{1 + \Gamma_b' \hat{\Sigma}_{-b}^{-1} \Gamma_b}
$$  \hfill (C.36)

$$
= \frac{1}{1 + \Gamma_b' \hat{\Sigma}_{-b}^{-1} \Gamma_b}
$$  \hfill (C.37)

$$
\Rightarrow C_1^{-1} = 1 + \Gamma_b' \hat{\Sigma}_{-b}^{-1} \Gamma_b
$$  \hfill (C.38)

$$
= 1 + \mu' \Sigma_{-b}^{-1} \mu_{-b}
$$  \hfill (C.39)

Therefore, the standard error of the test statistics $\hat{\alpha}_b$ is:

$$
se (\hat{\alpha}_b) = \sqrt{\frac{1}{T} C_1^{-1} \sigma^2}
$$  \hfill (C.40)

$$
= \frac{1}{\sqrt{T}} \sqrt{\sigma^2 \left( 1 + \mu' \Sigma_{-b}^{-1} \mu_{-b} \right)}
$$  \hfill (C.41)

From Appendix B.6, we derive the variance of the efficient portfolio for factor $b$ is $\sigma^2_{bb} + \frac{1}{N^2} \sigma^2 \sum_i (\delta'_b \gamma_i)^2$. Therefore, by decomposition of the total variance, we get:

$$
\sigma^2_{bb} + \frac{1}{N^2} \sigma^2 \sum_i (\delta'_b \gamma_i)^2 = \beta'_b \Sigma_{-b} \beta_b + \sigma^2_{\epsilon}
$$  \hfill (C.42)

$$
\Rightarrow \sigma^2_{\epsilon} = \left( 1 - R^2 \right) \left( \sigma^2_{bb} + \frac{1}{N^2} \sigma^2 \sum_i (\delta'_b \gamma_i)^2 \right)
$$  \hfill (C.43)

where $R^2$ corresponds the regression in Equation C.28.
Plug it back in the Equation for \( se(\hat{\alpha}_b) \) and get:

\[
se(\hat{\alpha}_b) = \frac{1}{\sqrt{T}} \sqrt{(1 - R^2) \left( \sigma_{bb}^2 + \frac{1}{N^2} \sigma^2 \sum_i (\delta_i' \gamma_i)^2 \right) (1 + SR^2)}
\]

**C.2.6 Using an Efficient Factor Portfolio in Equation 3.17**

Here, we derive the form of Equation 3.17 using an efficient portfolio for a test factor \( b \). This is essentially the same as forming a portfolio for test factor \( b \) using a cross section regression, as follows. We use optimization to construct a pure test factor portfolio \( P \) that is dollar neutral, delivers unit exposure to the test factor of interest, zero exposure to all other factors, and otherwise minimizes idiosyncratic risk.

Let the weight in portfolio \( P \) be \( w' = [w_1, w_2, \ldots, w_N] \). The portfolio has the following return properties:

\[
\begin{align*}
     r_{Pt} &= f_{bt} + \varepsilon_{Pt} & (C.44) \\
     \mu_P &= \mu_b & (C.45) \\
     \sigma^2_P &= \sigma^2_{bb} + \sigma^2_{P\varepsilon} & (C.46)
\end{align*}
\]

We then minimize idiosyncratic risk:

\[
\min \sigma^2_{P\varepsilon}
\]

subject to constraints of dollar neutrality, unit exposure to \( b \) and zero exposure to all other factors:

\[
\begin{cases}
    \sum_i w_i &= 0 \\
    \sum_i w_i \gamma_{ai} &= 0 \\
    \sum_i w_i \gamma_{bi} &= 1 \\
    \sum_i w_i \gamma_{ci} &= 0 \\
    \vdots
\end{cases}
\]

The solution has the same form as the unit exposure portfolios \( Q_{unit} = \Gamma (\Gamma' \Gamma)^{-1} \).

And
so, the portfolio variance can be written as:

\[ \sigma^2_P = \sigma^2_{\delta b} + \sigma^2_{\delta z} = \sigma^2_{\delta b} + \frac{1}{N^2} \sigma^2 \sum_i (\delta'_b \gamma_i)^2 \]

which is the same as the variance of the individual cross-sectional regression coefficient.