

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Three Essays in Finance

**Permalink**

<https://escholarship.org/uc/item/8tv1d4p2>

**Author**

Wang, Ye

**Publication Date**

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA  
Los Angeles

Three Essays in Finance

A dissertation submitted in partial satisfaction  
of the requirements for the degree  
Doctor of Philosophy in Management

by

Ye Wang

2019

© Copyright by

Ye Wang

2019

# ABSTRACT OF THE DISSERTATION

Three Essays in Finance

by

Ye Wang

Doctor of Philosophy in Management

University of California, Los Angeles, 2019

Professor Antonio E. Bernardo, Chair

The first chapter of my dissertation explores the roles “quants” or technically trained workers play at hedge funds. I quantify the impact of hiring “quants” on hedge fund strategy and risk taking by exploiting H-1B lottery results and a policy shock that significantly reduced the future supply of foreign high-skilled labor in the U.S. I find that the H-1B visa program allows hedge funds to pursue strategies that are more quantitative (e.g., hedging, systematic trading, etc.) as opposed to fundamentals-based (e.g., event driven). Although there is evidence of substitution between fund strategies, I also find that hedge funds overall become more diversified, both in terms of regional focus and fund investment style.

Traditional theories on corporate financial policy mostly focus on firm-specific determinants and assume that capital structure choices are made independently of the actions of peer firms. The second chapter of my dissertation introduces a theoretical model that embodies the dynamic of how peer behavior affects a firm’s optimal choice of debt level. Specifically, peer effect is captured by a measure of how much the liquidation value of assets are adversely impacted if peer firms choose to liquidate simultaneously, the effect of which is more prominent for industries with highly specialized assets – hence low asset redeployability. The theoretical model predicts that a firm’s optional debt level is lower when the liquidation prices of its assets are expected to be more severely impacted by the concurrent liquidation of peer firms. Moreover, in those situations where peer firms’ cash flows are more closely correlated, the optional debt level will be higher. Individual firms’ attempts to

maximize their own utility lead to an industry-wide over leverage, further exacerbating the risk of general crises in an already highly correlated industry.

The third chapter of my dissertation explores the potential of using deep learning models to enhance equity trading strategies such as momentum and reversal trading. A deep feed-forward neural network model (DFN) is fed with a training data set (1965-2000) that uses the rolling Z-scored cumulative returns of various horizons as predicative variables. With the model parameters generated by the training set, the model is applied to the validation set (2000-2016) to generate a signal for each stock predicting its likelihood of outperforming the cross-sectional medium. Based on this model-predicted signal, I have constructed a long-short investment portfolio out of the validation data set. The deep learning portfolios yield an annualized return of over 20 percent and generate significantly large alphas over commonly used factor models.

The dissertation of Ye Wang is approved.

Pierre-Olivier Weill

Barney P. Hartman-Glaser

Ivo I. Welch

Antonio E. Bernardo, Committee Chair

University of California, Los Angeles

2019

*To my mother Xin, my wife Meng, and my pets Wukong and Bajie*

## TABLE OF CONTENTS

|   |               |
|---|---------------|
| <b>1 Foreign High-Skilled Labor and Hedge Fund Performance: Evidence from H-1B Lotteries</b>                            | <b>1</b>      |
| 1.1 Background and Literature   | 3             |
| 1.1.1 Immigrant Labor in the U.S.   | 3             |
| 1.1.2 Labor Diversity and U.S. Corporations   | 4             |
| 1.1.3 Regional Diversity and Investment   | 5             |
| 1.2 The H-1B Program and Data Processing  | 7             |
| 1.3 Empirical Strategies  | 11            |
| 1.3.1 Q1: Does Immigrant Hiring Affect Hedge Fund Firm Investment Behavior?   | 11            |
| 1.3.2 Q2: Does Dependence on Immigrant Labor Affect Investment Strategy? Evidence From the H-1B Visa Reform Act of 2004 | 13            |
| 1.4 Results   | 14            |
| 1.4.1 Summary Statistics  | 14            |
| 1.4.2 How Does Immigrant Hiring Affect Hedge Funds?   | 16            |
| 1.4.3 H-1B Reform in 2004   | 19            |
| 1.5 Conclusion  | 20            |
| <br><b>2 Peer Effects on Firm-Level Strategic Debt Choices: A Theoretical Model</b>                                     | <br><b>33</b> |
| 2.1 The Model   | 35            |
| 2.1.1 The Benchmark Model (no peer effects)   | 35            |
| 2.1.2 Introducing Peer Effects on Liquidation Prices  | 36            |
| 2.2 Properties of the Baseline Peer-Effect Model  | 38            |
| 2.3 Variation: Adding Different Firm Sizes  | 42            |



|          |  |           |
|----------|--|-----------|
| 2.4      | Variation: Adding Correlation in Cash Flow . . . . .                   | 43        |
| 2.5      | Conclusion . . . . .   | 46        |
| <b>3</b> | <b>Applying Deep Learning to Momentum Trading Strategies . . . . .</b> | <b>53</b> |
| 3.1      | Data Description . . . . .   | 55        |
| 3.2      | Model . . . . .  | 57        |
| 3.2.1    | Deep Feedforward Network . . . . .                                     | 57        |
| 3.2.2    | Cost Function . . . . .  | 59        |
| 3.2.3    | Gradient Descent, Stochastic Gradient Descent, and Mini-Batching .     | 60        |
| 3.2.4    | Validation . . . . .   | 61        |
| 3.3      | Results . . . . .  | 62        |
| 3.3.1    | Portfolio Returns . . . . .  | 62        |
| 3.3.2    | Factor Analysis . . . . .  | 64        |
| 3.4      | Conclusion . . . . .   | 67        |
| <b>4</b> | <b>Bibliography . . . . .</b>  | <b>69</b> |

## LIST OF FIGURES

|     |  |    |
|-----|--|----|
| 2.1 | Optimal Debt vs. Tax Benefit . . . . .                                 | 39 |
| 2.2 | Optimal Debt and Price Discount ( $P_h$ fixed) . . . . .               | 40 |
| 2.3 | Optimal Debt and Price Spread ( $P_h + P_l$ fixed) . . . . .           | 42 |
| 2.4 | Optimal Debt and Firm Size . . . . .                                   | 43 |
| 2.5 | Responses to Peer Debt with Varying $\rho$ . . . . .                   | 45 |
| 3.1 | Example of a Simple Four-Layer (6-4-4-1) DFN . . . . .                 | 57 |
| 3.2 | Five-Layer (20-50-100-20-2) DFN Model Specification. . . . .           | 58 |
| 3.3 | Illustration of Gradient Descent . . . . .                             | 61 |
| 3.4 | Portfolio Value of Different Strategies (2001-2016) . . . . .          | 63 |
| 3.5 | Relative Importance of the Twenty Input Scores . . . . .               | 63 |
| 3.6 | Average Operating Profitability in Long and Short Portfolios . . . . . | 67 |

## LIST OF TABLES

|      |   |    |
|------|---|----|
| 1.1  | Number of Hedge Fund Firms by Year . . . . .  | 22 |
| 1.2  | Number of Hedge Fund Firms by Headquarter State . . . . .   | 23 |
| 1.3  | Summary statistics of firms that filed Labor Condition Applications . . . . .                             | 24 |
| 1.4  | Summary statistics of firms that never filed Labor Condition Applications . . . . .                       | 25 |
| 1.5  | Determinants of LCA Filings . . . . .   | 26 |
| 1.6  | Determinants of Capped H-1B Visa Issuance . . . . .   | 27 |
| 1.7  | Fraction of foreign high-skilled foreign labor demand met and quantitative funds . . . . .                | 28 |
| 1.8  | Fraction of foreign high-skilled foreign labor demand met and fund strategy diversification . . . . .     | 29 |
| 1.9  | Fraction of foreign high-skilled foreign labor demand met and fund region focus diversification . . . . . | 30 |
| 1.10 | Treatment dynamics on quant vs. fundamental fund strategies . . . . .                                     | 31 |
| 1.11 | Average treatment effect on quant vs. fundamental fund strategies (2001 to 2006) . . . . .                | 32 |
| 3.1  | Annualized Returns of Different Strategies . . . . .  | 62 |
| 3.2  | Portfolio Returns vs. FF Three Factors, Momentum, and Short-Term Reversal . . . . .                       | 65 |
| 3.3  | Portfolio Returns vs. Fama-French Five-Factor Model . . . . .   | 66 |

## ACKNOWLEDGMENTS

I thank my committee chair Tony Bernardo for never giving up on me over the years, without whom I would not be able to graduate. I thank my committee members for the precious suggestions on my research over the years.

Chapter one of my dissertation is a version of the working paper “Quants, Hedge Fund Strategy, and Risk” with Shenje Hshieh. I thank him for suggestions on writing and the data work.

Over the course of my eight-year life at UCLA, my friends have been affected me greatly. Not only did they keep me accompanied but also they provided me with great supports, especially mentally.

## VITA

- 2009-2011 M.S. in Financial Engineering, University of Michigan.
- 2005-2009 B.S. in Physics, Peking University.
- 2011-2014 UCLA Anderson Fellowship.
- 2012-2018 Teaching Assistant, UCLA Anderson School of Management.

# CHAPTER 1

## Foreign High-Skilled Labor and Hedge Fund Performance: Evidence from H-1B Lotteries

How do high-skilled immigrant employees affect the investment strategies of hedge funds? This paper endeavors to expand on the growing literature that intersects between the fields of labor economics and investment by exploring the key issue of whether immigrant labor can contribute to financial firms' investment strategy beyond the contribution of local labor. Exploring this issue is inherently difficult since labor composition of firms is often unobservable. Additionally, quantifying the impact of an additional immigrant worker is not meaningful when foreign labor only constitutes a minuscule portion of a firm's total employee headcount. This is especially problematic when analyzing large firms, since they may want to hire immigrant labor with varying degrees of skill levels suited for different tasks (e.g., secretaries vs. engineers). In other words, not all immigrant labor are the same, especially in large firms.

I overcome these challenges by focusing exclusively on hedge fund firms. This setting is ideal firstly because hedge fund firms are much smaller in size than mutual fund firms. Secondly, given its small size, the skill or talent levels of employees at hedge fund firms should be much closer to being uniform or symmetric. Therefore, the marginal effect of immigrant labor can be significant as we expect the average intended foreign hire to be highly skilled, trained, experienced, and specialized. This claim is verifiable by studying the job titles and highest academic degrees as reported in the U.S. temporary work visa applications filed by hedge fund firms.

The key outcome of interest in this project is the investment behavior of hedge fund firms.

Specifically, I explore the changes to a hedge fund firm's investment style, regional coverage, and risk behavior resulting from the hiring of high-skilled foreign labor. A central link under exploration is whether labor diversification would lead to more diversified investment strategies, both in terms of regional focus and fund investment styles. My null hypothesis is twofold: first, the hiring of high-skilled foreign labor contributes to the adoption of more quantitative strategies (e.g., hedging, systematic trading, etc.) as opposed to fundamental-based strategies (e.g., event driven); Second, the diverse backgrounds of foreign-born labor brings in broader exposure to different regional markets and therefore contributes to the diversification of the hedge fund firms' investment portfolios. Since the U.S. market is more predisposed to the employment of quantitative investment strategies, the combined effects of the aforementioned two mechanisms could generate an empirical observation that the hiring of foreign-born employees in hedge fund firms leads simultaneously to a more intensive focus on the U.S. market as well as a higher level of overall regional diversification of the investment portfolio.

The main empirical strategy in establishing a clear link between foreign labor hiring and hedge fund investment behavior is to exploit random firm-level variation in the supply of immigrant work visas in the U.S. Specifically, H-1B visas are often distributed by lottery due to excess demand for them beyond the fixed quota. For example, the number of H-1B visas available in 2008 was 85,000. The number of H-1B applications, however, exceeded 150,000 on the first day the Department of Homeland Security started accepting applications for the year. As a result, a lottery was held to randomly distribute H-1B visas in 2008. For every year, some hedge fund firms will arbitrarily not be able to hire (for the long-term) some of their intended immigrant labor. I interpret this failure in hiring as an excess demand shock for foreign skilled labor. Therefore, using this shock, I am able to cleanly test for the causal links between high-skilled foreign labor and hedge fund investment behavior.

## 1.1 Background and Literature

### 1.1.1 Immigrant Labor in the U.S.

Although the U.S. historically has stood as one of the most accommodating countries with respect to immigration, the topic has always been controversial due to the inherent avenues of abuse (e.g., displacement of native workers, suppression of wages, etc.) and risk to national security (e.g., corporate espionage, terrorism, etc.). Despite arguments for a more restrictive immigration policy, we have yet to fully understand the role of immigrants in facilitating economic growth and have only recently begun understanding the value of immigrant labor from the perspectives of corporations.

U.S. corporations hire most of their immigrant labor through the H-1B visa program. Firms sponsor their intended foreign-born hires for H-1B visas by filing two key documents: the Labor Condition Application (LCA), which is submitted to the U.S. Department of Labor, and the I-129 petition, which is submitted to the U.S. Department of Homeland Security. In doing so, however, they must provide supporting evidence that hiring them will not displace native workers. To demonstrate the necessity of hiring a foreign national, the firm needs to show that the intended foreign hire is uniquely qualified for the job position in question by his/her educational qualification, experience, and specialty and that an equally qualified local alternative cannot be found within a fixed period of time.

The empirical evidence on the crowding out effect are mixed. Kerr and Lincoln 2010 and Peri, Shih, and Sparber 2015, for example, find no evidence of local worker displacement at the state level and within computer-related firms, respectively. Peri, Shih, and Sparber 2015 suggest that the hiring of native and immigrant labor are in fact complements. Aobdia, Srivastava, and Wang 2018, who study the determinants of LCA submissions by large accounting firms, similarly find that the immigrant hires serve mostly as “gap fillers” and specialists. Likewise, Mithas and Lucas 2010 show that on average foreign IT professionals earn more than IT professionals with U.S. citizenship. This finding suggests that U.S. and foreign IT professionals may, again, be complements (or imperfect substitutes at the very least). On the other hand, Doran, Gelber, and Isen 2014 find that winning additional H-1B



visas via lottery in the years 2006 and 2007 led to fewer employment of local workers, lower employee earnings, and higher firm profits.

Whether immigrant labor actually make contributions beyond that of local labor is also ambiguous and understudied. Doran, Gelber, and Isen 2014, for example, demonstrate that additional H-1B visas approved at the firm level has an insignificant effect on patenting and R&D activity (measured by the use of R&D tax credits). However, this finding is at odds with the preponderance of evidence of an increasing proportion of innovating activities done by immigrants. Hunt and Gauthier-Loiselle 2010 cite several supporting facts: 25 percent of founders of public U.S. companies between 1990 and 2005 that received venture capital financing are immigrants, 25 percent of founders of high-tech companies in 2006 with more than \$1 million in sales are immigrants, and non-U.S. citizens account for 24 percent of all international patent applications from the U.S. I intend to extend this literature through the analysis of immigrant labor within hedge fund firms in an attempt to understand their roles and contributions to portfolio composition and investment strategies.

### **1.1.2 Labor Diversity and U.S. Corporations**

This project contributes to the branch of literature that explores how characteristics of labor within a firm can influence its strategy and financing decisions. While many recent research highlight the cost of labor within a firm (e.g., Simintzi, Vig, and Volpin 2015 show increase in employment protection can crowd out financial leverage), others such as Tate and Yang 2015b, Tate and Yang 2015a, Tate and Yang 2015c, Mueller, Ouimet, and Simintzi 2017, and Ahern, Daminelli, and Fracassi 2015 provide empirical evidence that labor diversity, in terms of skill, gender or ethnicity, for example, can potentially enhance firm value.

Tate and Yang 2015b and Tate and Yang 2015c explore the worker-firm matched data from the U.S. Census Bureau in order to understand the value of corporate diversification and its impact on human capital. Tate and Yang 2015b document that workers in diversified (operating in at least two distinct industries) firms develop a more diverse set of skills compared to single industry firms. Workers with a broader range of skills afford the firm more

adaptability: diversified firms can redeploy labor to sectors with higher marginal returns. As result, firms enjoy higher productivity and lower costs of hiring and firing. Tate and Yang 2015c find similar benefits in the context of mergers and acquisitions: industry pairs with higher human capital “transferability” leads to larger productivity gains during cross-industry acquisitions.

Diversity in labor skill levels, as represented in differential pay scales, has also been shown to benefit firms. Using a set of firms in the United Kingdom, Mueller, Ouimet, and Simintzi 2017 show that larger relative wage differentials lead to higher valuations, stronger operating performance, larger equity returns, and greater sensitivity to earnings surprises. As Mueller, Ouimet, and Simintzi 2017 argue, this increased productivity is the result of higher incentive provisions.

Lastly, Tate and Yang 2015a and Ahern, Daminelli, and Fracassi 2015 find evidence that labor diversity through gender and culture could limit bias in employee compensation and increase trust between firms, respectively. Tate and Yang 2015a find that the wage gap between men and women are much smaller under female leadership at plants. Ahern, Daminelli, and Fracassi 2015 underscores the importance of ethnic culture in cross-border mergers, as they find that cross-border mergers yield lower combined announcement returns when target and acquirer have greater cultural distance. This project expands on the benefits of diversity in labor and corporate culture through the exploration of the role of immigrant labor in firms.

### **1.1.3 Regional Diversity and Investment**

Past scholarship has explored the connection between investors’ place of origin or residence and the regional distribution of their investment portfolios. In particular, in both household and professionally managed portfolios, the home bias puzzle or the observation that investors often have a bias for local investments, is well documented. These findings, however, does not necessarily imply the presence of behavioral bias. Van Nieuwerburgh and Veldkamp 2010, for example, provide a model to show that investors can deviate from a diversified portfolio

if allowed to collect information first. In other words, permitting information acquisition can rationalize the emergence of both diversified and concentrated investing strategies.

Many of the empirical findings seem to support the superior information theory (i.e., “rational bias” theory), as documented in Coval and Moskowitz 1999, Coval and Moskowitz 2001, Ivkovic and Weisbenner 2005, and Bae, Stulz, and Tan 2008. Coval and Moskowitz 1999 find that U.S. investment managers have a strong preference for firms whose headquarters are located nearby. Coval and Moskowitz 2001 extend this finding by showing that fund managers additionally earn substantial abnormal returns on these nearby investments, which supports the information advantage theory. Ivkovic and Weisbenner 2005 draw similar conclusions when analyzing household portfolios, which also exhibit heavy bias towards local investments. Likewise, these local holdings generate significant higher returns than non-local holdings, which again suggests superior knowledge on local firms. Lastly, Bae, Stulz, and Tan 2008 show that an analyst makes more precise earnings forecasts for firms located in the same country where the analyst is residing.

Besides a “rational bias” explanation, cultural proximity and familiarity could be the behavioral factors that lead to home bias. Grinblatt and Keloharju 2001 show that investors are more inclined to trade Finnish firms that are located closer in proximity, communicate with them in their native language, and have CEOs of the same cultural background as they do. Huberman 2001 likewise show that household’s portfolios are concentrated towards firms that are geographically close or familiar (e.g., employees owning employer’s stock in their retirement portfolio). However, as found in Grinblatt and Keloharju 2000, home bias is inversely related to investor sophistication. These set of findings are difficult to reconcile with just the superior information theory.

Whether the cost of information acquisition or cultural familiarity is the explanatory factor for home bias, labor diversification stands as a plausible channel to mitigate this bias. The diverse backgrounds of foreign-born labor could change the corporate culture, style, or practices of hedge fund firms in a positive way by bringing new perspectives and therefore new or different investment proposals into consideration. Their language skills and cultural versatility could serve to reduce the cost of information acquisition. Therefore, one could

expect that the hiring of foreign labor in a hedge fund firm should lead to a more regionally diversified investment portfolio.

On the other hand, since foreign employees recruited through H1-B visas in the investment industry tend to be more highly trained in quantitative methods, we can also expect more quantitative-based investment strategies at those firms that depend on foreign labor. Given that quantitative strategies are more easily applicable to the U.S. market, another plausible outcome is that the hiring of high-skilled foreign labor leads to a more U.S.-focused investment portfolio, which would be the direct opposite of “home bias.”

## **1.2 The H-1B Program and Data Processing**

More than million immigrants enter the United States every year. Most have significantly lower skills than native workers, but a narrow category of highly-skilled immigrants have played an important role in the technological advances achieved by the corporate sector and in the research produced by academic institutions. The regulation concerning their employment by U.S. corporations is linked to the H-1B temporary visa program that has remained practically unchanged since 1990, despite dramatic changes in the U.S. business and economic environment. H1-B is a short-term work visa issued to employers to hire professional in specialty occupations that require at least a bachelor’s degree and the theoretical and practical application of highly specialized knowledge and skills. It is valid for an initial period of three years and can be renewed once, for a total of six years.

Since the the advent of the H-1B program in 1990, it has been subject to an annual quota on new visa issuance. The initial quota capped the annual new assurances of H-1B visas at 65,000, which was not surpassed until 1997. The period from 1999 to 2003 saw a temporary relaxation of the quota: the cap was increased to 115,000 for 1999 and 2000 and further to 195,000 for 2001 through 2003. However, in 2004, the cap was drastically reduced back to 65,000. In addition, the H-1B Visa Reform Act of 2004 allowed for an additional 20,000 visas each year reserved for foreigners who hold a master’s degree or above from an American institution of higher education. Moreover, since 2000 U.S. higher education and non-profit

institutions wishing to hire foreign skilled workers were exempt from H-1B limits, creating an open pathway for foreign masters and Ph.D. students to find academic and research-based employment. Still, the overall quota pales in comparison with the number of foreigners receiving advanced degrees in the U.S., which increased to 380,000 by 2016. The quota is filled on a first-come, first served basis beginning on April 1 of the prior fiscal year. Since 2004, the H-1B quota has been reached within a few days to a few months of the opening date.

Several steps are required for a foreign national to acquire an H-1B visa. The very first step is to find a U.S.-based employee that is willing to sponsor his/her visa application. For information on hedge fund firms, I have obtained the Hedge Fund Research (HFR) Assets/Performance Data, which includes around 1.8 million reports on hedge funds from 1998 to 2017. By restricting the sample to U.S.-based hedge funds for which regular monthly reports are available, the sample size is reduced to 780,000. Each report contains information on the hedge fund's address, investment strategies, leverage, assets, regional investment distribution, and performance.

A firm willing to hire a foreign-born specialty worker must first file an LCA with the Department of Labor. The LCA document outlines the nature of the job and attest that the firm will comply with H-1B regulations. The form includes information on the prevailing wage of the occupation, the wage to be paid to the prospective worker, and the address of the work site. Since the visa is temporary in nature, the LCA must also provide the beginning and end dates of the position. It's worth mentioning that a firm might file for a single or multiple potential hires in the same LCA form and that personal information of the employee (e.g. name, country of origin, age, sex, previous visa status) is not revealed in the LCAs. I have obtained the LCAs for H-1B visas from the U.S. Department of Labor, which include roughly 7 million records from 2001-2017. From these, those filings that are not certified or do not have valid employment start dates or decision dates are deleted, leaving 6.37 million entries. Based on firm name and address information, I matched the HFR data with the LCA data, generating a sample of 10,844 matched LCA records filed by hedge fund firms.

A prospective foreign employee's application for an H-1B visa is submitted in a I-129 form

to the U.S. Department of Homeland Security, which must be accompanied by an approved LCA. Therefore, while the visa belongs to an employee (not the firm), a person can only obtain a new H-1B visa when his/her employer has obtained LCA approval (there is no cap on LCA approvals).

I-129 applications are accepted beginning on April 1, six months prior to the October 1 start of the federal fiscal year, and throughout the fiscal year. Since H-1B processing works on a first-come, first-served basis, given the acute scarcity of the annual quota, applicants try to submit their cap-subjective I-129 applications on or shortly after April 1. For example, the U.S. Citizenship and Immigration Services (USCIS) in the Department of Homeland Security already received 119,193 cap-subject H-1B applications by April 3, 2007 and approximately 163,000 petitions in the first week of April 2008. USCIS randomly selected visas for processing until the 65,000 available quotas are all allocated. The unselected visa applications are not processed and returned. This lottery process induces a rationing of the supply of foreign skilled workers that is both significant in scale and randomly distributed among applicants. Firms submitting LCAs clearly have vacancies in specific occupations that they hope to fill with foreign skilled labor, but the lottery generates a randomly distributed negative shock to their supply.

I have obtained a total of roughly 5.5 million processed H-1B I-129 applications spanning from 1996 to 2016 from the U.S. Department of Labor. Each I-129 record contains the receipt date of the petition, employer's name and address, compensation, job code, job definition, and some personal information of the applicant, such as country of birth, country of citizenship, current visa class, field of study, etc. It also contains a series of questions that help identify the nature of the filing (e.g. new petition, revised previous petition). I first deleted all the filings that were meant to correct errors in a previous submission. Then, matching the remaining records with the HFR data based on firm name and address information, I got a sample of 7,791 processed I-129 applications from prospective foreign employees at hedge fund firms.

Notice that this sample includes three types of applications: 1) processed cap-dependent applications (these are the applications that have passed through the lotteries); 2) processed

cap-exempt applications; 3) processed applications for extensions of existing H-1B status. For an example of the second category, the processed I-129 applications from foreigners who hold advanced degrees and are to be employed in higher education or non-profit institutions are included in this sample, while these applications are not subject to the annual quota placed on new H-1B visas. The empirical specification of this study, which will be discussed in the next section, requires the processed cap-dependent I-129 applications (*I129\_CAP*) be separated from the other two categories of cap-exempt and extension applications (*I129\_Other*).

Applications for extensions are easily identifiable from a question in the form. The complicating matter is that the form does not always clearly indicate whether a new application is cap-subjective or cap-exempt. Specifically, older versions of the I-129 form, the last one being the revision of March 11, 2009, did not include a question on whether the application is cap-subjective or cap-exempt. A new version of the form introduced on November 23, 2010 and, for that matter, all subsequent versions include a required question that asks the applicant to specify the type of H-1B petition being filed by selecting from four options, three of which are cap-subjective (“cap H-1B bachelors’ degree,” “cap H-1B masters’ degree or higher,” and “cap H-1B1 Chile/Singapore”) and the other one is “cap exempt.” Therefore, for all the I-129 applications filed since 2011, I can rely on answers to this question to determine their cap status. For older applications, I can only rely on the information on the individual characteristics of the applicant to determine whether he/she would count toward the H-1B quota.

For older applications that didn’t include the self-identification question on filing type, I have employed the following criteria to approximately arrive at a sub-sample of cap-subjective I-129s, with the cap-exempt and extension applications in its complement. First, given that the processing of cap-subjective H-1B visas is first-come first-served and the scarcity of the quota, I assume that prospective cap-dependent employees would file I-129 petitions within a short span of time after April 1 to be eligible for working in the U.S. from the next fiscal year (from October 1). Based on this assumption I have filtered out all the I-129 applications other than those filed in April, May, and June. Second, since cap-dependent applicants are applying for H-1B visas for the next fiscal year, the start date

of their jobs cannot be before October 1; while a cap-exempt applicant may very well have the leisure of belatedly submitting an I-129 in May to start a job in the August of the same fiscal year. Based on this differentiation, I have further filtered the remaining sample by requiring that the stated job start date should be later than October 1. Third, based on their answers to certain questions in the I-129 form, I have filtered out those applicants who are to be employed at institutions of higher education, government research organizations, or non-profit organizations. Fourth, the checks of certain boxes in the I-129 form, such as checking "new employment" for the question on the basis of classification (part 2, question 2), are necessary (but not sufficient) indicators for cap-dependent applications. Those forms that contain responses other than such specifications are filtered out. Following such measures, the remaining records in the sample are assumed to represent processed cap-dependent I-129 applications (*I129\_CAP*), while its complement set should include both cap-exempt applications and applications for extensions (*I129\_Other*).

## 1.3 Empirical Strategies

### 1.3.1 Q1: Does Immigrant Hiring Affect Hedge Fund Firm Investment Behavior?

Our primary specification relies on the assumption that some hedge fund firms are randomly unable to hire foreign-born labor almost every year due to aggregate demand for H-1B visas exceeding the capped supply. Hedge fund firms essentially have to compete with large corporations that have the resources to hire large amounts of foreign labor every year. This compounds the difficulty of acquiring H-1B visas for hedge fund firms. Therefore, being able to hire or keep a foreign-born candidate employee is much more meaningful for hedge fund firms. However, we cannot use the total number of received H-1B visas outright as an exogenous variable, as it depends on the number of submitted H-1B visa applications, which is endogenous. Instead, I take advantage of the lottery of cap-subjective petitions as a random exogenous shock to a firm's foreign labor supply.



A firm that plans to sponsor an H-1B visa for an intended foreign hire needs to file the Labor Condition Application (LCA) to the U.S. Department of Labor. Hence the total number of LCAs filed by a firm indicates its expected foreign labor demand. I define  $LCA_{it}$  as the total number of LCAs for H-1B visas filed by firm  $i$  in year  $t$ ; of these,  $LCA\_CAP_{it}$  is the number of LCAs that correspond to cap-subjective H-1B visa applications. On the supply side, the total number of approved I-129 petitions indicates the actual number of H-1B visas acquired by a firm. Similarly, I define  $I129_{it}$  as the total number of approved I-129 petitions obtained by firm  $i$  in year  $t$ , and of these,  $I129\_CAP_{it}$  is the number of approved cap-subjective I-129 petitions.

Because of the random lottery, an exogenous measure of a hedge fund firm's employment of foreign labor can be defined as the fraction of satisfied foreign labor demand out of the total lottery-capped LCAs, which is defined as:

$$MD_{it} \equiv \frac{I129\_CAP_{it}}{LCA\_CAP_{it}} \quad (1.1)$$

The previous section has detailed how I have separated the sample of processed I-129 applications into a sub-sample of cap-subjective applications ( $I129\_CAP$ ) and another sub-sample of cap-exempt and extension applications ( $I129\_Other$ ). To estimate  $LCA\_CAP$ , I made the following assumption. Because there is no quota placed on new cap-exempt I-129 petitions or applications for extensions, it is reasonable to assume that in such categories, there should be a one-on-one correspondence between the LCA sample and the processed I-129 sample. It is in the category of the cap-dependent petitions that we might see an LCA in the sample but could not find a corresponding processed I-129, because this I-129 petition did not pass through the lottery and hence did not end up in the processed I-129 sample. Based on this assumption,  $LCA\_CAP_{it}$  should equal to  $LCA_{it}$  minus  $I129\_Other_{it}$ .

The baseline regression model estimates the effect of marginal changes in  $MD_{it}$  on several outcome measures of the hedge fund firms' investment behavior and performance:

$$y_{it} \equiv \alpha_t + \lambda_1 D\_CAP_{it-1} + \lambda_2 MD_{it-1} + \gamma' C_{it} + \epsilon_{it} \quad (1.2)$$

$y_{it}$  is a measure of hedge fund firm investment style or performance,  $\alpha_t$  is the year fixed effect,  $C_{it}$  is a vector of control variables for firm characteristics (e.g., age, size, location, etc.), and  $\epsilon_{it}$  is the error term.  $D\_CAP_{it}$  is a dummy variable that equals 1 if firm  $i$  has submitted any H1-B LCAs that are subject to the lottery cap in year  $t$  (i.e.  $LCA\_CAP_{it} > 0$ ). A statistically significant  $\lambda_2$  would support the proposed channel of impact.

The main outcomes of interest are the hedge fund firms' investment style and diversification:

1. Investment strategy style: I define two major investment styles: "quantitative" vis-a-vis "fundamental." Based on hedge funds' self-reported investment strategies and sub-strategies, I count the strategies of "equity market neutral," "quantitative directional," "active trading," "commodity-multi," "currency-systematic," and "systematic diversified" as quantitative investment style as they require highly sophisticated quantitative analysis. I count the strategies of "event-driven," "fundamental growth," and "fundamental value" as fundamental investment style as they rely on fundamental analysis.
2. Diversification: a hedge fund firm's diversification in terms of investment region (e.g. U.S., Asia, Africa, etc.) and strategy (e.g. macro, equity hedge, event-driven, etc.) is measured by the Herfindahl-Hirschman Index (HHI) of its constituent funds in that respect.

### **1.3.2 Q2: Does Dependence on Immigrant Labor Affect Investment Strategy? Evidence From the H-1B Visa Reform Act of 2004**

An alternative method to quantifying the impact of high-skilled immigrant labor on hedge fund investment is to compare hedge fund firms that relied on immigrant labor to those that did not around the time of a U.S. policy change that decreased the H-1B work visa supply cap. The H-1B Visa Reform Act of 2004 reduced the number of H-1B visa supply from 195,000 to 85,000 in 2005. I would evaluate investment style differences between these two types of hedge fund firms using the following difference-in-differences specification:

$$y_{it} = \alpha_i + \alpha_t + \tau(Post_t \times Foreign_i) + \gamma' C_{it} + \epsilon_{it} \quad (1.3)$$

where  $\alpha_i$  and  $\alpha_t$  are firm and calendar-year fixed effects,  $y_{it}$  is a measure of investment style (e.g., the count or percentage of quantitative-style funds in a firm),  $Post_t$  is a dummy variable that equals to 1 if  $t$  is greater than the year 2004 and 0 otherwise,  $Foreign_i$  is a dummy variable that equals to 1 if the hedge fund firm has previously hired immigrant labor (i.e., hired foreign worker that received H-1B visas) in the year 2005 and 0 otherwise, and  $C_{it}$  is a vector of control variables for hedge fund firm characteristics.  $\tau$  is the difference-in-differences estimator that measures the sensitivity of hedge fund firm investment style with respect to being dependent on the H-1B visa system.

For example, with respect to the percentage of quantitative-style funds within a hedge fund firm, a  $\tau < 0$  suggests that a negative supply shock of immigrant labor has an adverse effect on the adoption of quantitative investment strategies. This finding would imply that immigrant labor enhances the technical sophistication of hedge fund firms beyond local labor.

## 1.4 Results

### 1.4.1 Summary Statistics

Table 1.1 reports the number of unique firms that show up in the whole HFR sample in any given year and the number of firms that filed any LCA petition in that year. While the total number of hedge fund firms has varied between 1201 and 1944, the percentage of firms that demanded foreign labor in any given year varied between 6 to 12 percent. As expected, the total number of hedge fund firms as well as the number and percentage of firms that filed LCAs all reached the highest point in 2007 or 2008; thereafter, the percentage of hedge fund firms that filed LCAs fluctuated between 9 to 11 percent. Table 1.2 shows similar variables by state. Hedge fund firms that demand foreign labor are concentrated in New York, California, Connecticut, Illinois, and Massachusetts.

[Place table 1.1 and 1.2 about here]

Table 1.3 and 1.4 shows the summary characteristics of firms that have or have not ever filed LCA petition respectively. In my sample, there are 831 hedge fund firms that have filed a LCA at least once during the sample period and 3,059 firms that have never filed any LCAs. Those firms that have had demand for foreign labor are on average 4 years older and more likely to be located in New York; their managed asset pool is on average eight times bigger.

[Place table 1.3 and 1.4 about here]

Table 1.5 further delineates the correlations between a firm's demand for foreign labor in any given year and a list of firm characteristics, including firm age, asset under management, leverage, the number of constituent funds that specialize in quantitative or fundamental investment strategies. Column 1 shows the result of the Probit model. It seems that younger firms are more likely to file LCAs. One possible explanation is that older firms have enough immigrant labors that they acquired in earlier years. The probability of filing LCAs is positively correlated with a firm's asset under management (AUM), leverage, and the number of funds using quantitative strategies. More heavily leveraged funds tend to deploy more complicated algorithmic strategies, generating higher demand for skilled labor. The same goes for quantitative hedge funds. Since Greene 2004 shows that the maximum likelihood estimator in Probit model with too many fixed effects has a persistence bias, I have also reported the results from a direct ordinary least square (OLS) estimation with the same dependent dummy variable in column 2 and with the logarithm of the number of LCA-backed workers as the dependent variable in column 3. The results are consistent.

[Place table 1.5 about here]

### 1.4.2 How Does Immigrant Hiring Affect Hedge Funds?

I take advantage of the random lottery of cap-dependent H-1B visa petitions to examine the causal effect of hiring more immigrant labor on a hedge fund firm's investment behavior. First, Table 1.6 demonstrates that the constructed variable  $MD_{it}$ , which represent the percentage of a firm's satisfied cap-subjective foreign labor demand, is indeed an exogenous shock to its foreign labor supply. The table shows the correlations between a set of firm characteristics and the number of cap-subjective H-1B visas granted to a firm's prospective employees in any given year, i.e. the satisfied portion of the firm's demand for foreign labor. The result shows that the issuance of cap-dependent visa is indeed very random. The number of capped visas received by a firm is only strongly correlated with the number of filed LCAs – buying more lottery tickets, more chances to win. Other visible traits of the firm demonstrate no correlation with the number of received cap-dependent H-1B visas. Column 2 shows the results when the dependent variable is changed to  $MD_{it}$ . Again, it is not predictable by any firm characteristics. It supports my premise that  $MD_{it}$  represents an exogenous, random supply shock to a firm's demand for foreign labor.

[Place table 1.6 about here]

Using  $MD_{it}$  as the main independent variable, the following three tables demonstrate the impact of foreign labor hiring on hedge fund firms' investment style (Table 1.7), regional diversification (Table 1.8), and strategy diversification (Table 1.9). They are all based on post-2008 data, when the lotteries of H1-B visas became regularized and highly restrictive.

Table 1.7 reports the one-year forward effect of  $MD_{it}$  on the number of quantitative funds and fundamental funds within a hedge fund firm. Quantitative funds are defined based on the hedge funds strategies. These are funds that employ sophisticated quantitative techniques of analyzing price data to ascertain information about future price movement and relationships between securities to make selections for purchase and sale. Data is collapsed at the firm-year level. On average, firms whose demand for capped H-1B labors is fully satisfied will have ten percent more funds with quantitative strategies than those funds whose demand is not met at all. Also, firms with fully satisfied cap-dependent foreign labor demand are 6 percent

more likely to contain at least one fund that utilizes quantitative strategies than those whose demand for cap-subjective foreign labor is not met at all. These findings support the claim that hiring high-skilled foreign labor will cause the firm to implement more quantitative strategies.

Columns 4, 5, and 6 in Table 1.7 show the effect of hiring foreign immigrant labor on the number of fundamental funds within a firm. Fundamental funds are defined based on the strategies they deploy, which mostly focus on identifying attractive opportunities through the analysis of fundamental value and growth potential and responding to significant events. Column 5 shows that on average firms with their capped H-1B labor demand fully satisfied will have seven percent fewer funds using fundamental strategies, comparing to firms whose capped H-1B demand are not met at all. Column 6 shows that firms whose capped H-1B labor demand are fully met will allocate eight percent less of their asset to be invested by fundamental strategies, comparing to firms whose capped H-1B demand are not met at all. Here my assumption is that funds that utilize fundamental strategies are less technically sophisticated and therefore have less use for skilled labor. Based on this assumption, these results further verify that extra supply of foreign skilled labor within a firm will result in higher number of funds that require higher skills.

[Place table 1.7 about here]

Table 1.8 reports the impact of foreign labor supply on the regional diversification of a hedge fund firm's investments. Here I use Herfindahl-Hirschman Index (HHI) to measure regional diversification. Specifically, the HHI of a firm's regional diversification is calculated in the following way. Each of the firm's constituent funds has a reported regional focus of investment, which I have grouped into five big regions of "Asia," "Europe," "America," "Africa," and "Others." The HHI is calculated as the sum of squared asset weight under each region. For example, suppose that within in a firm, those funds that report "America" as their regional investment focus collectively manage 80 percent of the firm's total asset, those that report "Europe" manage 10 percent, and those that report "Asia" take up the remaining 10 percent. The HHI of this firm's regional diversification would be  $0.66 ((0.8)^2 +$

$(0.1)^2 + (0.1)^2$ ). A higher HHI would indicate lower diversification level.

The HHI of regional diversification is the dependent variable in Column 1. Extra supply of foreign labor will lead to a significant decrease in the HHI indicator, namely, a higher level of regional diversification in the firm's investments. Specifically, a firm whose demand for cap-dependent foreign labor is not met at all will have an HHI higher by 0.032 than a firm whose demand for cap-dependent foreign labor is fully satisfied. Column 2 provides the result of a robustness check by using the logarithm of the number of unique investment regions within a firm as the dependent variable and shows similar causal relations. As demonstrated in column 3, the increase in regional diversification in response to extra foreign labor supply is not because of an increase in the number of funds in this firm, but because of existing funds' incorporation of more investment regions.

[Place table 1.8 about here]

Table 1.9 reports the impact of foreign labor supply on a hedge fund firm's strategy diversification. A firm's strategy diversification is also measured by an HHI, which is calculated as the sum of squared asset weight of each unique main strategy as reported by its constituent funds. There are seven possible main strategies. For example, suppose that within a firm, those funds that report "equity hedge" as their main investment strategy collectively manage 50 percent of the firm's total asset, those that report "event driven" manage 30 percent, and those that report "relative value" manage the remaining 20 percent. The HHI of this firm's main strategy diversification would be 0.38 ( $(0.5)^2 + (0.3)^2 + (0.2)^2$ ). Again, a higher HHI would indicate lower diversification level, and vice versa.

The HHI of main strategy diversification is the dependent variable in column 1. Extra supply of foreign labor will lead to a significant decrease in the HHI indicator, which in this case means a higher level of main strategy diversification in the firm's funds. Specifically, a firm whose demand for cap-dependent foreign labor is not met at all will have an HHI higher by 0.045 than a firm whose demand for cap-dependent foreign labor is fully satisfied. As robustness checks, column 2 and 3 report the results when the dependent variable is the logarithm of the number of unique main strategies or the logarithm of the number of

unique sub-strategies as reported by a firm's funds. The results are consistent. A firm whose demand for cap-dependent foreign labor is fully satisfied will have 5 percent more unique main strategies and 8 percent more unique sub-strategies than a firm whose demand for cap-dependent foreign labor is not met at all. Here again, column 4 shows that the increase in strategy diversification in response to extra foreign labor supply is not because of an increase in the number of funds in the firm, but because of existing funds' adoption of more diverse main and sub-strategies.

[Place table 1.9 about here]

Up to this point, I have taken advantage of the external shock to firms' foreign labor supply generated by the random H-1B visa lotteries and demonstrated that hiring foreign skilled labor will cause a hedge fund firm to adopt more quantitative investment styles and diversify its regional investment focus and strategies. The following section shows the result from another empirical strategy, where I exploit a sudden U.S. policy change in 2004 to quantify the impact of high-skilled immigrant labor on hedge fund investment behavior.

### **1.4.3 H-1B Reform in 2004**

The H-1B Visa Reform Act of 2004 suddenly reduced the annual quota of new H-1B visa issuance from 195,000 to 85,000 in 2005. This policy change provides an opportunity to examine how the sudden restriction on H-1B visas might have had differential impacts on hedge fund firms that had demanded foreign labor previously and those that had never demanded foreign labor. The data sample includes all hedge fund firms from 2001 to 2006 and is collapsed on the firm-year level. Firms in the treatment group are those that had filed an LCA at least once before 2004.

Table 1.10 reports the difference in difference regression based on the H-1B visa reform in 2004. The outcome variable of interest is the investment style of the hedge fund firms. I first look at the number fundamental or quantitative funds within a firm. Column 4 shows that compared to the control group, treated firms saw 2 percent more increase in their number of fundamental funds from 2003 to 2004. The difference increases to 2.4 percent with two-year



accumulated effect (from 2003 to 2005) and 3.2 percent with three-year accumulated effect (from 2003 to 2006). If I replace the number of fundamental funds with its percentage or asset-weighted percentage as the dependent variable, the results are similar (column 5 and 6). Table 1.11 confirms these results. The interacted term of treatment and post-2004 is positively related to the increase in the number and percentage of fundamental funds within in a firm. This shows that the H-1B reform did have an impact on hedge fund strategy deployment. Because of the post-2004 restriction on the supply of high-skilled foreign labor, hedge fund firms relatively shifted more to fundamental strategies. Firms that had hired foreign labor before 2004 were more impacted by the sudden policy restriction and saw a greater increase in the use of fundamental strategies in the years immediately after the policy shift.

In both Table 1.10 and 1.11, the differential impact on the number of quantitative funds between controlled and treated firms is not significant. In fact, for both the treated and controlled groups, there is no significant difference in the number of quantitative funds before and after the 2004 shock. While firms that might have wanted to expand their quantitative sections through hiring high-skilled foreign labor were no longer able to do so as quickly as they had wished as a result of the restriction of the issuance of new H-1B visas, maintaining existing foreign labor would not be affected by the policy exchange. Moreover, the controlled firms are likely to be less ambitious about the adoption of quantitative investment strategies to begin with. The lack of significant difference in difference effect for the number of quantitative funds is therefore understandable.

[Place table 1.10 about here]

[Place table 1.11 about here]

## 1.5 Conclusion

By exploiting the exogenous impact of random H-1B lotteries on foreign labor supply, I have examined the role played by technically trained foreign-born employees at hedge fund

firms. The H-1B visa program allows hedge fund firms to hire from a larger pool of adequately trained “quants” to pursue more quantitative investment strategies as opposed to fundamental-based strategies. Labor diversification, as resulted from the hiring of foreign-born labor, also leads to more diversified investment regions and strategies. In other words, the two traits of H-1B visa holders, namely highly specialized skills and diverse background, both contributed to changes in hedge fund investment behavior towards a more technical and more diverse direction.

These results are also confirmed by a difference-in-difference study that quantifies how a 2004 policy shock that significantly reduced the supply of H-1B visas affected hedge fund firms that had developed a dependence on the H-1B program for foreign labor supply more than it impacted firms that had never hired any foreigners. With the tightened restriction on the issuance of new H-1B visas, hedge fund firms responded by shifting toward more fundamental strategies – this impact was more visible for those firms that had already developed a demand for foreign labor through the H-1B program.

## A Summary Statistics

| Year | No. of Firms | No. of Firms<br>Filed LCAs | Aggregate Fraction<br>of Foreign Labor Demand Met |
|------|--------------|----------------------------|---|
| 2001 | 1201         | 76                         | 1.00  |
| 2002 | 1349         | 94                         | 0.00  |
| 2003 | 1481         | 122                        | 0.33  |
| 2004 | 1644         | 166                        | 0.00  |
| 2005 | 1803         | 182                        | 0.30  |
| 2006 | 1896         | 198                        | 0.32  |
| 2007 | 1913         | 220                        | 0.53  |
| 2008 | 1944         | 211                        | 0.57  |
| 2009 | 1846         | 189                        | 0.39  |
| 2010 | 1811         | 160                        | 0.60  |
| 2011 | 1769         | 181                        | 0.38  |
| 2012 | 1747         | 170                        | 0.14  |
| 2013 | 1690         | 161                        | 0.75  |
| 2014 | 1627         | 173                        | 0.68  |
| 2015 | 1547         | 172                        | 0.51  |
| 2016 | 1414         | 158                        | 0.51  |

Table 1.1: Number of Hedge Fund Firms by Year

No. of firm is defined as the number of unique firms reported in HFR in a specific year. No. of firms filed LCAs is defined as the number of unique firms in HFR that filed LCA petition in that year. Aggregate fraction of foreign labor demand met is defined as the total number of H1-B lottery winners among firms in HFR divided by total number of H1-B lottery applications among firms in HFR in that year.

| State | No. of Firms | No. of Firms Filed LCAs | Aggregate Fraction of Foreign Labor Demand Met |
|-------|--------------|-------------------------|--|
| NY    | 1179         | 382                     | 0.45   |
| CA    | 496          | 103                     | 0.33   |
| IL    | 231          | 43                      | 0.43   |
| CT    | 229          | 73                      | 0.59   |
| TX    | 170          | 17                      | —  |
| FL    | 158          | 37                      | 0.00   |
| MA    | 157          | 43                      | 0.81   |
| NJ    | 136          | 24                      | 1.00   |
| PA    | 93           | 13                      | 0.25   |
| CO    | 57           | 6                       | —  |
| VA    | 50           | 12                      | 1.00   |
| GA    | 46           | 7                       | —  |
| NC    | 42           | 6                       | 0.00   |
| MN    | 38           | 9                       | 1.00   |
| WA    | 34           | 3                       | 0.00   |
| MD    | 30           | 7                       | 0.50   |
| TN    | 30           | 6                       | —  |
| OH    | 21           | 2                       | —  |
| WI    | 16           | 1                       | —  |
| NV    | 16           | 3                       | —  |
| OR    | 15           | 2                       | —  |
| DC    | 14           | 6                       | 0.00   |
| MO    | 12           | 1                       | —  |
| AZ    | 12           | 1                       | —  |
| KS    | 10           | 2                       | —  |
| UT    | 10           | 1                       | —  |
| KY    | 7            | 1                       | —  |
| NM    | 5            | 1                       | —  |
| HI    | 5            | 1                       | 1.00   |
| RI    | 5            | 1                       | 1.00   |
| WY    | 4            | 1                       | —  |
| IA    | 4            | 2                       | —  |
| AK    | 1            | 1                       | 1.00   |

Table 1.2: Number of Hedge Fund Firms by Headquarter State

No. of firm is defined as the number of unique firms reported in HFR in a specific state. No. of firms filed LCAs is defined as the number of unique firms in HFR that filed LCA petition in that state. Aggregate fraction of foreign labor demand met is defined as the total number of H1-B lottery winners among firms in HFR divided by total number of H1-B lottery applications among firms in HFR in that state.

Table 1.3: Summary statistics of firms that filed Labor Condition Applications

|                               | N   | Mean    | Std. Dev. | Min.    | 5-%ile  | 25-%ile | 50-%ile | 75-%ile | 95-%ile | Max     |
|-------------------------------|-----|---------|-----------|---------|---------|---------|---------|---------|---------|---------|
| Hedge Fund Firm Age           | 831 | 11.64   | 7.23      | 0.00    | 2.00    | 6.00    | 10.00   | 16.00   | 26.00   | 37.00   |
| Live-to-Dead Fund Count Ratio | 831 | 0.40    | 0.48      | 0.00    | 0.00    | 0.00    | 0.00    | 1.00    | 1.00    | 1.00    |
| Percent Fund using Leverage   | 758 | 0.66    | 0.45      | 0.00    | 0.00    | 0.00    | 1.00    | 1.00    | 1.00    | 1.00    |
| Percent Fund in U.S./Canada   | 758 | 0.42    | 0.48      | 0.00    | 0.00    | 0.00    | 0.00    | 1.00    | 1.00    | 1.00    |
| Percent Fund in Quantitative  | 758 | 0.15    | 0.35      | 0.00    | 0.00    | 0.00    | 0.00    | 0.00    | 1.00    | 1.00    |
| Percent Fund in Fundamentals  | 758 | 0.12    | 0.32      | 0.00    | 0.00    | 0.00    | 0.00    | 0.00    | 1.00    | 1.00    |
| In New York                   | 831 | 0.47    | 0.50      | 0.00    | 0.00    | 0.00    | 0.00    | 1.00    | 1.00    | 1.00    |
| Ln(Asset Under Management)    | 831 | 4.57    | 2.39      | 0.00    | 0.00    | 3.00    | 4.75    | 6.10    | 8.37    | 11.83   |
| Fund Inception                | 831 | 2001.83 | 6.67      | 1980.00 | 1990.00 | 1998.00 | 2003.00 | 2006.00 | 2012.00 | 2017.00 |

Hedge fund firm age is defined as the difference between current time and the first fund inception within the firm. Percent fund using leverage are defined as the average number of funds with leverage, weighted by assets. Percent fund in U.S./Canada are defined as the average number of funds with main investment region in the U.S. or Canada, weighted by assets. Percent fund in quantitative is defined as the percentage of funds that deploys quantitative strategies aforementioned, weighted by assets. Percent fund in quantitative is defined as the percentage of funds that deploys fundamental strategies aforementioned, weighted by assets. In New York indicates if the firm is located in New York state. Ln(Asset Under Management) is the sum of assets in each fund within the firm. Fund inception is defined as the first inception of funds within that firm.

Table 1.4: Summary statistics of firms that never filed Labor Condition Applications

|                               | N    | Mean    | Std. Dev. | Min. | 5-%ile | 25-%ile | 50-%ile | 75-%ile | 95-%ile | Max   |
|-------------------------------|------|---------|-----------|------|--------|---------|---------|---------|---------|-------|
| Hedge Fund Firm Age           | 3059 | 7.12    | 6.18      | 0.00 | 1.00   | 3.00    | 5.00    | 10.00   | 20.00   | 58.00 |
| Live-to-Dead Fund Count Ratio | 3059 | 0.24    | 0.42      | 0.00 | 0.00   | 0.00    | 0.00    | 0.00    | 1.00    | 1.00  |
| Percent Fund using Leverage   | 2743 | 0.62    | 0.48      | 0.00 | 0.00   | 0.00    | 1.00    | 1.00    | 1.00    | 1.00  |
| Percent Fund in U.S./Canada   | 2743 | 0.45    | 0.49      | 0.00 | 0.00   | 0.00    | 0.00    | 1.00    | 1.00    | 1.00  |
| Percent Fund in Quantitative  | 2743 | 0.19    | 0.39      | 0.00 | 0.00   | 0.00    | 0.00    | 0.00    | 1.00    | 1.00  |
| Percent Fund in Fundamentals  | 2743 | 0.09    | 0.28      | 0.00 | 0.00   | 0.00    | 0.00    | 0.00    | 1.00    | 1.00  |
| In New York                   | 3059 | 0.32    | 0.46      | 0.00 | 0.00   | 0.00    | 0.00    | 1.00    | 1.00    | 1.00  |
| ln(Asset Under Management)    | 3059 | 2.83    | 1.91      | 0.00 | 0.00   | 1.32    | 2.77    | 4.19    | 6.15    | 9.40  |
| Fund Inception                | 3059 | 2003.65 | 7.19      | 1959 | 1991   | 1999    | 2004    | 2009    | 2014    | 2017  |

Hedge fund firm age is defined as the difference between current time and the first fund inception within the firm. Percent fund using leverage are defined as the average number of funds with leverage, weighted by assets. Percent fund in U.S./Canada are defined as the average number of funds with main investment region in the U.S. or Canada, weighted by assets. Percent fund in quantitative is defined as the percentage of funds that deploys quantitative strategies aforementioned, weighted by assets. Percent fund in quantitative is defined as the percentage of funds that deploys fundamental strategies aforementioned, weighted by assets. In New York indicates if the firm is located in New York state.  $\ln(\text{Asset Under Management})$  is the sum of assets in each fund within the firm. Fund inception is defined as the first inception of funds within that firm.

Table 1.5: Determinants of LCA Filings

|                            | (1)                        | (2)                     | (3)<br>ln(No.<br>LCA<br>Workers)<br>OLS |
|----------------------------|----------------------------|-------------------------|---|
|                            | File<br>H-1B LCA<br>Probit | File<br>H-1B LCA<br>OLS |   |
| ln(Firm Age)               | -0.017**<br>(-2.28)        | -0.026***<br>(-3.69)    | -0.042***<br>(-3.25)                    |
| ln(Asset Under Management) | 0.040***<br>(17.77)        | 0.045***<br>(16.01)     | 0.057***<br>(10.26)                     |
| % Leverage                 | 0.026***<br>(3.28)         | 0.026***<br>(3.39)      | 0.041***<br>(3.59)                      |
| % Fund in Quant            | 0.044***<br>(4.01)         | 0.045***<br>(4.01)      | 0.057***<br>(3.56)                      |
| % Fund in Fundamentals     | -0.013<br>(-1.20)          | -0.020*<br>(-1.65)      | -0.038**<br>(-2.34)                     |
| Observations               | 23,626                     | 24,850                  | 24,850                                  |
| Adj. $R^2$                 | 0.178                      | 0.111                   | 0.107                                   |
| State FE                   | Yes                        | Yes                     | Yes                                     |
| Year FE                    | Yes                        | Yes                     | Yes                                     |
| Inception Year FE          | Yes                        | Yes                     | Yes                                     |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in columns 1 and 2 is the dummy variable that indicates if a firm files for LCAs in that specific year. The dependent variable in 3 is the number of workers a firm filed for a specific year. Ln(Firm Age) is defined as logarithm of firm's age. Ln(Asset Under Management) is defined as the logarithm of the sum of assets of funds within that firm. % Leverage is the percentage of funds that is levered, weighted by assets. % Fund in quant is the percentage of funds that is quantitative, weighted by assets. % Fund in fundamentals is the percentage of funds that is fundamental, weighted by assets. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.

Table 1.6: Determinants of Capped H-1B Visa Issuance

|                            | (1)<br>No. Cap<br>Visa<br>Granted | (2)<br>%H-1B<br>Demand Met |
|----------------------------|-----------------------------------|----------------------------|
| ln(LCA Cap Demand)         | 0.225***<br>(4.70)                |                            |
| ln(Firm Age)               | -0.028<br>(-1.16)                 | -0.036<br>(-1.61)          |
| ln(Asset Under Management) | 0.005<br>(0.93)                   | -0.004<br>(-1.03)          |
| % Leverage                 | 0.019<br>(1.30)                   | -0.009<br>(-0.53)          |
| % Fund in Quant            | 0.022<br>(0.96)                   | 0.002<br>(0.10)            |
| % Fund in Fundamentals     | -0.065***<br>(-4.04)              | -0.053*<br>(-1.95)         |
| Observations               | 2,316                             | 2,316                      |
| Adj. $R^2$                 | 0.172                             | 0.062                      |
| State FE                   | Yes                               | Yes                        |
| Year FE                    | Yes                               | Yes                        |
| Inception Year FE          | Yes                               | Yes                        |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in column 1 is the number of capped visa granted,  $I129\_CAP$ , as defined before. The dependent variable in 2 is the number of capped visa granted, scaled by the number of capped LCA filed for a specific year,  $MD$ .  $\ln(\text{LCA cap demand})$  is defined as the number of capped LCA filed for a specific year  $LCA\_CAP$ .  $\ln(\text{Firm Age})$  is defined as logarithm of firm's age.  $\ln(\text{Asset Under Management})$  is defined as the logarithm of the sum of assets of funds within that firm. % Leverage is the percentage of funds that is levered, weighted by assets. % Fund in quant is the percentage of funds that is quantitative, weighted by assets. % Fund in fundamentals is the percentage of funds that is fundamental, weighted by assets. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.



## B Natural Experiment: H-1B Lotteries

Table 1.7: Fraction of foreign high-skilled foreign labor demand met and quantitative funds

|                       | Quantitative          |                       |   | Fundamental           |                               |   |
|-----------------------|-----------------------|-----------------------|---|-----------------------|-------------------------------|---|
|                       | (1)<br>ln(Fund Count) | (2)<br>Fund Count > 0 | (3)<br>Asset-Weighted<br>Percentage<br>of Funds | (4)<br>ln(Fund Count) | (5)<br>Percentage<br>of Funds | (6)<br>Asset-Weighted<br>Percentage<br>of Funds |
| % H-1B Demand Met     | 0.108**<br>(2.30)     | 0.067*<br>(1.79)      | 0.037<br>(1.29)                                 | -0.069<br>(-1.64)     | -0.071***<br>(-2.82)          | -0.082***<br>(-3.17)                            |
| Filed Capped H-1B LCA | 0.181***<br>(4.01)    | 0.151***<br>(4.69)    | 0.077***<br>(3.16)                              | -0.035<br>(-1.05)     | -0.062***<br>(-3.14)          | -0.078***<br>(-4.47)                            |
| ln(Firm Age)          | 0.008<br>(0.76)       | 0.001<br>(0.09)       | -0.007<br>(-0.83)                               | -0.004<br>(-0.48)     | -0.009<br>(-1.25)             | -0.009<br>(-1.28)                               |
| ln(Total Firm Assets) | 0.004<br>(0.92)       | -0.009**<br>(-2.52)   | -0.015***<br>(-5.27)                            | 0.024***<br>(5.23)    | 0.009***<br>(2.87)            | 0.015***<br>(4.12)                              |
| Observations          | 16,688                | 16,688                | 16,688  | 16,688                | 16,688                        | 15,128  |
| Adj. $R^2$            | 0.086                 | 0.078                 | 0.080   | 0.069                 | 0.041                         | 0.046   |
| Year FE               | Yes                   | Yes                   | Yes   | Yes                   | Yes                           | Yes   |
| Inception Year FE     | Yes                   | Yes                   | Yes   | Yes                   | Yes                           | Yes   |
| State FE              | Yes                   | Yes                   | Yes   | Yes                   | Yes                           | Yes   |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in column 1 is the logarithm of the number of quantitative funds. The dependent variable in column 2 is the a dummy variable indicates if a firm contains a quantitative fund. The dependent variable in column 3 is the percentage of quantitative fund within a firm, weighted by assets. The dependance variable in column 4 is the logarithm of the number of fundamental funds within the firm. The dependent variable in column 5 is the percentage of fundamental fund within a firm. The dependent variable in column 6 is the percentage of capped LCA filed for a specific year,  $MD$ . Filed capped H-1B LCA is the dummy variable indicates if a firm filed capped LCA for that year  $DCAP$ , as defined before. Ln(Firm Age) is defined as logarithm of firm's age. Ln(Asset Under Management) is defined as the logarithm of the sum of assets of funds within that firm. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.

Table 1.8: Fraction of foreign high-skilled foreign labor demand met and fund strategy diversification

|                       | (1)<br>Region HHI    | (2)<br>ln(No. Unique<br>Investment Regions) | (3)<br>ln(Overall<br>Fund Count) |
|-----------------------|----------------------|---|----------------------------------|
| % H-1B Demand Met     | -0.032***<br>(-2.59) | 0.041**<br>(2.01)                           | 0.077<br>(1.62)                  |
| Filed Capped H-1B LCA | -0.039***<br>(-3.33) | 0.114***<br>(5.49)                          | 0.242***<br>(5.46)               |
| ln(Firm Age)          | 0.002<br>(0.56)      | 0.004<br>(0.74)                             | 0.005<br>(0.39)                  |
| ln(Total Firm Assets) | -0.008***<br>(-7.41) | 0.018***<br>(8.87)                          | 0.087***<br>(16.33)              |
| Observations          | 15,128               | 16,688                                      | 16,688                           |
| Adj. $R^2$            | 0.089                | 0.168                                       | 0.340                            |
| Year FE               | Yes                  | Yes   | Yes                              |
| Inception Year FE     | Yes                  | Yes   | Yes                              |
| State FE              | Yes                  | Yes   | Yes                              |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in column 1 is the HHI of investment region, weighted by assets. The dependent variable in column 2 is the logarithm of the number of unique investment regions within a firm. The dependent variable in column 3 is the logarithm of the total number of funds within a firm. % H-1B demand met is the number of capped visa granted, scaled by the number of capped LCA filed for a specific year,  $MD$ . Filed capped H-1B LCA is the dummy variable indicates if a firm filed capped LCA for that year  $D_{LCA}$ , as defined before. Ln(Firm Age) is defined as the logarithm of the firm's age. Ln(Asset Under Management) is defined as the logarithm of the sum of assets of funds within that firm. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.

Table 1.9: Fraction of foreign high-skilled foreign labor demand met and fund region focus diversification

|                       | (1)                  | (2)                            | (3)                           | (4)                    |
|-----------------------|----------------------|--------------------------------|-------------------------------|------------------------|
|                       | Main Strategy HHI    | ln(No. Unique Main Strategies) | ln(No. Unique Sub-Strategies) | ln(Overall Fund Count) |
| % H-1B Demand Met     | -0.045***<br>(-3.12) | 0.054***<br>(2.62)             | 0.087***<br>(3.04)            | 0.077<br>(1.62)        |
| Filed Capped H-1B LCA | -0.073***<br>(-4.89) | 0.139***<br>(6.18)             | 0.200***<br>(6.41)            | 0.242***<br>(5.46)     |
| ln(Firm Age)          | -0.000<br>(-0.02)    | 0.003<br>(0.79)                | 0.008<br>(1.17)               | 0.005<br>(0.39)        |
| ln(Total Firm Assets) | -0.010***<br>(-7.16) | 0.015***<br>(8.80)             | 0.031***<br>(11.18)           | 0.087***<br>(16.33)    |
| Observations          | 15,128               | 16,688                         | 16,688                        | 16,688                 |
| Adj. $R^2$            | 0.102                | 0.159                          | 0.221                         | 0.340                  |
| Year FE               | Yes                  | Yes                            | Yes                           | Yes                    |
| Inception Year FE     | Yes                  | Yes                            | Yes                           | Yes                    |
| State FE              | Yes                  | Yes                            | Yes                           | Yes                    |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in column 1 is the HHI of main investment strategy, weighted by assets. The dependent variable in column 2 is the logarithm of the number of unique investment main strategies within a firm. The dependent variable in column 3 is the logarithm of the number of unique investment sub strategies within a firm. The dependent variable in column 4 is the logarithm of the total number of funds within a firm. % H-1B demand met is the number of capped visa granted, scaled by the number of capped LCA filed for a specific year,  $MD$ . Filed capped H-1B LCA is the dummy variable indicates if a firm filed capped LCA for that year  $D_{LCA}$ , as defined before. Ln(Firm Age) is defined as logarithm of firm's age. Ln(Asset Under Management) is defined as the logarithm of the sum of assets of funds within that firm. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.

## C Natural Experiment: H-1B Visa Reform Act of 2004

Table 1.10: Treatment dynamics on quant vs. fundamental fund strategies

|                               | Quantitative      |                   |                                    | Fundamental       |                     |                                    |
|-------------------------------|-------------------|-------------------|------------------------------------|-------------------|---------------------|------------------------------------|
|                               | (1)               | (2)               | (3)                                | (4)               | (5)                 | (6)                                |
|                               | ln(Fund Count)    | Fund Count > 0    | Asset-Weighted Percentage of Funds | ln(Fund Count)    | Percentage of Funds | Asset-Weighted Percentage of Funds |
| Treatment × Year= 2001        | -0.014<br>(-0.73) | -0.009<br>(-0.55) | 0.011<br>(1.33)                    | -0.006<br>(-0.45) | -0.011**<br>(-2.07) | -0.008<br>(-1.24)                  |
| Treatment × Year= 2002        | -0.007<br>(-0.55) | -0.001<br>(-0.05) | 0.003<br>(0.45)                    | -0.007<br>(-0.73) | -0.004<br>(-1.04)   | -0.008*<br>(-1.71)                 |
| Treatment × Year= 2004        | -0.006<br>(-0.59) | -0.007<br>(-0.76) | -0.008<br>(-1.33)                  | 0.020**<br>(2.03) | 0.004*<br>(1.75)    | 0.008**<br>(2.12)                  |
| Treatment × Year= 2005        | 0.007<br>(0.34)   | -0.001<br>(-0.05) | -0.015<br>(-1.52)                  | 0.024**<br>(1.99) | 0.006**<br>(2.08)   | 0.015***<br>(3.02)                 |
| Treatment × Year= 2006        | 0.024<br>(0.90)   | -0.003<br>(-0.15) | -0.024*<br>(-1.65)                 | 0.032*<br>(1.93)  | 0.009*<br>(1.82)    | 0.020***<br>(3.36)                 |
| Observations                  | 7,905             | 7,905             | 7,905                              | 7,905             | 7,905               | 6,752                              |
| Adj. $R^2$                    | 0.898             | 0.906             | 0.957                              | 0.934             | 0.983               | 0.983                              |
| Firm Characteristics Controls | Yes               | Yes               | Yes                                | Yes               | Yes                 | Yes                                |
| Year FE                       | Yes               | Yes               | Yes                                | Yes               | Yes                 | Yes                                |
| Firm FE                       | Yes               | Yes               | Yes                                | Yes               | Yes                 | Yes                                |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in column 1 is the logarithm of the number of quantitative funds. The dependent variable in column 2 is the dummy variable indicates if a firm contains a quantitative fund. The dependent variable in column 3 is the percentage of quantitative fund within a firm, weighted by assets. The dependent variable in column 4 is the logarithm of the number of fundamental funds within the firm. The dependent variable in column 5 is the percentage of fundamental fund within a firm. The dependent variable in column 6 is the percentage of fundamental fund within a firm, weighted by assets. Treatment × Year = n is the indicator variable of year = n, interacted with the treatment dummy. The treatment dummy is defined as whether a firm has filed LCA prior to 2004. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.

Table 1.11: Average treatment effect on quant vs. fundamental fund strategies (2001 to 2006)

|                                     | Quantitative       |                    |                                    | Fundamental        |                     |                                    |
|-------------------------------------|--------------------|--------------------|------------------------------------|--------------------|---------------------|------------------------------------|
|                                     | (1)                | (2)                | (3)                                | (4)                | (5)                 | (6)                                |
|                                     | ln(Fund Count)     | Fund Count > 0     | Asset-Weighted Percentage of Funds | ln(Fund Count)     | Percentage of Funds | Asset-Weighted Percentage of Funds |
| Treatment $\times$ Year $\geq$ 2004 | 0.013<br>(0.71)    | -0.001<br>(-0.07)  | -0.019*<br>(-1.81)                 | 0.029**<br>(2.00)  | 0.011***<br>(2.59)  | 0.018***<br>(3.68)                 |
| ln(Firm Age)                        | -0.003<br>(-0.24)  | -0.001<br>(-0.11)  | 0.001<br>(0.20)                    | -0.008<br>(-0.92)  | 0.007**<br>(2.09)   | 0.009**<br>(2.16)                  |
| ln(Total Firm Assets)               | 0.022***<br>(5.25) | 0.010***<br>(2.78) | 0.001<br>(0.36)                    | 0.014***<br>(5.21) | 0.001<br>(0.67)     | -0.001<br>(-0.50)                  |
| Observations                        | 7,905              | 7,905              | 7,905                              | 7,905              | 7,905               | 7,341                              |
| Adj. $R^2$                          | 0.898              | 0.906              | 0.957                              | 0.934              | 0.983               | 0.984                              |
| Year FE                             | Yes                | Yes                | Yes                                | Yes                | Yes                 | Yes                                |
| Firm FE                             | Yes                | Yes                | Yes                                | Yes                | Yes                 | Yes                                |

Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively. T-statistics are in parentheses. Standard errors are clustered at the hedge fund level. The dependent variable in column 1 is the logarithm of the number of quantitative funds. The dependent variable in column 2 is the a dummy variable indicates if a firm contains a quantitative fund. The dependent variable in column 3 is the percentage of quantitative fund within a firm, weighted by assets. The dependent variable in column 4 is the logarithm of the number of fundamental funds within the firm. The dependent variable in column 5 is the percentage of fundamental fund within a firm. The dependent variable in column 6 is the percentage of fundamental fund within a firm, weighted by assets. Treatment  $\times$  Year  $\geq$  2004 = 1 if a firm is in the treatment group and the year is after 2004, otherwise 0. The treatment group is defined as firms that had filed LCA prior to 2004. Ln(Firm Age) is defined as the logarithm of the firm's age. Ln(Asset Under Management) is defined as the logarithm of the sum of assets of funds within that firm. Reported F statistic is the Kleibergen-Paap rk Wald F statistic.

## CHAPTER 2

# Peer Effects on Firm-Level Strategic Debt Choices: A Theoretical Model

Modigliani and Miller 1958 built the foundation of theories on corporate capital structure. In that paper the authors argued that in complete markets a firm's capital structure does not affect its value. Robichek and Myers 1966 introduced the trade-off theory by accounting for the existence of bankruptcy cost, along with corporate taxes. Firms would choose a leverage level to balance the marginal tax benefit and the marginal bankruptcy cost of raising debt. Jensen and Meckling 1976 then proposed a model with agency problems, and concluded with the pecking order theory which ranked the cost of financing by cash, debt and equity as ascending. Myers 1977 suggested that issuing risky debt reduces the present market value of a firm that holds real options by inducing a sub-optimal investment strategy or by forcing the firm and its creditors to bear the costs of avoiding the sub-optimal strategy. Overall, most theories on corporate capital structure predict that a firm's optimal debt choices are influenced by firm-specific determinants such as its marginal tax rate, asset illiquidity, incentive structure, information environment, and other industry- or economy-wide determinants. However, a firm's debt choice is often assumed to be made independently of the choices of their peers.

Recent empirical research has begun to explore the role of peer firm behavior in affecting capital structures. Industry average leverage ratios are found to be an important determinant of firms' capital structures (Welch 2004, MacKay and Phillips 2005, and Frank and Goyal 2009). Leary and Roberts 2014 find that firms' capital structures are responsive to peer firms' idiosyncratic equity shocks, which are used as exogenous indicators of changes in peer firm financial choices. The action channel of peer effects in an industry also amplifies the

impact of any exogenous capital structure determinant.

Theoretically, peer effects in capital structure can arise from a variety of mechanisms. Several models are presented to account for mimicking behavior in capital structure. For example, some models predict that firms are motivated to mimic the more conservative leverage policies of their peers, either because high leverage invites predatory price competition from less-levered rivals (Bolton and Scharfstein 1990) or because high-leverage firms under-invest during industry downturn and loss market shares to more conservatively financed competitors (Chevalier and Scharfstein 1996). Additionally, rational herding models (Devenow and Welch 1996) are also used to explain mimicking. Informational factors may play a big role. Free-riding in information acquisition or relative performance evaluation for managers may lead to herd behavior in capital structure policies (Zeckhauser, Patel, and Hendricks 1991). When a firm's own signal is noisy and optimization is costly, managers may rationally put more weight on the decision of other firms in the industry that are perceived as having greater expertise (Banerjee 1992, Conlisk 1980, Bikhchandani, Hirshleifer, and Welch 1998). In addition, learning mechanisms may be at work. Managers need not completely ignore their own information. Rather, it is sufficient that they update their priors in a Bayesian manner based on the observed actions of other firms (Romer 1993 and Trueman 1994). As a result, their decision will be pulled toward those of their peers, relative to what it would be if they solely relied on their own information. Reputation is another explanatory factor for mimicking, as managers may mimic other firms' policies to influence their perceived quality in the labor market (Scharfstein and Stein 1990 and Zwiebel 1995). Empirically, Leary and Roberts 2014 find that smaller and more financial constrained firms are more sensitive to the financial choices of their peers, suggesting that mimicking behavior is strongest among those firms with the greatest learning motive and the greatest need to build reputation.

However, most of the existing models only serve to explain converging financial choices among peer firms. Moreover, the general equilibrium implications of such responses to peer actions are also unclear. The model in this paper presents a mechanism derived from Robichek and Myers 1966 through which the financial choices of peer firms enter a firm's strategic consideration without assuming mimicking behavior. Rather, it explores the chan-

nel of peer influence through how liquidation values in times of financial distress might be affected by peer choices of leverage level. Strategic responses to peer debt choices generate optimal debt levels at Nash equilibrium.

Shleifer and Vishny 1992 presents a general equilibrium model that explores the economy- and industry-wide determinants of asset illiquidity. When a firm in financial distress needs to sell assets, its industry peers are likely to be experiencing problems themselves, leading to asset sales at prices below value in best use. Such asset illiquidity makes assets cheap in bad times, and so ex ante is a cost of leverage. As a result, optimal debt levels are restricted by asset illiquidity (e.g. cyclical and growth assets have a lower optimal level of debt finance) and dependent on the leverage of other firms in the same industry. Moreover, this model suggests that the debt choices of the firms in the same industry could be very divergent instead of mimicking each other. While each individual firm can have a high or a low debt level depending on which equilibrium obtains, there is an aggregate industry-level optimal debt capacity.

Base on such understandings of the impact of asset illiquidity on optimal debt choices, the model in this paper considers a two-period two-firm game. Each strategically chooses a debt level in time 0 in response to expected utility of time 1, which is a function of its cash flow and expected cost of financial distress, which is further influenced by peer choices of liquidation. The next section introduces the benchmark model, while following sections explore its implications and consider further variations to the basic model.

## 2.1 The Model

### 2.1.1 The Benchmark Model (no peer effects)

The model has two periods, 0 and 1. For simplicity, assume zero interest rate. There are two firms in the industry,  $F$  and  $F'$ , which are completely symmetrical. The capital structure of each is determined in period 0 by choosing a debt level,  $D$  and  $D'$ , which bring immediate tax benefits of  $\tau D$  and  $\tau D'$ . In period 1, each firm realizes a cash flow,  $V$  and  $V'$ , which are



i.i.d  $U[0, 1]$ . The distribution of each other's cash flow is common knowledge in period 0; in period 1, the uncertainty of a firm's own cash flow is resolved, but they still don't know the realized cash flow of the opponent.

In period 1, if a firm's realized cash flow is above its debt level, i.e.  $V \geq D$ , then the firm faces no financial distress and its value is  $V + \tau D$ . The firm will face financial distress if  $V < D$ , in which case it has two options, continuation or liquidation. If the firm chooses to continue, it will incur a cost of  $\phi(D - V)$ . In this case, the firm's value would be  $V - \phi(D - V) + \tau D$ . In the case of liquidation, the firm liquidates its assets for a price of  $P$  and leaves with a total value of  $P + \tau D$ . Notice that this baseline model does not yet contain peer effects on liquidation values; therefore, a firm's optimal debt level is not yet affected its peer's debt choice.

Firm  $F$ 's objective is to choose an optimal debt level  $D$  in period 0 to maximize its expected value in period 1. The same holds for firm  $F'$ . The solution is straightforward. We can conclude that a firm will continue for  $\forall V \geq \Lambda$  and liquidate for  $\forall V < \Lambda$ , where  $\Lambda = \frac{P + \phi D}{1 + \phi} < D$ . So in period 0, a firm's expected utility can be written as

$$U(D) = \int_0^\Lambda P dV + \int_\Lambda^D (V - \phi(V - D)) dV + \int_D^1 V dV + \tau D \quad (2.1)$$

As a result, the optimal debt level for both firms is  $D^* = P + (1 + \phi)/\phi\tau$ .

### 2.1.2 Introducing Peer Effects on Liquidation Prices

Now we can introduce peer effects on liquidation prices to the baseline model. According to Shleifer and Vishny 1992, asset illiquidity is exacerbated when firms experience economy- or industry-wide shocks. In bad times, when more firms in the same industry experience difficulties and may be forced to liquidate their assets, liquidation prices will plummet not only because of more distressed assets are being offered for sale but also because industry insiders, who would be the highest valuation buyers in normal times, are likely to have cash flow problems of their own and thus constrained in their purchasing capacity, leading to fire

sales of the assets to inefficient industry outsiders at prices much below value in best use.

Based on this understanding, in our model let's suppose that the liquidation price of a firm's assets is  $P_h$  when only one firm is liquidating and  $P_l$  when both firms are liquidating, where  $P_h > P_l$ .

In period 1, firms play a one shot game by choosing whether to liquidate. Let the function  $L(V, D, D')$  represents firm  $F$ 's liquidation choice.  $L$  equals 1 if firm  $F$  chooses to liquidate, equals 0 if it chooses to continue. Similarly,  $L'$  represents the liquidation choice of firm  $F'$ .

Consider the liquidation choice of firm  $F$ . It will liquidate (i.e.  $L = 1$ ) if and only if the expected utility of liquidation is greater than the expected utility of continuation:

$$\mathbb{E}[P_l \times L' + P_h \times (1 - L')] - [V - \phi(D - V)] > 0 \quad (2.2)$$

In period 0, firm  $F$  chooses its debt policy  $D(D')$  so that given the opponent's debt choice  $D'$ ,  $D(D')$  maximizes its expected utility in period 1. Since the two firms are perfectly symmetrical, the same conditions apply for firm  $F'$ .

**Lemma 1** *There exists a single cutoff of the cash flow  $V_c$ , above which the firm will choose to continue, and below which the firm will choose to liquidate.*

**Proof 1** *Consider Equation 2.2. Since the cash flow  $V$  is i.i.d., so the first part, which can be written as  $E[P_h - (P_h - P_l) \times L']$ , is independent of  $V$ . And the second part  $[V - \phi(D - V)]$  is strictly continuously increasing w.r.t.  $V$  (linear function). Therefore, the whole left-hand-side expression is a strictly decreasing linear function w.r.t.  $V$ . Thus, if there exists a  $\tilde{V}$  that makes the left-hand-side expression negative, then  $\forall V > \tilde{V}$  the expression should be negative. If  $\tilde{V}$  makes this expression positive, then  $\forall V < \tilde{V}$  the expression should be positive. Furthermore, when  $V = 0$  the expression is positive and when  $V = D$  it is negative. Hence,  $\exists V_c \in [0, D]$  such that when  $V = V_c$  the expression equals 0. Then,  $\forall V > V_c$  the firm will choose to continue, and  $\forall V < V_c$  the firm will choose to liquidate.*

The next step is to solve the response function  $D(D')$  and the optimal equilibrium. Since

two firms are symmetrical, their  $V_c$  should be the same given debt choices. We can get  $V_c$  as a function of  $D$ ,  $D'$ ,  $\phi$ ,  $P_h$ , and  $P_l$ :

$$\begin{aligned}
V_c &= A \cdot D + B \cdot D' + C \\
A &= \frac{\phi \cdot (1 + \phi)}{(1 + \phi)^2 - (P_h - P_l)^2} > 0 \\
B &= \frac{-\phi \cdot (P_h - P_l)}{(1 + \phi)^2 - (P_h - P_l)^2} < 0 \\
C &= \frac{P_h}{1 + \phi + P_h - P_l} > 0
\end{aligned} \tag{2.3}$$

This result is very intuitive. The cutoff  $V_c$  increases w.r.t.  $D$  and decreases w.r.t.  $D'$ . A higher  $D$  means higher costs in financial distress, hence the firm is more likely to liquidate even under a higher cash flow, hence the higher  $V_c$ . A higher  $D'$  means the opponent firm is more likely to liquidate, which lowers the expected utility of the liquidation option, thus the firm will be less likely to liquidate, hence the lower  $V_c$ .

Then we can solve for the optimal debt level:

$$\begin{aligned}
D^* &= \frac{\beta}{1 - \alpha} \\
\alpha &= -\frac{-A^2(P_h - P_l) - B^2(P_h - P_l) - (1 + \phi)A \cdot B + \phi B}{2[-A \cdot B(P_h - P_l) + \frac{1+\phi}{2}(1 - A^2) - \phi(1 - A) - \frac{1}{2}]} \\
\beta &= -\frac{-A \cdot C(P_h - P_l) + A \cdot P_l - B \cdot C(P_h - P_l) - (1 + \phi)A \cdot C + \phi C + \tau}{2[-A \cdot B(P_h - P_l) + \frac{1+\phi}{2}(1 - A^2) - \phi(1 - A) - \frac{1}{2}]}
\end{aligned} \tag{2.4}$$

From now on this will be our baseline peer-effect model.

## 2.2 Properties of the Baseline Peer-Effect Model

1. Optimal debt level increases with tax benefit.

As shown in figure 2.1, the optimal debt level  $D^*$  increases with the tax benefit of debt

$\tau$ , which is consistent with classical theories of capital structure.

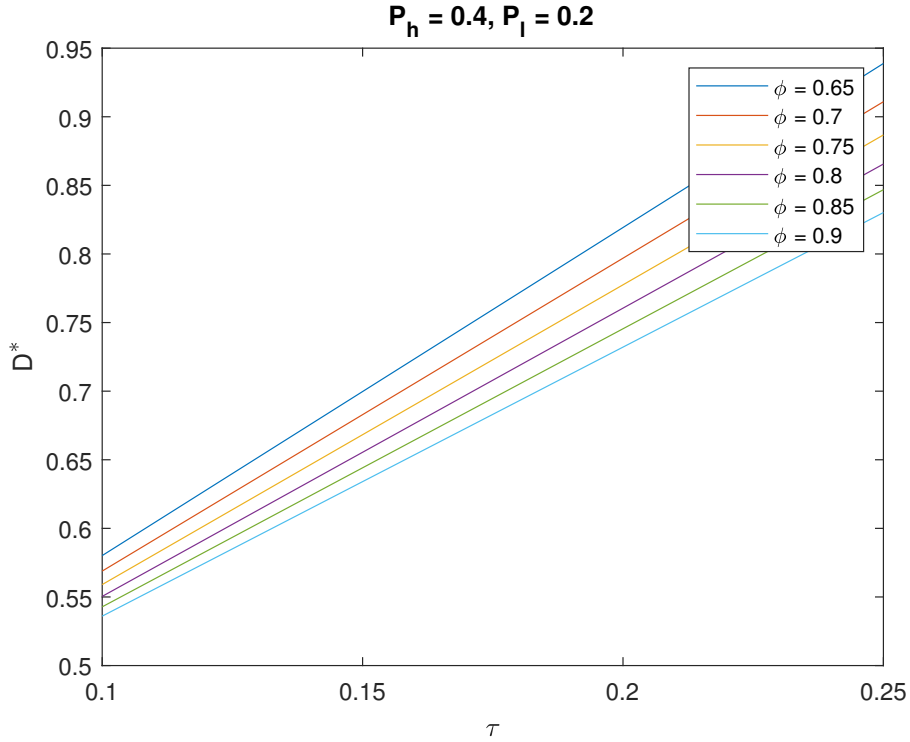


Figure 2.1: Optimal Debt vs. Tax Benefit

## 2. Optimal debt vs. peer effects on liquidation prices

Peer effects on liquidation prices can be captured by the difference between  $P_h$  and  $P_l$ . The following two scenarios discuss the dynamic between  $D^*$  and the difference of the two liquidation prices.

### (a) Scenario 1. Fix $P_h$ , vary the rate of $P_l/P_h$ .

This is the more intuitive approach. According to Shleifer and Vishny 1992's theory about the determinants of asset illiquidity, when a firm is liquidating due to idiosyncratic factors (e.g. mismanagement), it is more likely that other firms in the industry will compete to drive the asset price in liquidation up to the fundamental value in best use. By contrast, during an industry-wide recession, more firms experience financial distress and are forced to liquidate, liquidation prices tend to be much depressed below value in best use when assets are forced to be sold to less efficient industry outsiders.

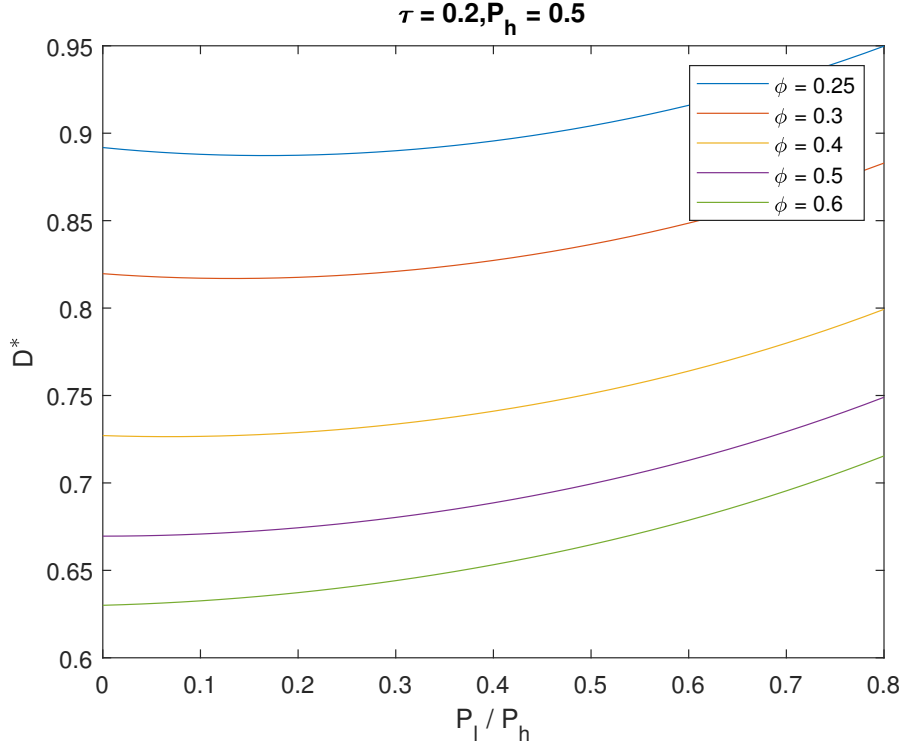


Figure 2.2: Optimal Debt and Price Discount ( $P_h$  fixed)

Corresponding to our mode,  $P_h$  can be interpreted as being more approximate to value in best use, while  $P_l$  as the depressed liquidation price in recession. Hence, a lower ratio of  $P_l/P_h$  represents a higher extent to which simultaneous liquidations in the same industry (often due to industry-wide negative shocks to cash flow) distort the liquidation prices below value of best use due to asset illiquidity. Figure 2.2 shows that given a level of  $P_h$  and  $\phi$  (the rate of financial distress cost if the firm chooses to continue), the optimal debt level  $D^*$  decreases with a lower  $P_l/P_h$ . This means that as the discount of liquidation price becomes more severe when both firms choose to liquidate, the optimal level of firm leverage is lowered because of the worsened asset illiquidity.

- (b) Scenario 2. Fix  $P_h + P_l$ , vary the spread of  $P_h - P_l$ . In other words,  $P_h$  can be written as  $mean + spread$ ,  $P_l = mean - spread$ , where  $mean$  is fixed.

If we don't think of liquidation as depressed sales of firm assets that tend to be at prices lower than value of best use to varying degrees, but as transactions

conforming more to normal supply-demand relations, we may capture the price response to supply by following set-up. Suppose there is a mean value the represents the assets' value of best use. When only one firm is liquidating, the price  $P_h$  is higher than the mean value because of supply scarcity. When both firms are liquidating, the price  $P_l$  is lower than the mean value because of over supply. In other words, a bigger spread ( $P_h - P_l$ ) around the mean represents higher price elasticity to changes in supply.

Figure 2.3 shows the dynamic between optimal debt level  $D^*$  and the price spread around a fixed mean. That is,  $P_h + P_l$  is fixed at 0.7, while  $P_h - P_l$  varies. Interestingly, this result seems to be the opposite of that in scenario 1. When the mean is fixed, a bigger price spread between  $P_h$  and  $P_l$  leads to a higher optimal debt level. How do we interpret this result? The key is that under this setting, not only is liquidation price negatively affected by more firms choosing to liquidate, it is also positively affected by fewer firms choose to liquidate. Whenever the probability of the opponent firm choosing to continue under financial distress ( $Prob$ , indicated by the upward tilting red line in Figure 2.3) is strictly higher than 0.5, the original firm's expected liquidation price  $Prob \times P_h + (1 - Prob) \times P_l$  is increasing w.r.t. the price spread  $P_h - P_l$ . As a result, when the mean is fixed, a bigger price spread means a higher expected liquidation price, and consequently a higher optimal debt level. Following this logic, when the probability of continuing falls below 0.5, the relationship between  $D^*$  and  $P_h - P_l$  will be reversed. However, that situation is much rarer.

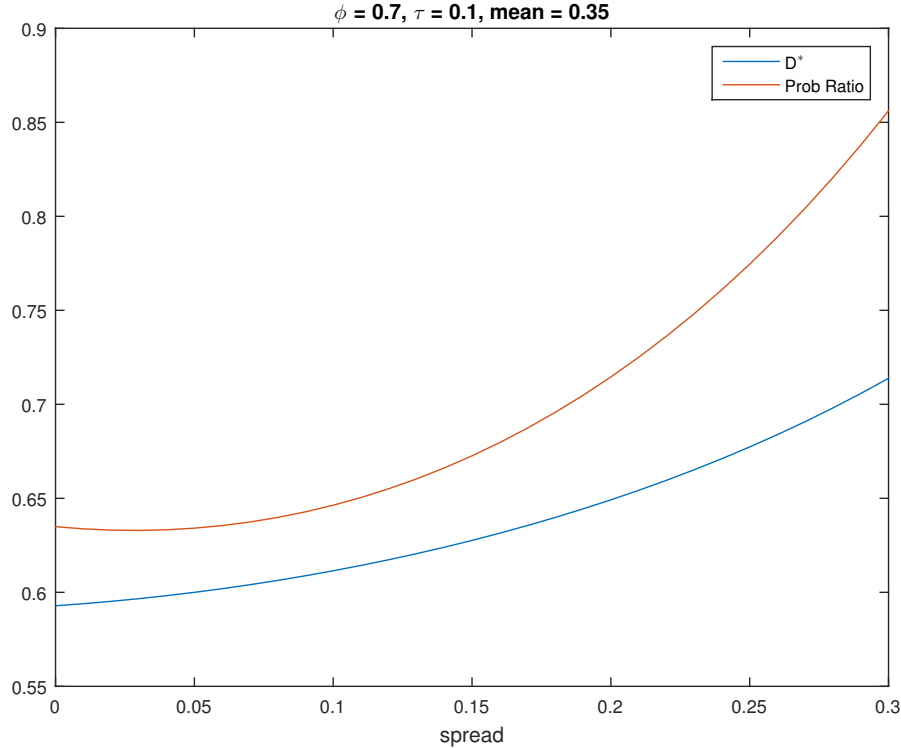


Figure 2.3: Optimal Debt and Price Spread ( $P_h + P_l$  fixed)

### 2.3 Variation: Adding Different Firm Sizes

Up till now the two firms in the baseline model are entirely symmetrical. Now we can add in the variation that two firms are not of the same size. Suppose firm  $F$  is the larger one, with its scale being 1 and its cash flow  $V \sim U[0, 1]$ .  $F'$  is the smaller firm, with its scale being  $\Gamma$  and its cash flow  $V' \sim U[0, \Gamma]$ . How does the optimal leverage ratio of each firm (now defined as the absolute optimal debt amount scaled by firm size:  $D^*$  and  $D'^*/\Gamma$ ) change in response to the varying size difference between the two firms?

Figure 2.4 shows the result. As the size difference between the two firms enlarges ( $\Gamma$  decreases), the optimal leverage ratio for the larger firm goes down while that of the smaller firm goes up from the common  $D^*$  when the two firms are symmetrical. The bigger the difference in firm size, the more their optimal leverage ratios diverge.

How do we interpret this result? Apparently, this is contrary to the predictions of learning and reputation models that smaller firms are more likely to mimic the capital structure of

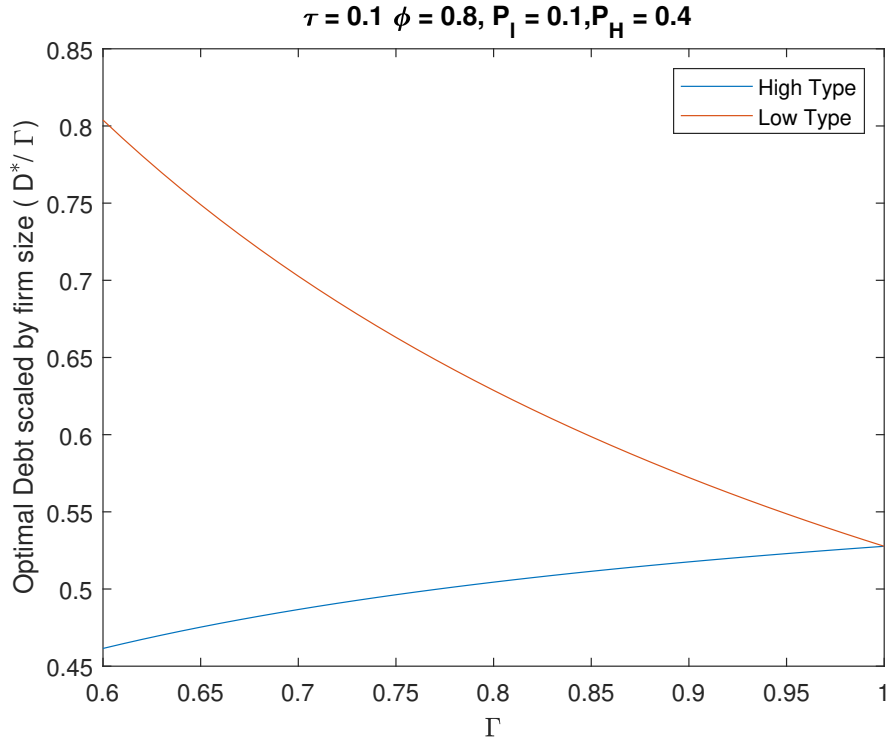


Figure 2.4: Optimal Debt and Firm Size

industry leaders. Rather, Shleifer and Vishny 1992 has considered a situation where firms have divergent optimal leverage, while the industry as a whole has a more or less stable overall debt capacity. My finding corresponds to an equilibrium where the smaller firm has a lot of debt and is liquidated in the depression, while the bigger firm, in recognition that there is an opportunity to buy the liquidating firm in the depression, chooses to forgo the debt overhang just to take this opportunity. In turn, the fact that the bigger firm has little debt and is more likely to continue in period 1 makes it attractive for the smaller firm to have more debt.

## 2.4 Variation: Adding Correlation in Cash Flow

It is natural to assume that firms in the same industry should receive positively correlated cash flows. Not only would this correlation affect firms' debt choices through the channel of correlated liquidation, its effect also work through the channel of information, as higher correlation means that firms can better infer each other's cash flow based on their own



outcome. Correlated cash flows can be introduced to the baseline model by the following transformation:

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \triangleq \begin{bmatrix} \Phi^{-1}(V_1) \\ \Phi^{-1}(V_2) \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \quad (2.5)$$

Where  $\Phi(\cdot)$  is the CDF of standard normal distribution.  $X_1$  and  $X_2$  are jointly normally distributed. In period 1, firm 1 will receive a cash flow of  $V_1 = \Phi(X_1)$  and firm 2 will get  $V_2 = \Phi(X_2)$ . The distributions of  $X_i$  and  $V_i$  are common knowledge in period 0. Notice that the marginal distribution of  $V_i$  is still a standard uniform distribution on  $[0, 1]$ . When  $\rho = 0$ , this model converges to the baseline model; and when  $\rho = 1$ , both firms receive the same amount of cash flow. Basically, by introducing  $X_i$ , this transformation imposes a normal copula on the joint distribution of the two firms' cash flows. Correlated cash flows also mean that one's own cash flow now contains information about the opponent's cash flow. Notice that while  $X_i \sim N(0, 1)$ ,  $X_{-i}|X_i \sim N(\rho X_i, 1 - \rho^2)$ . So in period 1, firms now have more accurate information about each other's cash flow based on their own realized cash flow.

It is worth noting that this information effect of cash flow correlation means that there are now possibly multiple equilibrium with regard to the firms' choices of liquidation. For an extreme example, assume  $\rho = 1$ , meaning that two firms receive the exact same cash flow in period 1. So both firms know precisely each other's cash flow once its own cash flow is realized. For the parameter setting of  $\phi|D_1 - D_2| < P_h - P_l$  and WOLOG assume  $D_1 > D_2$ , for  $\frac{P_l + \phi D_1}{1 + \phi} < \forall V < \frac{P_h + \phi D_2}{1 + \phi}$ , both the continuation values of firm 1 and firm 2 will be between  $P_l$  and  $P_h$ , because  $P_l < V - \phi(D_1 - V) < V - \phi(D_2 - V) < P_h$ . In this case, either of the firm has the incentive to liquidate, but not both. Hence there are two equilibria: firm 1 liquidates and firm 2 continues, or firm 1 continues and firm 2 liquidates. For now we assume that the firm with higher debt will liquidate.

Another complexity is that multiple cutoff points of  $V_c$  also become a possibility. Recall the firm's liquidation choice expression in equation 2.2. Again the  $[V - \phi(D - V)]$  is increasing w.r.t.  $V$ . Now in case of a positive correlation between the cash flows, the first term

$E[P_h - (P_h - P_l) \times L']$  is no longer independent of  $V$  but also becomes increasing w.r.t  $V$ . As a result, the argument of a single cutoff point of  $V_c$  in lemma 1 is not applicable anymore.

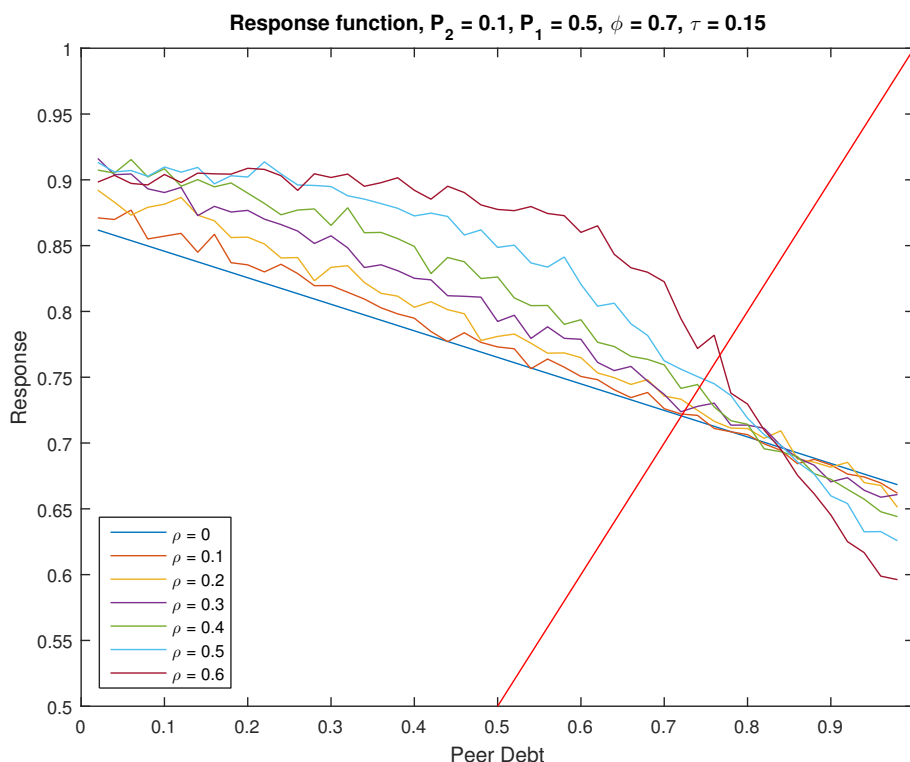


Figure 2.5: Responses to Peer Debt with Varying  $\rho$

Setting aside these extreme cases, let's now look at how different levels of correlation affect the optimal debt levels under the assumption of single equilibrium and single cutoff point. As shown in Figure 2.5, with a higher correlation  $\rho$ , the response function of  $D(D')$  becomes more steep around equilibrium points. The equilibrium optimal debt level, as indicated by the crossing point of the response function and the 45 degree diagonal line, increases with the correlation level  $\rho$ .

Intuitively, one might expect that a higher level of cash flow correlation should be associated with a lower optimal debt level, since higher correlation implies higher possibility of both firms facing financial distress and liquidating. However, Figure 2.5 apparently shows the opposite, that the equilibrium of both firms' optimal debt level increases with the correlation of their cash flows. The reason behind this phenomenon is that the two firms are locked in a situation similar to the prisoner's dilemma. When  $\rho$  increases, the probability of

both firms liquidating increases while the probability of only one firm liquidating decreases. From either firm's perspective, itself being the liquidating firm when the opponent chooses to continue is definitely a better case scenario than a situation where both firms choose to liquidate. If the strategy and debt level of firm 2 remain unchanged, firm 1 has the incentive to marginally increase its own debt level to increase the probability of the situation where itself is the only firm to liquidate. The same rationale applies to firm 2. As a result, when  $\rho$  increases, the equilibrium debt levels of both firms increase, which increases the risk of industry-wide financial distress and simultaneous liquidations. The implication is alarming. It means that in those industries where peer firms' performance are more influenced by common factors (business cycle, supply shock, etc.), there is a tendency for the whole industry to be overly leveraged. Highly correlated sectors are more susceptible to the risk of industry-wide crises to begin with, which, unfortunately, is not remedied but further exacerbated by their debt choices.

## 2.5 Conclusion

This paper has presented a theoretical model to explore how peer effects on liquidation prices influence firms' debt choices. The model predicts that optimal debt levels will be lower if asset liquidation prices are expected to be more severely impacted by simultaneous peer liquidations. Somewhat counter-intuitively, higher correlation in firms' cash flows will lead to higher debt levels in equilibrium. Individual firms' attempts to maximize their own utility lead to an industry-wide over leverage, further exacerbating the risk of general crises in an already highly correlated industry. This result cautions us against the presumption that firms' capital structure choices can ameliorate industry-level risks. In certain situations, optimal choices at the individual firm level only worsen the industry-level exposure to systematic risks.

## Appendix : The Benchmark Model (No Peer Effect)

The firm's expected utility can be written as the following.

$$\begin{aligned} U(D) &= \int_0^\Lambda PdV + \int_\Lambda^D (V - \phi(V - D))dV + \int_D^1 VdV + \tau D \\ &= P\Lambda + 1/2(1 + \phi)(D^2 - \Lambda^2) - \phi D(D - \Lambda) + 1/2(1 - D^2) + \tau D \end{aligned}$$

And  $\Lambda = \frac{P+\phi D}{1+\phi}$ , Then

$$\begin{aligned} U(D) &= D^2 \frac{-\phi}{2(1+\phi)} + D\left(\tau + \frac{\phi PD}{1+\phi}\right) + const \\ D^* &= -\frac{\tau + \frac{\phi PD}{1+\phi}}{2 \frac{-\phi}{2(1+\phi)}} = P + \frac{1+\phi}{\phi} \tau \end{aligned}$$

## Appendix : Baseline Peer-Effect Model

Lemma 1 states that there exists a cutoff for  $V$  for each firm. So it must be true that when  $V = V^c$  the firm is indifferent between liquidation and continuation. That is

$$\begin{cases} \mathbb{E}[P_l \times L' + P_h \times (1 - L') | V = V_c] - [V_c - \phi(D - V_c)] = 0 \\ \mathbb{E}[P_l \times L + P_h \times (1 - L) | V' = V'_c] - [V'_c - \phi(D' - V'_c)] = 0 \end{cases}$$

Given the cutoff setup, we have

$$\begin{cases} \mathbb{E}[L' | V = V_c] = \mathbb{P}[V' < V'_c | V = V_c] = P[V' < V'_c] = V'_c \\ E[L | V' = V'_c] = \mathbb{P}[V < V_c | V' = V'_c] = P[V < V_c] = V_c \end{cases}$$

Then

$$\begin{cases} P_h - (P_h - P_l) \times V'_c - [V_c - \phi(D - V_c)] = 0 \\ P_h - (P_h - P_l) \times V_c - [V'_c - \phi(D' - V'_c)] = 0 \end{cases}$$

This is a linear equation set and after solving for  $V_c$  and  $V'_c$  we have

$$\begin{aligned} V_c &= A \cdot D + B \cdot D' + C \\ V'_c &= A \cdot D' + B \cdot D + C \\ A &= \frac{\phi \cdot (1 + \phi)}{(1 + \phi)^2 - (P_h - P_l)^2} > 0 \\ B &= \frac{-\phi \cdot (P_h - P_l)}{(1 + \phi)^2 - (P_h - P_l)^2} < 0 \\ C &= \frac{P_h}{1 + \phi + P_h - P_l} > 0 \end{aligned}$$

Then firm's expected utility becomes

$$\begin{aligned}
U(D) &= \int_0^{V_c} [P_l V'_c + P_l(1 - V'_c)]dV + \int_{V_c}^D (V - \phi(V - D))dV + \int_D^1 V dV + \tau D \\
&= [P_l V'_c + P_h(1 - V'_c)]V_c + 1/2(1 + \phi)(D^2 - V_c^2) - \phi D(D - V_c) + 1/2(1 - D^2) + \tau D \\
&= [P_h - (P_h - P_l)(AD' + BD + C)](AD + BD' + C) + 1/2(1 + \phi)(D^2 - (AD + BD' + C)^2) \\
&\quad - \phi D(D - (AD + BD' + C)) + 1/2(1 - D^2) + \tau D \\
&= D^2((P_h - P_l)AB + 1/2(1 + \phi)(1 - A^2) - \phi(1 - A) - 1/2) \\
&\quad + D(D'(-(P_h - P_l)(A^2 + B^2) - (1 + \phi)AB + \phi B) \\
&\quad + (P_h A - (P_h - P_l)(BC + AC) - (1 + \phi)AC + \phi C + \tau)) + const
\end{aligned}$$

This is quadratic w.r.t. D. So the response function can be written as

$$\begin{aligned}
D(D') &= \alpha D' + \beta \\
D'(D) &= \alpha D + \beta \\
\alpha &= -\frac{-A^2(P_h - P_l) - B^2(P_h - P_l) - (1 + \phi)A \cdot B + \phi B}{2[-A \cdot B(P_h - P_l) + \frac{1+\phi}{2}(1 - A^2) - \phi(1 - A) - \frac{1}{2}]} \\
\beta &= -\frac{-A \cdot C(P_h - P_l) + A \cdot P_l - B \cdot C(P_h - P_l) - (1 + \phi)A \cdot C + \phi C + \tau}{2[-A \cdot B(P_h - P_l) + \frac{1+\phi}{2}(1 - A^2) - \phi(1 - A) - \frac{1}{2}]}
\end{aligned}$$

So in equilibrium it must be true that  $D^* = \alpha D^* + \beta$ . Hence  $D^* = \beta/(1 - \alpha)$

## Appendix : Model with Different Firm Sizes

Lemma 1 is still valid in this case. To solve a more general case, let's assume the size of the larger firm is  $\Gamma_H$  while the size of the smaller firm is  $\Gamma_L$ . In my model  $\Gamma_H = 1$  and  $\Gamma_L = \Gamma$ .

$$\begin{cases} \mathbb{E}[P_l \times L' + P_h \times (1 - L')|V = V_c] - [V_c - \phi(D - V_c)] = 0 \\ \mathbb{E}[P_l \times L + P_h \times (1 - L)|V' = V'_c] - [V'_c - \phi(D' - V'_c)] = 0 \end{cases}$$

Given the cutoff setup, we have

$$\begin{cases} \mathbb{E}[L'|V = V_c] = \mathbb{P}[V' < V'_c|V = V_c] = P[V' < V'_c] = V'_c/\Gamma_L \\ \mathbb{E}[L|V' = V'_c] = \mathbb{P}[V < V_c|V' = V'_c] = P[V < V_c] = V_c/\Gamma_H \end{cases}$$

Then

$$\begin{cases} P_h - (P_h - P_l) \times V'_c/\Gamma_L - [V_c - \phi(D - V_c)] = 0 \\ P_h - (P_h - P_l) \times V_c/\Gamma_H - [V'_c - \phi(D' - V'_c)] = 0 \end{cases}$$

This is a linear equation set and after solving for  $V_c$  and  $V'_c$  we have

$$\begin{aligned} V_c &= A_H \cdot D + B_H \cdot D' + C_H \\ V'_c &= A_L \cdot D' + B_L \cdot D + C_L \\ A_H &= A_L = \frac{\phi \cdot (1 + \phi)}{(1 + \phi)^2 - (P_h - P_l)^2/\Gamma_H/\Gamma_L} > 0 \\ B_H &= \frac{-\phi \cdot (P_h - P_l)/\Gamma_L}{(1 + \phi)^2 - (P_h - P_l)^2/\Gamma_H/\Gamma_L} < 0 \\ B_L &= \frac{-\phi \cdot (P_h - P_l)/\Gamma_H}{(1 + \phi)^2 - (P_h - P_l)^2/\Gamma_H/\Gamma_L} < 0 \\ C_H &= \frac{P_h(1 - (P_h - P_l)/\Gamma_L/\phi)}{(1 + \phi)^2 - (P_h - P_l)^2/\Gamma_H/\Gamma_L} > 0 \\ C_L &= \frac{P_h(1 - (P_h - P_l)/\Gamma_H/\phi)}{(1 + \phi)^2 - (P_h - P_l)^2/\Gamma_H/\Gamma_L} > 0 \end{aligned}$$

Now the expected utility for firm F becomes

$$\begin{aligned}
U(D) &= \int_0^{V_c} [P_l V_c' / \Gamma_L + P_l (1 - V_c' / \Gamma_L)] \frac{dV}{\Gamma_H} + \int_{V_c}^D (V - \phi(V - D)) \frac{dV}{\Gamma_H} + \int_D^{\Gamma_H} V \frac{dV}{\Gamma_H} + \tau D \\
&= [P_l V_c' / \Gamma_L + P_h (1 - V_c' / \Gamma_L)] V_c / \Gamma_H + 1/2(1 + \phi)(D^2 - V_c^2) / \Gamma_H \\
&\quad - \phi D(D - V_c) / \Gamma_H + 1/2(\Gamma_H^2 - D^2) / \Gamma_H + \tau D \\
&= [P_h - (P_h - P_l)(A_L D' + B_L D + C_L) / \Gamma_L] (A_H D + B_H D' + C_H) / \Gamma_H \\
&\quad + 1/2(1 + \phi)(D^2 - (A_H D + B_H D' + C_H)^2) / \Gamma_H \\
&\quad - \phi D(D - (A_H D + B_H D' + C_H)) / \Gamma_H + 1/2(\Gamma_H^2 - D^2) / \Gamma_H + \tau D \\
&= D^2 \left( \frac{(P_h - P_l) A_H B_L}{\Gamma_H \Gamma_L} + 1/2(1 + \phi)(1 - A_H^2) / \Gamma_H - \phi(1 - A_H) / \Gamma_H - 1/2 / \Gamma_H \right) \\
&\quad + D(D' \left( \frac{-(P_h - P_l)(A_H A_L + B_H B_L)}{\Gamma_H \Gamma_L} - (1 + \phi) A_H B_H / \Gamma_H + \phi B_H / \Gamma_H \right) \\
&\quad + (P_h A_H / \Gamma_H - (P_h - P_l) \frac{(B_L C_H + A_H C_L)}{\Gamma_H \Gamma_L} - (1 + \phi) A_H C_H / \Gamma_H + \phi C_H / \Gamma_H + \tau)) \\
&\quad + \text{const}
\end{aligned}$$

Similarly the expected utility for firm F' is

$$\begin{aligned}
U'(D') &= \int_0^{V_c'} [P_l V_c / \Gamma_H + P_l (1 - V_c / \Gamma_H)] \frac{dV}{\Gamma_L} + \int_{V_c'}^{D'} (V - \phi(V - D')) \frac{dV}{\Gamma_L} + \int_{D'}^{\Gamma_L} V \frac{dV}{\Gamma_L} + \tau D' \\
&= [P_l V_c / \Gamma_H + P_h (1 - V_c / \Gamma_H)] V_c' / \Gamma_L + 1/2(1 + \phi)(D'^2 - V_c'^2) / \Gamma_L \\
&\quad - \phi D'(D' - V_c') / \Gamma_L + 1/2(\Gamma_L^2 - D'^2) / \Gamma_L + \tau D' \\
&= [P_h - (P_h - P_l)(A_H D + B_H D' + C_H) / \Gamma_H] (A_L D' + B_L D + C_L) / \Gamma_L \\
&\quad + 1/2(1 + \phi)(D'^2 - (A_L D' + B_L D + C_L)^2) / \Gamma_L \\
&\quad - \phi D'(D' - (A_L D' + B_L D + C_L)) / \Gamma_L + 1/2(\Gamma_L^2 - D'^2) / \Gamma_L + \tau D' \\
&= D'^2 \left( \frac{(P_h - P_l) A_L B_H}{\Gamma_H \Gamma_L} + 1/2(1 + \phi)(1 - A_L^2) / \Gamma_L - \phi(1 - A_L) / \Gamma_L - 1/2 / \Gamma_L \right) \\
&\quad + D'(D' \left( \frac{-(P_h - P_l)(A_H A_L + B_H B_L)}{\Gamma_H \Gamma_L} - (1 + \phi) A_L B_L / \Gamma_L + \phi B_L / \Gamma_L \right) \\
&\quad + (P_h A_L / \Gamma_L - (P_h - P_l) \frac{(B_H C_L + A_L C_H)}{\Gamma_H \Gamma_L} - (1 + \phi) A_L C_L / \Gamma_L + \phi C_L / \Gamma_L + \tau)) \\
&\quad + \text{const}
\end{aligned}$$



Hence

$$D(D') = \alpha_H D' + \beta_H$$

$$D'(D) = \alpha_L D + \beta_L$$

$$\alpha_H = -\frac{\frac{-(P_h-P_l)(A_H A_L+B_H B_L)}{\Gamma_H \Gamma_L} - (1+\phi)A_H B_H/\Gamma_H + \phi B_H/\Gamma_H}{2\left[\frac{(P_h-P_l)A_H B_L}{\Gamma_H \Gamma_L} + 1/2(1+\phi)(1-A_H^2)/\Gamma_H - \phi(1-A_H)/\Gamma_H - 1/2/\Gamma_H\right]}$$

$$\alpha_L = -\frac{\frac{-(P_h-P_l)(A_H A_L+B_H B_L)}{\Gamma_H \Gamma_L} - (1+\phi)A_L B_L/\Gamma_L + \phi B_L/\Gamma_L}{2\left[\frac{(P_h-P_l)A_L B_H}{\Gamma_H \Gamma_L} + 1/2(1+\phi)(1-A_L^2)/\Gamma_L - \phi(1-A_L)/\Gamma_L - 1/2/\Gamma_L\right]}$$

$$\beta_H = -\frac{P_h A_H/\Gamma_H - (P_h - P_l)\frac{(B_L C_H + A_H C_L)}{\Gamma_H \Gamma_L} - (1+\phi)A_H C_H/\Gamma_H + \phi C_H/\Gamma_H + \tau}{2\left[\frac{(P_h-P_l)A_H B_L}{\Gamma_H \Gamma_L} + 1/2(1+\phi)(1-A_H^2)/\Gamma_H - \phi(1-A_H)/\Gamma_H - 1/2/\Gamma_H\right]}$$

$$\beta_L = -\frac{P_h A_L/\Gamma_L - (P_h - P_l)\frac{(B_H C_L + A_L C_H)}{\Gamma_H \Gamma_L} - (1+\phi)A_L C_L/\Gamma_L + \phi C_L/\Gamma_L + \tau}{2\left[\frac{(P_h-P_l)A_L B_H}{\Gamma_H \Gamma_L} + 1/2(1+\phi)(1-A_L^2)/\Gamma_L - \phi(1-A_L)/\Gamma_L - 1/2/\Gamma_L\right]}$$

After solving for  $D = D(D'(D))$ , the result is

$$D^* = \frac{\alpha_H \beta_L + \beta_H}{1 - \alpha_H \alpha_L}$$

$$D'^* = \frac{\alpha_L \beta_H + \beta_L}{1 - \alpha_H \alpha_L}$$

## CHAPTER 3

# Applying Deep Learning to Momentum Trading Strategies

The financial literature on the cross-section prediction of stock returns is voluminous. Price momentum and reversals have been documented for various time horizons. Bondt and Thaler 1985 document long-term reversals over a period of 3 to 5 years while Lehmann 1990 and Jegadeesh 1990 document reversals for horizons of one-week and one-month, respectively. Over the horizons of 3 to 12 months, return continuation (momentum) is observed (Jegadeesh and Titman 1993, 2001). Examining returns in weekly frequencies, Gutierrez Jr and Kelley 2008 also find reversals in short horizons and momentum in the 3-to-12 month horizons, which have gradually become a widely accepted wisdom.

The cross-section return predictability exists between and within industries (Moskowitz and Grinblatt 1999; Hameed and Mian 2015) and, beyond the US market, appears to be prevalent in many other markets to varying degrees (Rouwenhorst 1998; Muga and Santamara 2007; Naughton, Truong, and Veeraraghavan 2008; Hhn and Scholz 2019; Tang and Zhang 2014). Prevalence of momentum or reversal is found to be associated with many factors, including firm characteristics such as market cap and volatility (Wei and Yang 2012), past trading volume (Connolly and Stivers 2003), and liquidity provision by active institutional investors (Cheng et al. 2017).

The existence of momentum and reversal effects provides a possible ground for constructing profitable trading strategies. In reality, however, it turns out to be challenging to transfer research outputs into real-world investment trading, as the findings of existing studies tend to be contingent on specific market periods and holding horizons. The momentum/reversal

trading strategies are largely the products of a vast effort by finance academics and practitioners to hand-engineer features from historical stock prices.

With the recent development in deep learning, historical financial prices can be a fertile ground to test if machine learning algorithms can generate more reliable trading strategies than the traditional statistical methods used in financial research and practice. In general, deep learning refers to a rich set of neural network models that show promising results in improving the prediction accuracy of regression and classification problems in many different areas of scientific research such as image processing, speech recognition, etc (Goodfellow, Bengio, and Courville 2016).

The idea of applying deep learning models to financial predictions is not entirely new. Trippi and DeSieno 1992 have experimented with trading equity index futures with a neural network. Leung, Daouk, and Chen 2000 also used neural networks as part of a forecasting model for stock indices. These early works are prone to the problem of over-fitting and as a result haven't generated much impact in the financial literature. More recently, Takeuchi 2013 have analyzed cross-sectional momentum strategies using a restricted Boltzmann machine, which shows performance improvements compared to classical counterparts. Dixon, Klabjan, and Bang 2017 have applied deep learning to predict the returns of currency and commodity futures based on high-frequency data. Messmer 2017 has trained deep Deep Feedforward Networks with firm characteristics to predict the US cross-section of stock returns and found that the model-generated long-short portfolios can outperform a linear benchmark. Li and Tam 2018 have experimented with various machine learning techniques to predict the momentum and reversal effects in the stock market of mainland China and highlighted the promising use of machine learning approaches to build real-world profitable trading strategies.

This study aims to further explore how recent innovations in deep learning can enhance momentum- and reversal-based trading strategies. I have designed a Deep Feedforward Network (DFN) to look at prediction changes purely from variations in the input data of historical prices. To capture the momentum and reversal effects in all the possible short- and mid-horizons documented in the literature, for any stock  $i$ 's performance in month  $m$ ,

twenty z-scored returns accumulated over horizons of 1 to 10 days and 1 to 10 months are fed to the deep learning machine as predictive variables. The model is trained to generate a predictive signal for each stock that ranges from 0 to 1, where a higher value indicating a higher probability that the said stock will outperform the market median in the next period.

Once the model has finished learning from the training data set, which covers the period from 1965-2000, it is tested on the validation set (2000-2016). Based on the ranking of the model-generated predictive signals, I then construct a long-short portfolio with the top and bottom 20, 30, or 40 percentiles of the stocks. These portfolios generated by the deep learning machine significantly outperform traditional momentum strategies by both mean returns and Sharpe ratios. Over the course of the 15-year validation period, the long-short portfolio using the top and bottom quintiles would have performed the best, seeing a total value increase of 50 times, in comparison to a 2.5 times increase of the S&P500 portfolio over the same period.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 sets up and trains the Deep Feedforward Neural Network Model. In Section 4, I report the results of applying the trained model to the validation set and try to explore the factors behind the better performance of deep-learning-generated portfolios over traditional trading strategies. Section 5 concludes.

### **3.1 Data Description**

The stock price data is obtained from The Center for Research in Security Price (CRSP). I restrict my analysis to US shares trading on NYSE, AMEX, and Nasdaq (with exchange codes of 1, 2 or 3). PERMNO is used as the stock identifier. Stock prices are adjusted by dividing a cumulative price adjusting factor to account for stock split or dividend. Data for the Fama-French five factors, momentum factor, and short-term reversal factor are downloaded from Kenneth French's website. The price data is split into two sets by date. Data from the beginning of 1965 to the end of 2000 is used as the training set, and data from the beginning of 2000 to the end of 2016 is used for validation.

To construct the training data set, for any stock  $i$ 's return in month  $m$ , the dependent variable is defined as a binary that takes the value of 1 if the stock's return is higher than the cross-sectional median and 0 otherwise. On the predictive side, twenty z-scores are constructed cross-sectionally from the preceding window of 12 months:

- First, with the starting moment of the window fixed at one year before the current trading month  $m$ , compute the z-scored return accumulated over a horizon of 1 to 10 months. This generates 10 scores, respectively for a 1-month horizon ( $m - 12$ ), a 2-month horizon ( $m - 12$  to  $m - 11$ ), a 3-month horizon ( $m - 12$  to  $m - 10$ ), and so forth, until a 10-month horizon ( $m - 12$  to  $m - 2$ ).
- Second, from the first day  $d$  in month  $m - 1$ , compute the z-scored return accumulated over a horizon of 1 to 10 days. This generates another 10 scores, respectively for a 1-day horizon ( $d$ ), a 2-day horizon ( $d$  to  $d + 1$ ), a 3-day horizon ( $d$  to  $d + 3$ ), and so forth, until a 10-day horizon ( $d$  to  $d + 10$ ). So the 'cooling period' of this strategy is around 20 days, from  $d + 11$  to the end of month  $m - 1$ .

To deal with missing and invalid data, entries with no value for the dependent variable or more than four missing values for independent variables are dropped. The remaining missing points in the independent variables are back-filled with the mean value of all the valid independent variables within the same window. The window then rolls forward by a month and the process above is repeated. The entries from different windows are vertically stacked to form the training data set, which includes in total 1,783,402 entries.

For the validation data set, the predictive variables for each stock  $i$  in month  $m$  are constructed in the exact same way. There are 189 months in the validation data set. The predictive variables are sent into the model, which returns a predictive signal between 0 and 1, indicating the machined-predicted likelihood for stock  $i$  to outperform the cross-sectional median in month  $m$ . The outcome in reality, i.e. stock  $i$ 's actual return in month  $m$ , is used to evaluate the hypothetical performance of the investment portfolios that are constructed according to the machine-generated signals.

## 3.2 Model

### 3.2.1 Deep Feedforward Network

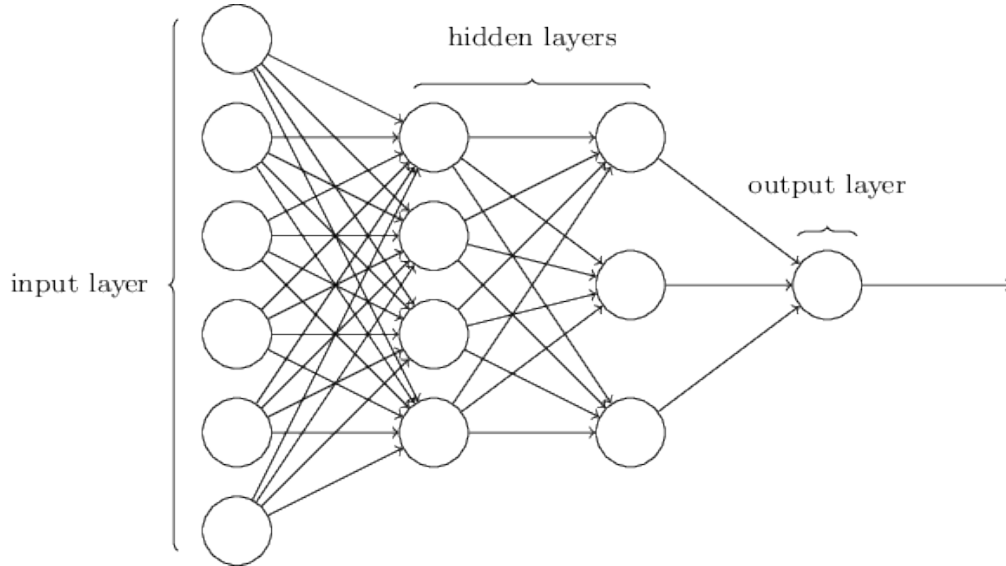


Figure 3.1: Example of a Simple Four-Layer (6-4-4-1) DFN

For the architecture of the deep learning model, I have chosen the Deep Feedforward Networks (DFNs), also known as feedforward neural networks or multilayer perceptrons (MLPs). The basic mechanism of a DFN model works as follows. The goal of a DFN is to approximate some function  $y = G^*(\mathbf{x})$ . In the case of this study,  $G^*(\mathbf{x})$  is a classifier function that maps a stock's historical performance  $\mathbf{x}$  to a binary output  $y$  of whether its return in the current month is above the cross-sectional median. The DFN defines a mapping  $y = G(\mathbf{x}, \boldsymbol{\theta})$  and learns the value of the model parameters  $\boldsymbol{\theta}$  that result in the best approximation.

To do the approximation, the model consists of multiple layers. The input data enters into the input layer. The number of nodes in the input layer equals to the dimension of the independent variable; in our case, that is 20. Then they go through a series of hidden layers that connect various inputs in a nested and hidden structure. The output layer reflects the prediction target. Figure 3.1 illustrates a simple four-layer (6-4-4-1) DFN. The model architecture used in this study is a five-layer (20-50-100-20-2) fully connected DFN, which is shown in Fig 3.2, where the input layer (with 20 nodes) is at the bottom, the output layer

(with 2 nodes) at the top, and three hidden layers in between.

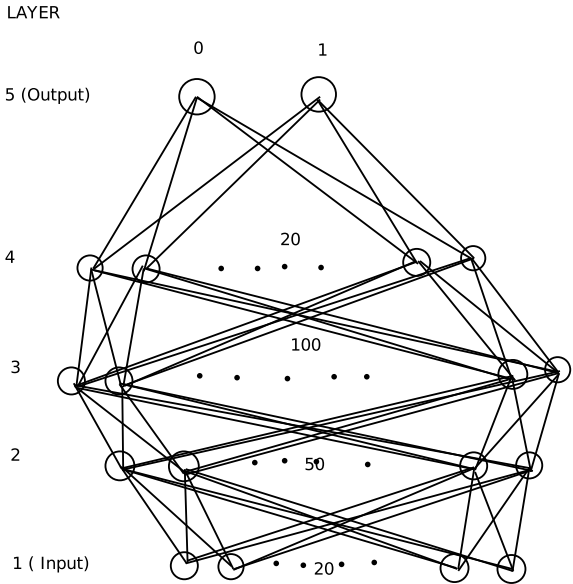


Figure 3.2: Five-Layer (20-50-100-20-2) DFN Model Specification.

Between two adjacent layers, any node  $i$  in the upper layer is fully connected with any node  $j$  in the lower layer by an edge with weight  $w_{ij}$ . Node  $i$  in the upper layer has a value of  $v_i = f(\sum_j u_j w_{ij} + b_i)$ , where  $u_j$  is the value of the  $j$ th node in the lower layer, and  $b_i$  is a node-specific bias term. The  $f$  function is called an activation function. Usually it's "S" shaped. Examples of activation functions are logistic function  $f(z) = \frac{1}{1+e^{-z}}$ , arctangent function  $f(z) = \arctan(z)$ , step function  $f(z) = \text{sgn}(z)$ , etc. In our five-layer model, the activation function for layers 2, 3, 4 is a rectifier function:

$$f(z) = z^+ = \max(0, z) \tag{3.1}$$

For a classification problem, the top layer contains as many nodes as the number of classes. There are two classes for the model used here: stock performance either higher or lower than the median. Hence there are two nodes in the top layer. The value of each node indicates the probability that the input data ends up in that node. So we need a way to make sure that the value of each node is between 0 and 1 and their values add up to 1.

Let  $y_i$  be the the value of node  $i$  in the output layer ( $i = 0, 1$ ). Therefore,  $y_1$  represents

the predicted probability that the stock will outperform the median, while  $y_0$  is the predicted probability that it will perform below the median.  $y_i$  is calculated by a softmax classifier as follows:

$$y_i = \sigma_i(z_0, z_1) = \frac{e^{z_i}}{e^{z_0} + e^{z_1}} \quad \text{for } i = 0, 1 \quad (3.2)$$

in which

$$z_i = \sum_j u_j w_{ij} + b_i \quad \text{for } i = 0, 1 \quad (3.3)$$

where  $u_j$  is the value of node  $j$  in the layer below (i.e. layer 4), and  $b_i$  is the bias term specific to node  $i$ .

One can easily verify that  $y_0$  and  $y_1$  are between 0 and 1 and add up to 1. With the two being equivalent,  $y_1$  is the only output variable that we need to look at. The softmax classifier has the additional advantage that it works well with the cross-entropy loss function, which will be detailed in the next part.

The model output that we are interested in, which is the predicted probability of the a stock outperforming the cross-sectional median, can be expressed as

$$y = G(\mathbf{x}, \boldsymbol{\theta}) = g^{(4)}(g^{(3)}(g^{(2)}(g^{(1)}(\mathbf{x})))) \quad (3.4)$$

Where  $g^{(n)}$  is the transformation between the layer  $n - 1$  and layer  $n$ . With the training data set, the deep learning machine then tries to generate the model parameters  $\boldsymbol{\theta}$ , which includes all the edge weights and node-specific bias terms, to best predict the outcome.

### 3.2.2 Cost Function

To train the model, an objective function, also called a cost function, needs to be defined. Here I use the loss function of maximum likelihood estimation (MLE). In the case of classification, the loss function of MLE,  $\sum \log(\mathbb{P}[D|X; \theta])$ , becomes



$$C = -\frac{1}{N} \sum_{d \in D} (y_d^* \ln(y_d) + (1 - y_d^*) \ln(1 - y_d)) \quad (3.5)$$

where  $D$  is the whole training dataset,  $d$  is an element in the set,  $N$  is the number of elements in the dataset,  $y_d^*$  is the actual binary outcome for  $d$ , and  $y_d$  is the predicted probabilistic outcome.

The training process is to find the edge weights and node-specific biases that minimize the cross-entropy given a certain dataset. Since the model is very complicated and non-linear, it's hard to acquire analytical solutions to these parameters like in a linear regression. Luckily, the gradient of the loss function can be directly acquired, given each node's value, by back-propagation, which is an application of the chain-rule when calculating derivatives.

### 3.2.3 Gradient Descent, Stochastic Gradient Descent, and Mini-Batching

Gradient descent is a iterative tool for finding the minimum of a function. Each iteration the optimization is done by moving along the gradient by one step. Fig 3.3 shows the basic idea of gradient descent. The circled blue lines are the contour plot of the loss function. Gradients are perpendicular to the blue lines. To find the local minimum, we start at point  $x_0$ , and calculate the gradient of the loss function at  $x_0$ . Then move along the gradient for one step from  $x_0$  to  $x_1$ . Then at  $x_1$ , calculate the local gradient, move along it to  $x_2$ , and so on. Eventually we should reach to a point where it's close to the local minimum.

The downside of a “full” gradient descent in machine learning is that for each iteration, the gradient has to be recalculated, by using the full dataset. So when the dataset is considerably large, this method becomes very slow. The solution to this problem is to split data into plenty of “mini-batches.” In each iteration, the optimizer uses one batch to calculate the gradient and move a step, which hugely speeds up the training process. I reshuffle the whole dataset and re-batch for certain iterations. In my code, I used a learning rate of 0.005, reshuffled 150 times, and the batch size is 1000.

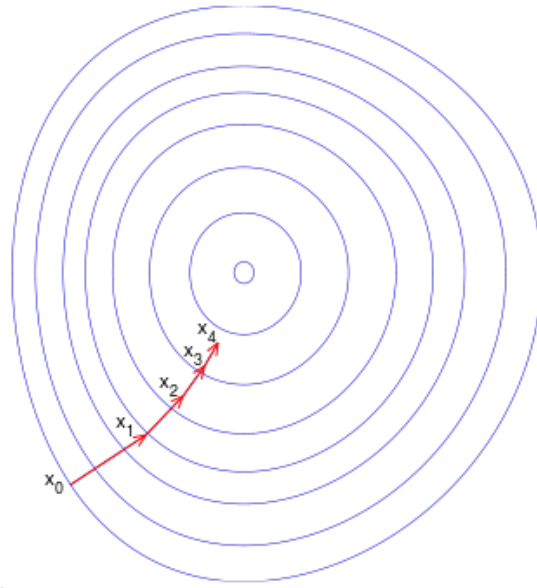


Figure 3.3: Illustration of Gradient Descent

### 3.2.4 Validation

Once the training is finished and the model parameters are generated, I apply the model to the validation data set, which includes the twenty input variables constructed from the period 2000 to 2016. The model's outputs are monthly predicted signals for each stock indicating its probability of performing above the cross-sectional median from March 2001 onward. Based on the predicted signals, I construct an investment portfolio by longing stocks with signals in the top  $x$  percentile and shorting stocks with signals in the bottom  $x$  percentile. The portfolio is re-balanced every month. The stocks are equal-weighted. The monthly return of the long-short portfolio from March 2001 to November 2016 (189 months in total) can be computed with the actual stock returns in the validation period.

Table 3.1: Annualized Returns of Different Strategies

| Portfolio | Mean  | Stdev | Sharpe Ratio |
|-----------|-------|-------|--------------|
| ML_40     | 0.17  | 0.135 | 1.53         |
| ML_30     | 0.206 | 0.158 | 1.31         |
| ML_20     | 0.267 | 0.175 | 1.25         |
| MOM_40    | 0.109 | 0.135 | 0.8          |
| MOM_30    | 0.131 | 0.158 | 0.83         |
| MOM_20    | 0.167 | 0.186 | 0.9          |

### 3.3 Results

#### 3.3.1 Portfolio Returns

Table 3.1 shows the portfolio returns throughout the span of 189 months. Returns are annualized as well as the Sharpe Ratio. ML\_40, ML\_30, and ML\_20 represent the long-short portfolios created by the deep learning signals with the top and bottom 40, 30, and 20 percentiles of the stocks, respectively. In addition, I have also created benchmark portfolios for comparison using a simple momentum strategy. MOM\_40, MOM\_30, and MOM\_20 represent, respectively, the long-short portfolios that include the top and bottom 40, 30, and 20 percentiles of stocks based on their returns in the preceding 12 months.

Overall, we can see that the portfolios generated by the deep learning machine outperform the simple momentum strategies by both mean returns and Sharpe ratios. ML\_20, ML\_30, ML\_40 achieve annualized returns of 0.267, 0.206, and 0.17 respectively. The back-tested Sharpe ratios are 1.25, 1.31, and 1.53 respectively. Since the signals are the same across different portfolios in the ML category, the difference is only the number of stocks they pick up. ML\_20 takes in a smaller number of stocks, therefore it has a higher return but also suffers from higher volatility. As a result, the Sharpe ratio is lower for the ML\_20 strategy.

Figure 3.4 shows the progression of portfolio values over time. “sprtrn” is a portfolio invested in the S&P 500 index. All portfolio values start at 1. Over the course of fifteen years, the ML\_20 portfolio performed the best. ML\_20, ML\_30, and ML\_40 would have grown by about 50 times, 20 times and 10 times respectively, in comparison to the S&P 500 portfolio that has increased by 2.5 times over the same period.

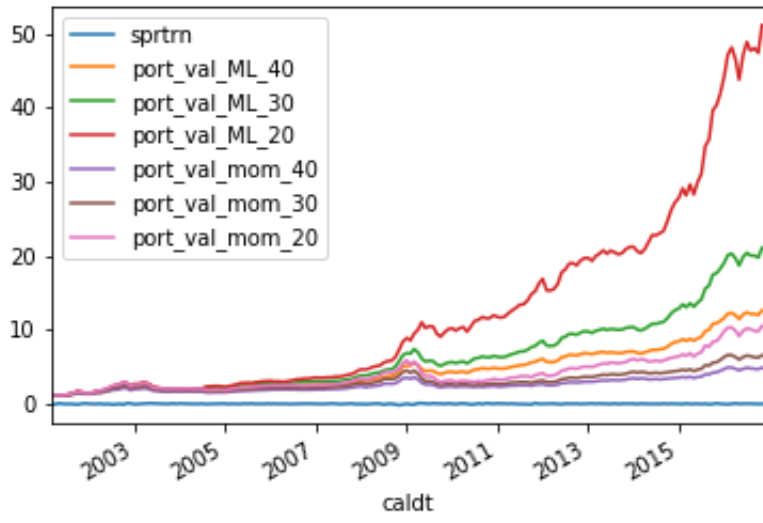


Figure 3.4: Portfolio Value of Different Strategies (2001-2016)

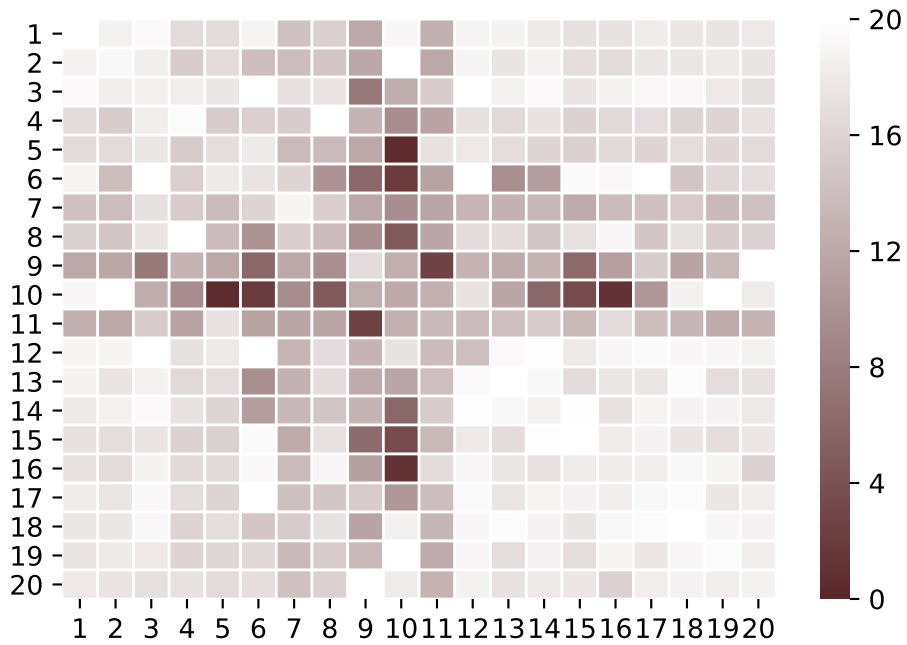


Figure 3.5: Relative Importance of the Twenty Input Scores

Which of the twenty input scores fed into the deep learning machine turn out to be most relevant for the prediction? Figure 3.5 shows the relative importance of the twenty input scores. Recall that input variables 1 to 10 represent the z-scored returns accumulated over 1 to 10 months starting from  $m - 12$ ; input variables 11 to 20 represent the z-scored returns accumulated over 1 to 10 days starting from the first day of the previous month. The color of each square  $(i, j)$  in Figure 3.5 represents the portfolio return of ML\_30 when input variables  $i$  and  $j$  are masked. When  $i = j$ , it represents the portfolio return when only input variable  $i$  is masked. The figure is symmetrical along the negative-sloping diagonal line. The benchmark return of ML\_30, when no input is masked, is 20.6%. The darker the square color, the lower the portfolio return.

Input variable 10, which is the 10-month accumulated return from  $m - 12$  to  $m - 2$ , seems to exert the greatest impact on the model prediction when it is masked together some other input variables. The highest impact comes from the simultaneous masking of input 10 and input 5, which is the 5-month accumulated return from  $m - 12$  to  $m - 7$ , and the simultaneous masking of input 10 and input 16, which is the 6-day accumulated return starting from the first day of the previous month.

### 3.3.2 Factor Analysis

To better understand the characteristics of the composition of the machine-generated portfolios, I run the portfolio returns against known factors that explain variations in stock/portfolio prices.

Table 3.2 shows the results of regressing the monthly portfolio returns of ML\_20, ML\_30, and ML\_40 on the Fama-French three factors plus a momentum or reversal factor (Fama and French 1993). Returns are monthly and not annualized.  $R_m - R_f$  is the excess return of the value-weight market portfolio over the risk-free return.  $SMB$  is the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks.  $HML$  is the difference between the returns on diversified portfolios of high and low book-to-market ratio stocks.  $Mom$ , the momentum factor, is the difference between returns on

Table 3.2: Portfolio Returns vs. FF Three Factors, Momentum, and Short-Term Reversal

|       | Intercept               | Rm-Rf                   | SMB                     | HML                    | Mom                    | ST_Rev                  | $R^2$ |
|-------|-------------------------|-------------------------|-------------------------|------------------------|------------------------|-------------------------|-------|
| ML_40 | 0.015***<br>(1.62e-18)  | -0.348***<br>(1.35e-14) | -0.324***<br>(7.11e-07) | 0.195***<br>(1.31e-03) | 0.422***<br>(3.74e-25) |                         | 0.73  |
| ML_30 | 0.0183***<br>(5.77e-19) | -0.399***<br>(1.08e-13) | -0.36***<br>(3.76e-06)  | 0.233***<br>(1.37e-03) | 0.49***<br>(3.84e-24)  |                         | 0.71  |
| ML_20 | 0.0235***<br>(3.01e-15) | -0.467***<br>(1.87e-09) | -0.335***<br>(3.26e-03) | 0.356***<br>(1.01e-03) | 0.349***<br>(7.14e-08) |                         | 0.478 |
| ML_40 | 0.0168***<br>(2.54e-14) | -0.549***<br>(3.11e-20) | -0.27***<br>(1.59e-03)  | 0.103<br>(0.193)       |                        | -0.099<br>(0.136)       | 0.513 |
| ML_30 | 0.0204***<br>(7.39e-15) | -0.63***<br>(1.7e-19)   | -0.297***<br>(3.24e-03) | 0.126<br>(0.18)        |                        | -0.117<br>(0.138)       | 0.498 |
| ML_20 | 0.0251***<br>(1.38e-15) | -0.577***<br>(6.11e-13) | -0.269**<br>(2.49e-02)  | 0.269**<br>(1.72e-02)  |                        | -0.281***<br>(3.16e-03) |       |

diversified portfolios of high and low accumulated returns over ten months from  $m - 12$  to  $m - 2$ .  $ST\_Rev$ , the short-term reversal factor, is the difference between returns on diversified portfolios of low and high accumulated returns over the prior month  $m - 1$ .

The coefficients on the momentum factor ( $Mom$ ) are of the highest scale and strongly significant, making it the top explanatory factor for the deep learning strategy. The momentum factor is also much more significant than the short-term reversal factor, indicating the neural network has picked up a strategy that is more momentum-reliant, which is also confirmed by the larger  $R^2$  for  $Mom$ . It also makes sense that ML\_30 has the highest momentum coefficient since the break points in constructing the momentum factor are the 30th and 70th percentile. The market ( $R_m - R_f$ ) and size ( $SMB$ ) factors carry significant negative coefficients. Book-to-market ratio ( $HML$ ) has positive coefficients, especially significant when the momentum factor is added instead of the short-term reversal factor.

Taken together, it means that the deep learning strategies, to varying degrees, are in effect longing the high-momentum portfolio, the big-cap portfolio, and the high book-to-market portfolio, while shorting the market portfolio. Apparently, these factors do not capture all variation in expected returns for the deep learning portfolios, all three of which have positive and significant alphas. ML\_40, ML\_30, and ML\_20 generate monthly alphas of 1.5, 1.8, and 2.3 percent points respectively, which are all significant at the 0.001 level.

Table 3.3: Portfolio Returns vs. Fama-French Five-Factor Model

| Item               | Intercept               | Rm-Rf                   | SMB                    | HML               | RMW                     | CWA                  |
|--------------------|-------------------------|-------------------------|------------------------|-------------------|-------------------------|----------------------|
| ML_40              | 0.012***<br>(1.79e-10)  | -0.33***<br>(6.26e-10)  | -0.156**<br>(3.34e-02) | -0.068<br>(0.389) | 0.839***<br>(2.07e-15)  | 0.174<br>(0.134)     |
| ML_30              | 0.015***<br>(4.59e-11)  | -0.378***<br>(2.16e-09) | -0.162*<br>(6.22e-02)  | -0.066<br>(0.481) | 0.986***<br>(3.427e-15) | 0.175<br>(0.204)     |
| ML_20<br>(p-value) | 0.0189***<br>(1.06e-11) | -0.318***<br>(2.07e-05) | -0.109<br>(0.299)      | 0.111<br>(0.326)  | 1.2***<br>(2.14e-15)    | 2.179e-04<br>(0.999) |

Alternatively, Table 3.3 shows the result of using the Fama-French Five-Factor model to explain the monthly portfolio returns (Fama and French 2015). In addition to the original three factors of  $R_m - R_f$ ,  $SMB$ , and  $HML$ , the five-factor model adds two more factors to capture operational profitability and investment.  $RMW$  is the difference between the returns on diversified portfolios of firms with robust and weak operational profitability.  $CWA$  is the difference between the returns on diversified portfolios of firms with low and high investment.

The Five-Factor Model has better prediction powers than the three factors plus momentum. The intercepts in Table 3.3 are smaller than those in Table 3.2, but still very significant statistically and economically. The coefficients shown in Table 3.3 suggest that the most prominent factor captured by the deep learning strategies is  $RMW$ . Operating profitability is measured with accounting data as revenue minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense on all divided by book equity. The coefficient for  $RMW$  is close to 1 for the ML\_30 and varies from 0.84 for ML\_40 to 1.2 for ML\_20, all at extremely high levels of significance.

To test if operating profitability is indeed a differentiating factor in the deep learning portfolios, we can compare the statistics of the longed stocks and the shorted stocks. Figure 3.6 shows the trend of the average operating profitability of the stocks that are longed in ML\_30 vs. those that are shorted. The t-test confirms that the difference is extremely significant.

It seems that the deep learning portfolios, which are constructed purely based on inputs of momentum/reversal information, resemble to a great extent portfolios that long firms with robust profitability and short those with weak profitability. It suggests important connections

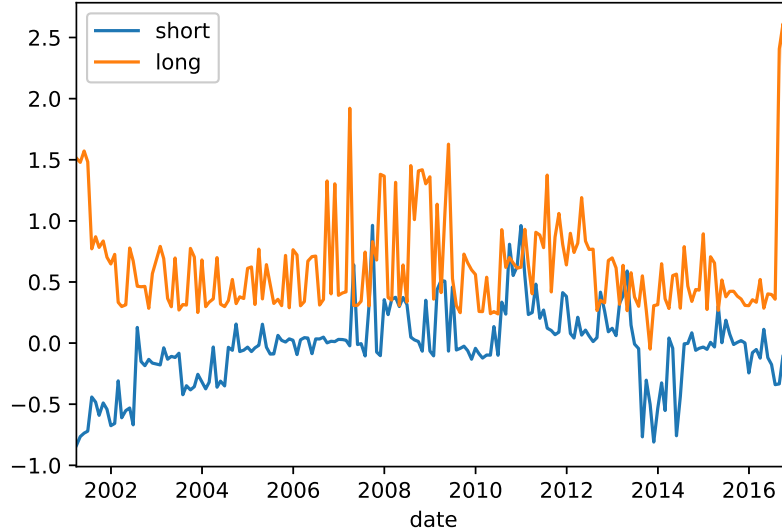


Figure 3.6: Average Operating Profitability in Long and Short Portfolios

between the predictability of stock prices and firm profitability. Indeed, the *RMW* factor has a correlation coefficient of 0.46 with the momentum factor. A recent article by Liang, Tang, and Xu 2019 also takes notice of the correlation between momentum and profitability factors and points to uncertainty as a possible common source. More research needs to be done in the future to fully explore the mechanism of such connections.

In addition to operating profitability, market ( $R_m - R_f$ ) and size (*SMB*) still carry significant negative coefficients, same as in Table 3.2. The conservativeness/aggressiveness of firm investment (*CWA*) is not a significant factor, nor is book-to-market ratio (*HML*) anymore. The insignificance of *HML* is in line with Fama and French’s observation that the addition of *RMW* and *CWA* to the factor model tends to absorb the explanatory power of *HML*, making it a redundant factor.

### 3.4 Conclusion

This study has explored the potential of using deep learning models to predict equity prices and generate profitable trading portfolios. I have designed a Deep Feedforward Network (DFN) to predict a stock’s likelihood of performing over the cross-sectional median purely



from variations in the input data of historical prices. Rolling Z-scored cumulative returns of various horizons are fed to the model as predicative variables. With the model parameters generated by the training dataset, the model is applied to the validation set to generate a monthly predicative signal for each stock. Based on this model-predicted signal, I have constructed long-short investment portfolios out of the validation dataset.

The deep learning portfolios yield an annualized return of over 20 percent and outperform simple momentum strategies. The most important input variable for the deep learning model seems to be the 10-month accumulative return from  $m-12$  to  $m-2$ , especially in combination with the 5-month accumulative return from  $m-12$  to  $m-7$  or the 6-day accumulative return starting from the first day of the previous month.

When regressed on commonly used factor models, the characteristics of the deep learning portfolios can be approached. Operating profitability seems to be the most prominent factor underlying the machine-generated strategies. It seems that the deep learning portfolios, which are constructed purely based on inputs of historical price information, resemble to a great extent portfolios that long firms with robust profitability and short those with weak profitability. Future research can further explore the connection between price predictability and profitability factors. Of lesser importance than operating profitability are the factors of market, size, momentum, and, to an even smaller degree, book-to-market ratio. The deep learning strategies, to varying degrees, are in effect longing the high-momentum portfolio, the big-cap portfolio, and the high book-to-market portfolio, while shorting the market portfolio. Still, even after accounting for these common factors, the deep learning portfolios still generate significantly large alphas. In general, the paper concludes with the great potential of using machine learning methods for portfolio management and asset pricing research.

## CHAPTER 4

### Bibliography

- Ahern, Kenneth R., Daniele Daminelli, and Cesare Fracassi. 2015. “Lost in translation? The effect of cultural values on mergers around the world”. *Journal of Financial Economics* 117, no. 1 (): 165–189. ISSN: 0304-405X. doi:10.1016/J.JFINECO.2012.08.006. <https://www.sciencedirect.com/science/article/pii/S0304405X12001602>!
- Aobdia, Daniel, Anup Srivastava, and Erqiu Wang. 2018. “Are Immigrants Complements or Substitutes? Evidence from the Audit Industry”. *Management Science* 64, no. 5 (): 1997–2012. ISSN: 0025-1909. doi:10.1287/mnsc.2016.2707. <http://pubsonline.informs.org/doi/10.1287/mnsc.2016.2707>.
- Bae, Kee-Hong, René M. Stulz, and Hongping Tan. 2008. “Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts”. *Journal of Financial Economics* 88, no. 3 (): 581–606. ISSN: 0304-405X. doi:10.1016/J.JFINECO.2007.02.004. <https://www.sciencedirect.com/science/article/pii/S0304405X07001973>.
- Banerjee, Abhijit V. 1992. “A Simple Model of Herd Behavior\*”. *The Quarterly Journal of Economics* 107, no. 3 (): 797–817. ISSN: 0033-5533. doi:10.2307/2118364. eprint: <http://oup.prod.sis.lan/qje/article-pdf/107/3/797/5298496/107-3-797.pdf>. <https://doi.org/10.2307/2118364>.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1998. “Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades”. *The Journal of Economic Perspectives* 12 (3): 151–170. ISSN: 08953309. <http://www.jstor.org/stable/2647037>.

- Bolton, Patrick, and David S Scharfstein. 1990. “A Theory of Predation Based on Agency Problems in Financial Contracting”. *American Economic Review* 80, no. 1 (): 93–106. <https://ideas.repec.org/a/aea/aecrev/v80y1990i1p93-106.html>.
- Bondt, Werner F. M. De, and Richard Thaler. 1985. “Does the Stock Market Overreact?” *The Journal of Finance* 40 (3): 793–805. ISSN: 00221082, 15406261. <http://www.jstor.org/stable/2327804>.
- Cheng, Si, et al. 2017. “Short-Term Reversals: The Effects of Past Returns and Institutional Exits”. *Journal of Financial and Quantitative Analysis* 52 (1): 143173. doi:10.1017/S0022109016000958.
- Chevalier, Judith A., and David S. Scharfstein. 1996. “Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence”. *The American Economic Review* 86 (4): 703–725. ISSN: 00028282. <http://www.jstor.org/stable/2118301>.
- Conlisk, John. 1980. “Costly optimizers versus cheap imitators”. *Journal of Economic Behavior & Organization* 1 (3): 275–293. <https://EconPapers.repec.org/RePEc:eee:jeborg:v:1:y:1980:i:3:p:275-293>.
- Connolly, Robert, and Chris Stivers. 2003. “Momentum and Reversals in Equity-Index Returns during Periods of Abnormal Turnover and Return Dispersion”. *The Journal of Finance* 58 (4): 1521–1555. ISSN: 00221082, 15406261. <http://www.jstor.org/stable/3648220>.
- Coval, Joshua D., and Tobias J. Moskowitz. 1999. “Home Bias at Home: Local Equity Preference in Domestic Portfolios”. *The Journal of Finance* 54, no. 6 (): 2045–2073. doi:10.1111/0022-1082.00181. <http://doi.wiley.com/10.1111/0022-1082.00181>.
- . 2001. “The Geography of Investment: Informed Trading and Asset Prices”. *Journal of Political Economy* 109, no. 4 (): 811–841. ISSN: 0022-3808. doi:10.1086/322088. <https://www.journals.uchicago.edu/doi/10.1086/322088>.
- Devenow, Andrea, and Ivo Welch. 1996. “Rational herding in financial economics”. Papers and Proceedings of the Tenth Annual Congress of the European Economic Association, *European Economic Review* 40 (3): 603–615. ISSN: 0014-2921. doi:<https://doi.org/10.1086/322088>.

1016/0014-2921(95)00073-9. <http://www.sciencedirect.com/science/article/pii/S0014292195000739>.

Dixon, Matthew, Diego Klabjan, and Jin Hoon Bang. 2017. "Classification-Based Financial Markets Prediction Using Deep Neural Networks". *Algorithmic Finance*: 67–77. doi:10.2139/ssrn.2756331.

Doran, Kirk, Alexander Gelber, and Adam Isen. 2014. *The Effects of High-Skilled Immigration Policy on Firms: Evidence from H-1B Visa Lotteries*. Tech. rep. Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w20668. <http://www.nber.org/papers/w20668.pdf>.

Fama, Eugene F., and Kenneth R. French. 2015. "A five-factor asset pricing model". *Journal of Financial Economics* 116 (1): 1–22. ISSN: 0304-405X. doi:<https://doi.org/10.1016/j.jfineco.2014.10.010>. <http://www.sciencedirect.com/science/article/pii/S0304405X14002323>.

— . 1993. "Common risk factors in the returns on stocks and bonds". *Journal of Financial Economics* 33 (1): 3–56. ISSN: 0304-405X. doi:[https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5). <http://www.sciencedirect.com/science/article/pii/S0304405X93900235>.

Frank, Murray Z., and Vidhan K. Goyal. 2009. "Capital Structure Decisions: Which Factors Are Reliably Important?" *Financial Management* 38 (1): 1–37. doi:10.1111/j.1755-053X.2009.01026.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1755-053X.2009.01026.x>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-053X.2009.01026.x>.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press.

Greene, William. 2004. "The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects". *The Econometrics Journal* 7 (1): 98–119. doi:10.1111/j.1368-423X.2004.00123.x. eprint: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1368-423X.2004.00123.x>.

wiley.com/doi/pdf/10.1111/j.1368-423X.2004.00123.x. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1368-423X.2004.00123.x>.

- Grinblatt, Mark, and Matti Keloharju. 2001. "How Distance, Language, and Culture Influence Stockholdings and Trades". *The Journal of Finance* 56, no. 3 (): 1053–1073. ISSN: 00221082. doi:10.1111/0022-1082.00355. <http://doi.wiley.com/10.1111/0022-1082.00355>.
- . 2000. "The investment behavior and performance of various investor types: a study of Finland's unique data set". *Journal of Financial Economics* 55 (1): 43–67.
- Gutierrez Jr, Roberto C., and Eric K. Kelley. 2008. "The Long-Lasting Momentum in Weekly Returns". *The Journal of Finance* 63 (1): 415–447. doi:10.1111/j.1540-6261.2008.01320.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2008.01320.x>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2008.01320.x>.
- Hameed, Allaudeen, and G. Mujtaba Mian. 2015. "Industries and Stock Return Reversals". *Journal of Financial and Quantitative Analysis* 50 (1-2): 89117. doi:10.1017/S0022109014000404.
- Huberman, Gur. 2001. "Familiarity Breeds Investment". *Review of Financial Studies* 14, no. 3 (): 659–680. ISSN: 0893-9454. doi:10.1093/rfs/14.3.659. <https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/14.3.659>.
- Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics* 2, no. 2 (): 31–56. ISSN: 1945-7707. doi:10.1257/mac.2.2.31. <http://pubs.aeaweb.org/doi/10.1257/mac.2.2.31>.
- Hhn, Hannah Lea, and Hendrik Scholz. 2019. "Reversal and momentum patterns in weekly stock returns: European evidence". *Review of Financial Economics* 37 (2): 272–296. doi:10.1002/rfe.1037. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/rfe.1037>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/rfe.1037>.

- Ivkovic, Zoran, and Scott Weisbenner. 2005. "Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments". *The Journal of Finance* 60, no. 1 (): 267–306. ISSN: 00221082. doi:10.1111/j.1540-6261.2005.00730.x. <http://doi.wiley.com/10.1111/j.1540-6261.2005.00730.x>.
- Jegadeesh, Narasimhan. 1990. "Evidence of Predictable Behavior of Security Returns". *The Journal of Finance* 45 (3): 881–898. ISSN: 00221082, 15406261. doi:10.2307/2328797. <http://www.jstor.org/stable/2328797>.
- Jegadeesh, Narasimhan, and Sheridan Titman. 2001. "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations". *The Journal of Finance* 56 (2): 699–720. ISSN: 00221082, 15406261. <http://www.jstor.org/stable/222579>.
- . 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency". *The Journal of Finance* 48 (1): 65–91. ISSN: 00221082, 15406261. <http://www.jstor.org/stable/2328882>.
- Jensen, Michael C., and William H. Meckling. 1976. "Theory of the firm: Managerial behavior, agency costs and ownership structure". *Journal of Financial Economics* 3 (4): 305–360. <https://ideas.repec.org/a/eee/jfinec/v3y1976i4p305-360.html>.
- Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H1B Visa Reforms and U.S. Ethnic Invention". *Journal of Labor Economics* 28, no. 3 (): 473–508. ISSN: 0734-306X. doi:10.1086/651934. <https://www.journals.uchicago.edu/doi/10.1086/651934>.
- Leary, Mark T., and Michael R. Roberts. 2014. "Do Peer Firms Affect Corporate Financial Policy?" *The Journal of Finance* 69 (1): 139–178. doi:10.1111/jofi.12094. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12094>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12094>.
- Lehmann, Bruce N. 1990. "Fads, Martingales, and Market Efficiency". *The Quarterly Journal of Economics* 105 (1): 1–28. ISSN: 00335533, 15314650. <http://www.jstor.org/stable/2937816>.

- Leung, Mark T., Hazem Daouk, and An-Sing Chen. 2000. “Forecasting stock indices: a comparison of classification and level estimation models”. *International Journal of Forecasting* 16 (2): 173–190. <https://EconPapers.repec.org/RePEc:eee:intfor:v:16:y:2000:i:2:p:173-190>.
- Li, Zhixi, and Vincent Tam. 2018. “A Machine Learning View on Momentum and Reversal Trading”. *Algorithms* 11 (): 170. doi:10.3390/a11110170.
- Liang, Claire Y.C., Zhenyang (David) Tang, and Xiaowei Xu. 2019. “Uncertainty, Momentum, and Profitability”. *The Journal of Portfolio Management*. ISSN: 0095-4918. doi:10.3905/jpm.2019.1.085. eprint: <https://jpm.ijournals.com/content/early/2019/03/06/jpm.2019.1.085>. [full.pdf](https://jpm.ijournals.com/content/early/2019/03/06/jpm.2019.1.085.full.pdf). <https://jpm.ijournals.com/content/early/2019/03/06/jpm.2019.1.085>.
- MacKay, Peter, and Gordon M. Phillips. 2005. “How Does Industry Affect Firm Financial Structure?” *The Review of Financial Studies* 18, no. 4 (): 1433–1466. ISSN: 0893-9454. doi:10.1093/rfs/hhi032. eprint: <http://oup.prod.sis.lan/rfs/article-pdf/18/4/1433/24421130/hhi032.pdf>. <https://doi.org/10.1093/rfs/hhi032>.
- Messmer, Marcial. 2017. “Deep Learning and the Cross-Section of Expected Returns”. *SSRN Electronic Journal*. doi:10.2139/ssrn.3081555.
- Mithas, Sunil, and Henry C. Lucas. 2010. “Are Foreign IT Workers Cheaper? U.S. Visa Policies and Compensation of Information Technology Professionals”. *Management Science* 56, no. 5 (): 745–765. ISSN: 0025-1909. doi:10.1287/mnsc.1100.1149. <http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1100.1149>.
- Modigliani, Franco, and Merton H. Miller. 1958. “The Cost of Capital, Corporation Finance and the Theory of Investment”. *The American Economic Review* 48 (3): 261–297. ISSN: 00028282. <http://www.jstor.org/stable/1809766>.
- Moskowitz, Tobias J., and Mark Grinblatt. 1999. “Do Industries Explain Momentum?” *The Journal of Finance* 54 (4): 1249–1290. doi:10.1111/0022-1082.00146. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/0022-1082.00146>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/0022-1082.00146>.

- Mueller, Holger M., Paige P. Ouimet, and Elena Simintzi. 2017. "Within-Firm Pay Inequality". *The Review of Financial Studies* 30, no. 10 (): 3605–3635. ISSN: 0893-9454. doi:10.1093/rfs/hhx032. <https://academic.oup.com/rfs/article/30/10/3605/3855264>.
- Muga, Luis, and Rafael Santamara. 2007. "The Momentum Effect in Latin American Emerging Markets". *Emerging Markets Finance & Trade* 43 (4): 24–45. ISSN: 1540496X, 15580938. doi:10.2307/27750557. <http://www.jstor.org/stable/27750557>.
- Myers, Stewart C. 1977. "Determinants of corporate borrowing". *Journal of Financial Economics* 5 (2): 147–175. <https://EconPapers.repec.org/RePEc:eee:jfinec:v:5:y:1977:i:2:p:147-175>.
- Naughton, Tony, Cameron Truong, and Madhu Veeraraghavan. 2008. "Momentum strategies and stock returns: Chinese evidence". *Pacific-Basin Finance Journal* 16 (): 476–492. doi:10.1016/j.pacfin.2007.10.001.
- Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. *Foreign and Native Skilled Workers: What Can We Learn from H-1B Lotteries?* Tech. rep. Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w21175. <http://www.nber.org/papers/w21175.pdf>.
- Robichek, Alexander A., and Stewart C. Myers. 1966. "Problems in the Theory of Optimal Capital Structure". *The Journal of Financial and Quantitative Analysis* 1 (2): 1–35. ISSN: 00221090, 17566916. <http://www.jstor.org/stable/2329989>.
- Romer, David. 1993. "Rational Asset-Price Movements without News". *American Economic Review* 83, no. 5 (): 1112–1130. <https://ideas.repec.org/a/aea/aecrev/v83y1993i5p1112-30.html>.
- Rouwenhorst, K. Geert. 1998. "International Momentum Strategies". *The Journal of Finance* 53 (1): 267–284. ISSN: 00221082, 15406261. doi:10.2307/117441. <http://www.jstor.org/stable/117441>.
- Scharfstein, David S., and Jeremy C. Stein. 1990. "Herd Behavior and Investment". *The American Economic Review* 80 (3): 465–479. ISSN: 00028282. <http://www.jstor.org/stable/2006678>.



- Shleifer, Andrei, and Robert W. Vishny. 1992. "Liquidation Values and Debt Capacity: A Market Equilibrium Approach". *The Journal of Finance* 47 (4): 1343–1366. doi:10.1111/j.1540-6261.1992.tb04661.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1992.tb04661.x>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1992.tb04661.x>.
- Simintzi, Elena, Vikrant Vig, and Paolo Volpin. 2015. "Labor Protection and Leverage". *Review of Financial Studies* 28, no. 2 (): 561–591. ISSN: 0893-9454. doi:10.1093/rfs/hhu053. <https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhu053>.
- Takeuchi, L. Nozomu. 2013. "Applying Deep Learning to Enhance Momentum Trading Strategies in Stocks".
- Tang, Gordon Y. N., and Haomin Zhang. 2014. "Stock return reversal and continuance anomaly: new evidence from Hong Kong". *Applied Economics* 46 (12): 1335–1349. doi:10.1080/00036846.2013.872767. eprint: <https://doi.org/10.1080/00036846.2013.872767>. <https://doi.org/10.1080/00036846.2013.872767>.
- Tate, Geoffrey A., and Liu Yang. 2015a. "Female leadership and gender equity: Evidence from plant closure". *Journal of Financial Economics* 117, no. 1 (): 77–97. ISSN: 0304-405X. doi:10.1016/J.JFINECO.2014.01.004. <https://www.sciencedirect.com/science/article/pii/S0304405X1400018X>.
- . 2015b. "The Bright Side of Corporate Diversification: Evidence from Internal Labor Markets". *Review of Financial Studies* 28, no. 8 (): 2203–2249. ISSN: 0893-9454. doi:10.1093/rfs/hhv012. <https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhv012>.
- . 2015c. "The Human Factor in Acquisitions: Cross-Industry Labor Mobility and Corporate Diversification". *SSRN Electronic Journal* (). ISSN: 1556-5068. doi:10.2139/ssrn.2667940. <http://www.ssrn.com/abstract=2667940>.

- Trippi, Robert R., and Duane. DeSieno. 1992. "Trading Equity Index Futures With a Neural Network". *The Journal of Portfolio Management* 19 (1): 27–33. ISSN: 0095-4918. doi:10.3905/jpm.1992.409432. eprint: <https://jpm.iijournals.com/content/19/1/27.full.pdf>. <https://jpm.iijournals.com/content/19/1/27>.
- Trueman, Brett. 1994. "Analyst Forecasts and Herding Behavior". *The Review of Financial Studies* 7 (1): 97–124. ISSN: 08939454, 14657368. <http://www.jstor.org/stable/2962287>.
- Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2010. "Information Acquisition and Under-Diversification". *The Review of Economic Studies* 77, no. 2 (): 779–805.
- Wei, Zhanshun, and Liyan Yang. 2012. "Short-Term Momentum and Reversals in Large Stocks". *SSRN Electronic Journal* (). doi:10.2139/ssrn.2029984.
- Welch, Ivo. 2004. "Capital Structure and Stock Returns". *Journal of Political Economy* 112 (1): 106–131. ISSN: 00223808, 1537534X. <http://www.jstor.org/stable/10.1086/379933>.
- Zeckhauser, Richard, Jayendu Patel, and Darryll Hendricks. 1991. "Nonrational actors and financial market behavior". *Theory and Decision* 31, no. 2 (): 257–287. ISSN: 1573-7187. doi:10.1007/BF00132995. <https://doi.org/10.1007/BF00132995>.
- Zwiebel, Jeffrey. 1995. "Corporate Conservatism and Relative Compensation". *Journal of Political Economy* 103 (1): 1–25. ISSN: 00223808, 1537534X. <http://www.jstor.org/stable/2138716>.