

ESSAYS IN ENTREPRENEURSHIP AND FINANCE

A Dissertation
Submitted to
the Temple University Graduate Board

In partial fulfillment of the
requirements of the degree
of

DOCTOR OF PHILOSOPHY IN ECONOMICS

by
R. David Ratigan
December, 2019

Examining Committee Members:

Douglas Webber, Advisor, Department of Economics

Michael Leeds, Department of Economics

Erwin Blackstone, Department of Economics

Peter Zaleski, External Member, Villanova University Economics Department

ABSTRACT

ESSAYS IN ENTREPRENEURSHIP AND FINANCE

by

R. David Ratigan

In this dissertation, I study three economic issues in the field of entrepreneurship and a fourth that looks at the network structure of financial markets.

Chapter 1, titled "Patenting and Nascent Firm Performance", examines how patents interact with nascent firm outcomes. The primary motivation follows a puzzle present in the data and discussed in the literature. Patents are prized possessions for business owners but the value that they confer is dubious. In this sample, patenting firms have no better survival rates and have significantly worse profit outcomes. A Oaxaca decomposition demonstrates the reason for this lies with patents' attraction of equity financing. The burden of the equity relationship proves to be a strain on the young firms.

Chapter 2, titled "Signal vs Appropriation Value in Patenting" provides insight into the causal connection between patents and equity financing. This chapter also weighs in on the source of the equity firms' attraction to patents. Controlling for unobservable influences via fixed effects regression, the value measured by the coefficient on patent, demonstrates the appropriative value of the patent, rather than the signal value. Further examination reveals that equity firms are very sensitive to the legal status of the firm, eschewing general partnerships.

Chapter 3, titled "Peer Effects in Entrepreneurship", studies gender differences in firm financing as a function of response to peers. This research borrows from the social interaction modeling pioneered by Charles Manski by adopting methods to include peer variables. The main finding is that males are positively and significantly affected by peer financing levels but females are not. This result is robust to industry, experience, credit risk, and other financing sources.

Chapter 4, "Network Analysis of the S&P 500", examines the financial market as a network. The literature takes two approaches to performing this analysis but no study has compared the two approaches. This chapter is the first study to compare them and provide insight as to which provides the more descriptive model. The threshold correlation approach provides much more useful results for any kind of analysis involving market connectedness and shock transition dynamics.

DEDICATION

To my wife and daughter. I am grateful to Ellen for her patience and continual support through this long process. I am grateful to Lila for being my motivation to finish.

TABLE OF CONTENTS

ABSTRACT	ii
DEDICATION	iv
LIST OF FIGURES	viii
LIST OF TABLES	ix
 CHAPTER	
1 PATENTS AND NASCENT FIRM PERFORMANCE	1
1.1 Introduction	1
1.2 Background and Literature Review	3
1.2.1 Nascent Ventures and Funding	3
1.2.2 Adverse Selection and Pecking Order in Funding	4
1.2.3 IP and Overcoming Asymmetrical Information	7
1.2.4 Patenting Effects on Firm Outcomes	8
1.2.5 Investment Effects on Firm Outcomes	9
1.3 Gaps in the Literature and Testable Hypotheses	12
1.4 Data and Variables	12
1.4.1 Outcome Variables	14
1.4.2 Key Explanatory Variable	14
1.4.3 Control Variables	15
1.4.4 Variable Transformations	17
1.5 Analysis	17
1.5.1 Descriptive Statistics	17
1.5.2 H_1 : The Effect of Patenting on Profits	19
1.5.3 H_2 : Adverse Selection	24
1.6 Conclusion	27
 2 SIGNAL VS APPROPRIATION VALUE OF PATENTS	 29
2.1 Introduction	29
2.2 Background and Literature Review	31
2.2.1 Nascent Ventures, IP, and Funding	31
2.2.2 Patenting Effects on Investment	33
2.2.3 Appropriation vs. Signaling	35
2.3 Testable Hypotheses	38

2.4	Data and Variables	38
2.4.1	Dependent Variable	39
2.4.2	Key Explanatory Variable	40
2.4.3	Control Variables	40
2.4.4	Variable Transformations	43
2.5	Analysis	43
2.5.1	Descriptive Statistics	43
2.5.2	Patenting and Investment	44
2.5.3	Model Diagnostics	48
2.5.4	Heterogeneous Effects	50
2.5.5	Robustness Checks	53
2.6	Conclusion	54
3	PEER EFFECTS IN ENTREPRENEURSHIP	55
3.1	Introduction	55
3.2	Background	57
3.2.1	Industry	57
3.2.2	Risk Preference	58
3.2.3	Firm Financing	59
3.2.4	Prospect Theory	60
3.3	Data and Variables	61
3.3.1	Outcome Variable	62
3.3.2	Main Explanatory Variable	62
3.3.3	Descriptive Statistics	62
3.4	Estimation and Results	64
3.4.1	Diagnositics and Robustness	67
3.4.2	Heterogeneous Effects	69
3.5	Discussion and Conclusion	71
4	NETWORK ANALYSIS OF THE S&P 500	73
4.1	Introduction	73
4.2	Background	75
4.2.1	Previous Studies	75
4.2.2	Theoretical Perspective	78
4.3	Data and Methodology	81
4.3.1	Data and Variables	81
4.3.2	Network Formation	82
4.4	Empirical Results	84
4.4.1	Step One - Data Preparation	84
4.4.2	Step Two - Network Graphs	84
4.4.3	Step Three - Machine Learning Classification	91
4.5	Simulation	93
4.6	Conclusion	94
	BIBLIOGRAPHY	96

APPENDIX A: THREEFOLD OAXACA DECOMPOSITION	110
APPENDIX B: LOG EQUITY RESULTS	114

LIST OF FIGURES

4.1	Kernel density plots of Pearson and Spearman Correlation	84
4.2	A histogram of Augmented Dickey Fuller p-values on the residuals of pairwise regressions.	85
4.3	A spring network of Pearson correlations for the S&P 500.	85
4.4	Positive correlation network graphs at four threshold values.	86
4.5	Correlation and Cointegration Approach Graphs	87
4.6	Negative Correlation and Cointegration Approach Graphs	87
4.7	Edge density as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.	88
4.8	Cointegration Graph: Edge density as the threshold value changes on the cointegration-formed network.	88
4.9	Node degree as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.	89
4.10	Cointegration Graph: Mean node degree as the threshold value changes on the cointegration-formed network.	89
4.11	Clustering coefficient as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.	90
4.12	Mean cluster coefficient as the threshold value changes on the cointegration-formed network.	90
4.13	Number of connected components as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.	91
4.14	The number of connected components as the threshold value changes on the cointegration-formed network.	91
4.15	Positive correlation network graph histograms of degree with graph insets.	92
4.16	KNN confusion matrix heatmap for correlation network.	92
4.17	KNN confusion matrix heatmap for cointegration network.	93
4.18	Simulated Cointegrated Series.	94

LIST OF TABLES

1.1	Patent Descriptives	18
1.2	Finance Descriptives	18
1.3	Patenting Firms vs Non-Patenting Firms Financing	19
1.4	Conditional Outcomes	20
1.5	Dependent Variable: Log Transform of Profits	21
1.6	Oaxaca Decomposition on Ln(External Equity)	27
2.1	Patent Descriptives	44
2.2	Finance Descriptives	44
2.3	Patenting Firms vs Non-Patenting Firms Financing	45
2.4	Legal Status vs Patent Ownership in 2011	45
2.5	Average Venture Capital Financing by Year and Legal Status	45
2.6	Dependent Variable: Binary Indicator of Positive Equity Financing	48
2.7	Time Varying Coefficients	49
2.8	Patent Interacted with Legal Status on Binary Dependent	52
2.9	Patent Estimates on Subset Data by Legal Status on Binary Dependent	53
2.10	Dependent Variable: Log of External Financing and Internal Equity	53
2.11	Dependent Variable: Binary of External Financing and Internal Equity	54
3.1	Gender Ownership	63
3.2	Finance Descriptives	63
3.3	Average Total Financing by Gender	64
3.4	External Equity by Gender	64
3.5	External Debt by Gender	65
3.6	Dependent Variable: Log of Total Financing	67
3.7	Time Interaction with all Control Variables	68
3.8	Legal Status Interaction	70
3.9	Industry Interaction	71
A1	Dependent Variable: Log of External Equity Financing	114

CHAPTER 1

PATENTS AND NASCENT FIRM PERFORMANCE

with Michael Risch
Vice Dean, Villanova Law

1.1 Introduction

In the quest for empirical verification of the role of patents in entrepreneurial firm success, venture funding, and survival tracking present a conundrum of sorts. Adverse selection, moral hazard, and pecking order theories predict that venture-funded firms will fail more often, but patenting can solve asymmetric information problems, allowing venture capitalists to choose only the best ventures. If this is so, then patenting should lead to more debt financing rather than equity, as the better firms should prefer debt to equity in the pecking order.

Thus, two primary questions remain stubbornly difficult to answer. First, do patents improve firm outcomes? Second, why or why not? This article uses a novel approach to a rich dataset to answer these questions from the perspective of venture funding and survival. More particularly, we explore the interaction between pecking order theory and patenting, an area rarely directly addressed. Answering these questions presents thorny, difficult to resolve issues. Three sources of variation can confound the estimation of the value of the patent. The first is that patent holders

might receive different valuations just for holding a patent. The second is that patent holders might all have correlated unobserved traits that are driving outcomes (Gartner, Frid, and Alexander 2012; Hoenig and Henkel 2015; Engel and Keilbach 2007). The third is that patent holders might, on average, have different levels of observed covariate levels compared to non-patent holders. This paper resolves those issues to estimate what remains, the true effect of holding a patent in venture funding and outcomes.

This article makes two contributions to the literature. First, it tests the role of adverse selection in venture funding, and finds little evidence to support adverse selection in patenting. In other words, patenting firms do not appear to obtain venture funding rather than debt because they are riskier investments. Whether signaling innovation or providing some other positive characteristic, patents appear to do more than hide an otherwise lesser investment.

Second, despite these positives, patent-holding funded firms have worse survival than all other firms. But the problem is not the patent, *per se*. Instead, heterogeneity analysis shows the culprit to be excessive growth by venture funded firms. Employment grows without a concomitant growth in revenue or profits, thus starving firms of cash. We discuss these findings in terms of moral hazard. Though the results are a bit surprising, our findings are consistent with prior work showing that venture capitalist firms often do not meet the expectations of their portfolio firms (Berg-Utby, Sorheim, and Widding 2007).

We support these results using a novel method—the Oaxaca decomposition. We use this decomposition to isolate the role of patenting, equity, and other endowments in firm profits.

This article proceeds as follows. We begin with an overview of startups, IP’s role in financing, and a brief discussion of adverse selection, moral hazard, and pecking order. We then report on prior findings that tie IP and venture funding (or both)

to firm outcomes. This leads to two testable hypotheses, which we outline. We then discuss our data, model and estimates used to test our hypotheses. The article concludes with a discussion of the implications.

1.2 Background and Literature Review

1.2.1 Nascent Ventures and Funding

The constant churn of the labor market, firms and sectors rising and falling, and new technologies pushing out the old necessitate a constant need for new entrepreneurs to fill any gaps and explore new territory. Over the last twenty years, there have been an average of 650,000 new firms each year, many of which enter as another is exiting. Those nascent¹ firms have accounted for the creation of more than 3.5 million jobs over the same period. Even though large firms, those with more than 250 employees, account for a large percentage of national employment, the entrepreneur is a central figure in the economy and vital to its health. As of 2015, small firms account for over 46% of national employment (BLS 2016).

A great deal of literature looks at the factors that drive and enable an entrepreneur to start a business. Those factors include intellectual property, taxes, number of regulations, and many others. Entrepreneurs must travel a complicated and difficult landscape. Early firm characteristics and owner characteristics have been shown to greatly shape a firm's capital structure. That structure in turn, shapes the development of that firm, influencing its likelihood of survival.

In theory, cash reserves are necessary to firm survival (Liu et al. 2006). Thus, Parker and Belghitar (2006) look at entrepreneurs in the planning phase before operations begin and find that access to personal or external sources of financing are

¹Nascent is used to define newly founded firms as opposed to firms in the development and pre-operation stage.

crucial to a firm's success.

Robb and Robinson (2010) use the same data source we do and conclude that firms rely heavily on external sources of financing as opposed to family and friends. Coleman and Robb (2009) also study firm financing in the Kauffman Survey but focus on differences between genders. Their paper is interesting since it considers supply-side constraints, such as investor preferences for certain industries and demand-side constraints, such as entrepreneur preferences for personal financing.

A number of studies reference access to capital as an important growth constraint, such as Beck and Demirguc-Kunt (2006) who find that an increased access to external financing helps level the playing field among firms of different sizes. Paglia and Harjoto (2014) focus on private equity and venture capital and show that access to such sources is very restricted to small and medium sized firms but those who do attain it benefits along a number of growth metrics.

1.2.2 Adverse Selection and Pecking Order in Funding

Financing is not free, however, and the rational entrepreneur seeks out the cheapest forms of financing. This is the notion behind the pecking order hypothesis - firms utilize the cheapest forms of capital first, only pursuing venture capital, for example, if not supplied with business loans. This also gives rise to adverse selection because venture capital firms are likely to be left with a pool of riskier firms, those not able to secure business loans. Myers (1984) and Myers and Majluf (1984) present the basic arguments and model for adverse selection: that due to information asymmetry, firms with debt will do better than firms with equity because the founders have better information. Because founders seek debt only when they expect to survive and earn high profits, venture capitalists are left with only the worst opportunities. This leads to the traditional pecking order, as entrepreneurs seek the lowest cost of capital in the presence of information asymmetry. Internal funding is preferred to debt, which

is preferred to venture funding (Scellato 2007; Hechavarría, Matthews, and Reynolds 2016; Paul, Whittam, and Wyper 2007a).

Studies of adverse selection are too numerous to discuss in full, but the theory has been tested many times with varying results. Several studies show a clear pecking order among startup firms (Michaelas, Chittenden, and Poutziouris 1999; Ou and Haynes 2006; Frid 2009; Vanacker and Manigart 2010). Similarly, findings that funded firms had no other funding sources (Peneder 2010) are consistent with adverse selection theory. Other work showing that debt funded firms survive longer than VC funded firms also supports adverse selection theory (Cole and Sokolyk 2018; Åstebro and Bernhardt 2003). Still other studies find the distribution of actual funding to match expected funding distribution (Mina, Lahr, and Hughes 2013).

There is a limit to the inference here, however, because personal debt does not lead to longer survival. If personal debt (that is, internal funds) are associated with firms that fail, this implies a different order of selection: firms prefer bank lenders at the lowest risk, then venture capitalists, and then internal financing and personal debt at the highest risk. This is a counter-pecking order.

Based on this apparent preference for external rather than internal funds among startups, another study finds similar results counter to the traditional pecking order. Moskowitz and Vissing-Jørgensen (2002) shows that most private equity is concentrated in individuals who manage their own firms. This equity is non-diversified and thus more risky, and the returns to that equity are no better than returns available for publicly traded equities. This seems contrary to adverse selection theory and the traditional pecking order; one would expect that non-diversified private equity would seek VC funding to mitigate the risk and grow returns, but they either do not or are unable to do so. The authors consider reasons for their results: benefits of control, preference for skewness, and misperception of failure probabilities. Each of these (and especially misperception) weakens the tenets of adverse selection theory, which

assumes that founders know which firms are likely to succeed and which are not.

Another study found counterintuitive results: that firms expecting large revenues obtain more personal and external financing, while firms expecting to grow do not (presumably obtaining debt) (Gartner, Frid, and Alexander 2012). This is contrary to the adverse selection story, because firms expecting revenues would have receivables to secure debt and would not need to dilute capital. Nonetheless, these results are somewhat consistent with our anecdotal evidence that firms expecting to make it big sought out venture capital funding because they sought growth and public offerings that are less likely with debt. Two foreign surveys that support this view (Boyer and Blazy 2014; Paul, Whittam, and Wyper 2007b).

In another contrary study, Huyghebaert and Gucht (2007) find that start-up firms with higher risk profiles have less bank debt, not more. Instead, they resort to personal debt. This makes sense, but implies that venture funding is not a result of asymmetric information. Rather it is a result of the inability to obtain bank debt. Similarly, others find a preference for debt, but suggest that the reasons are efficiency and not adverse selection (Rocca, Rocca, and Cariola 2011).

The discussion here suggests a testable hypothesis: where the pecking order is followed, the presence of venture capital implies information asymmetry and weaker ventures (Myers 1984; Hechavarría, Matthews, and Reynolds 2016). In short, we might expect venture funded firms to fail earlier, have lower profits, or otherwise show higher risk. But this is not a static system (Myers 1984). Both entrepreneurs and venture capitalists know the pecking order, and do what they can to overcome adverse selection (in addition to other factors that lead to the pecking order, such as moral hazard).

1.2.3 IP and Overcoming Asymmetrical Information

Intellectual property plays an important role in how firms overcome the adverse selection problem in a dynamic system. One of the characteristics of early stage firms is the type of intellectual property they hold. In the U.S. (and just about everywhere else), companies use four primary types of IP to appropriate value. Patents grant the right to exclude others from using claimed inventions, whether others copied them or developed them independently. Copyrights protect expression, from books to music to computer software, from being copied by others. Trade secrets protect all manner of information from the prying eyes of others—the primary defendants are usually ex-employees. Trademarks protect brand names from potentially confusing uses.

Myers and Maljuf make clear the importance of verifiable information that the firm is not selecting for external capital due to low quality. Where VC firms are able to mitigate information asymmetries, their target firms are more likely to receive investment (Lockett, Murray, and Wright 2002; Cumming and Johan 2008). Based on the theoretical framework, IP should aid in attracting financing (Long 2002; Mazzoleni and Nelson 1998; Hahn, Kim, and Kwon 2017; Tondo-Kramer 2009; Fabrizi et al. 2013).

While there is a debate about the signal value versus the appropriation value of patents (Hall 2018), the focus of this paper does not turn on the signaling v. appropriability question. The literature tends to focus on the signal aspect of patenting. Assuming that signal correlates with quality, venture firms obtain verifiable information, the same as if they were investing in the appropriability of the patent. Patents, therefore, reduce the inefficiencies of adverse selection. For example, Peneder (2010) report that venture capital firms select for innovative firms. Either benefit of patenting would serve to offset the adverse selection we might expect in venture funding.

1.2.4 Patenting Effects on Firm Outcomes

A related question is whether patents improve firm outcomes. Some studies find employment growth in patenting firms after venture capital investment (Engel and Keilbach 2007; Farre-Mensa, Hegde, and Ljungqvist 2017), but this is not a helpful outcome measure. Venture capital funds are almost always associated with employment growth—that’s how firms spend the money they receive. The final result of that employment growth matters more.

Farre-Mensa, Hegde, and Ljungqvist (2017) find that startups that have an easier time obtaining a patent have substantially higher sales (80%) growth five years later, primarily due to venture capital investment. In the Crunchbase study, Lerman (2015) finds that firms with patents are more likely to be acquired. Useche (2014) finds that IPO pricing is greater for those firms with patents. Cao and Hsu (2011) find that firms with patents are more likely to reach IPO and less likely to declare bankruptcy, but the firms in their study had all already received some financing. Balasubramanian and Sivadasan (2011) examine 48000 companies over 22 years to find that patenting is associated with increased firm capital, skill, and size. They find that firm growth is caused by sales around the time of patenting, and that patenting is associated with the introduction of new products. Mann and Sager (2007) find correlations between patenting and firm longevity among those firms that have already received at least one round of financing. Gurmur and Perez-Sebastian (2008) find growth in R&D spending after patenting in a ten year longitudinal panel. Czarnitzki and Kraft (2010) link profitability growth to patent stock among German firms. From a signaling vs. appropriability point of view, this might be attributed to patent and other intangible asset reduction of market uncertainty; R&D falls with greater uncertainty (Czarnitzki and Toole 2011; Brown and Kimbrough 2011). As such, patent portfolio size can increase returns to R&D (Noel and Schankerman 2013). Similarly, firms tend to patent more in large markets or when they have long development time (Lahr and

Mina 2016). One study found a negative effect on survival for patent applications, but a positive effect on survival for granted applications among Australian firms (Buddelmeyer, Jensen, and Webster 2010).

Hellmann and Puri (2000) examine "innovator" and "imitator" firms (without respect to patenting) and find that innovator firms are more likely to obtain venture capital, and—in turn—bring products to market more quickly. Similarly, Triguero, Corcoles, and Cuerva (2014) find that product or service innovations (not necessarily patents) led to increased employment in Spanish firms. Colombelli, Krafft, and Vivarelli (2016) found similar results for French firms. Cefis and Marsili (2011) reported that product innovation led to longer firm survival, though patents had no discernible effect. Stucki (2014) also find that innovation was correlated with firm success. Interesting, Boyer and Blazy (2014) find that innovation was associated with a substantial decrease in survival rate among French start-up firms.

One study (Tsvetkova, Thill, and Strumsky 2014) attempted to consider the interactive effect of patents on survival for VC funded firms for a subset of firms in the computer and electronics industry. It found that patents had an effect on survival, but it removed VC funding from estimates due to insignificant results. Baum and Silverman (2004) asked whether startups benefit from investment because venture capitalists are better at spotting talent or are better at nurturing the firms in which they invest. They find that patenting and alliances both lead to investment and improved firm success, but that management experience leads to investment but not necessarily better outcomes.

1.2.5 Investment Effects on Firm Outcomes

Whether due to adverse selection or not, only a fraction of firms funded by VCs succeed (Gorman and Sahlman 1989). Even so, VC funded firms tend to do better than non-funded firms. For example, VC funded firms often have more revenues (Colombo

and Murtinu 2017; Bhattacharya, Borisov, and Yu 2015). In times of uncertainty, VC funded firms are able to extract more rents (Alperovych and Hübner 2013). Further, they often experience more growth, though portfolio firms are often selected for growth, even among those firms that had no other funding sources (Peneder 2010; Puri and Zarutskie 2012). In one contrary study, Stucki (2014) compared financial constraints in Swiss firms and found that debt constraints affect firm survival, but that VC constraints do not; the authors suggest that the findings may be limited to the Swiss financing market.

Sometimes venture capital leads to more patenting. Kortum and Lerner first introduced this notion, and others have replicated it in a variety of contexts (Kortum and Lerner 2000; Popov and Roosenboom 2012; Hall and Lerner 2009; Bertoni, Croce, and D’Adda 2010). Some have gone so far to find this effect even for firms in close proximity to venture capital (Kolympiris and Kalaitzandonakes 2013). Still others have failed to replicate these findings, at least at the firm level (Hirukawa Masayuki and Ueda Masako 2011; Arvanitis and Stucki 2014; Engel and Keilbach 2007; Lahr and Mina 2016).

Davidsson and Gordon (2012) provide a quality literature review of VC outcome studies. They pay special attention to the difficulties of defining firm success. There are very few firm survival studies considering the differences between firms that receive VC funding and those that do not. Most of the studies that consider non-venture funded companies select those firms that are acquired or reach an IPO and consider survival or other metrics from that point (Megginson and Weiss 1991; Jain and Kini 2000; Ivanov and Xie 2010; Tian 2011; Chou, Cheng, and Chien 2013; Ragozzino and Blevins 2016).

Three studies compare survival rates of VC and non-VC funded companies from formation. The earliest, Manigart, Baeyens, and Van Hyfte (2002), matched non-VC funded companies with the VC funded companies in Belgium. They found that non-

funded companies had a better survival rate, and that there were differences among the types of VC firms: firms in seasoned government funds—which did not have a high growth/”cut your losses” drive—had better survival, while firms in other government funds were much, much worse. Private VC firms had survival rates about the same as those of non-funded firms. This is a mixed story for adverse selection. On the one hand, it appears that worse firms received venture funding, supporting pecking order theory. On the other hand, some firms performed better than others (and better than non-funded firms), which implies that venture capitalist screening and management goals are more important than inherent information asymmetry. The authors also found that higher debt ratios led to more bankruptcy, which is contrary to findings in other studies of better survival for companies financing with debt rather than equity. The study did not consider the role of patents.

The second study considers twenty-five years of VC funded firms and matched non-VC counterparts, finding that VC firms have no better profitability at exit than non-VC firms (Puri and Zarutskie 2012). They also find that VC firms have a lower cumulative failure rate, though the differences are only slight once initial factors, such as initial size, industry, and geography, are considered. Importantly, the authors examine VC behavior toward failed firms, which they consider to be not well known. They find that the survival rate difference is driven by differences in early life failures (presumably because similarly situated funded firms have the cash to survive). In short, their results imply that VC firms are ”patient” with portfolio firms during the first five years after funding, but after that time the failure rate is much higher than non-VC funded companies as the venture capitalists abandon firms that are not trending toward a lucrative exit. They also find that VC firms have more employees and sales, but not more profits, than non-VC companies at the time of failure. However, this study did not consider patenting, and performed little testing on innovation.

The third study is Chemmanur, Krishnan, and Nandy (2011), which used the

U.S. Census Longitudinal Research Database and found that non-service VC funded companies performed better both before and after they received funding, and that non-VC funded firms had a lower survival rate. This study, in particular, contradicts the adverse selection theory. The authors distinguish Puri and Zarutskie (2012) by focusing on screening and monitoring, but the results are somewhat similar. This study did not consider patenting either.

1.3 Gaps in the Literature and Testable Hypotheses

The theoretical background and empirical landscape yield two testable hypotheses that remain mostly unanswered in the literature:

H₁: Patents and Outcomes: Patenting firms have no greater profits

H₂: Patents and Adverse Selection: If patenting firms do not have greater profits, the reason is adverse selection

Our approach to these hypotheses fills a gap in the literature. Few studies compare profits of unfunded companies with funded companies to test adverse selection, and none of those studies consider the role of patents in reducing information asymmetry.

1.4 Data and Variables

The Kauffman Firm Survey is unique micro-panel of for-profit, non-subsidary, non-inherited businesses formed in 2004. There were 4,928 firms at the start of the survey, but due to attrition, only 2,032 firms survived for the length of the survey.

The survey oversamples firms classified as high or medium tech based on the amount of research and development employment in that company's industry. The survey also over-samples female owned firms. Roughly 25% of initial entrepreneurs are

women and roughly 13% are considered to be hightech². Despite the oversampling, the Kauffman Survey was intended to be representative of the full U.S. economy. The survey represents 25 two-digit industry codes (industry sector) and 85 three-digit industry codes (industry subsectors) by the North American Industry Classification System (NAICS)³.

Due to attrition and a high non-response rate, the Kauffman Foundation provides a few versions of the data. We use the multiply imputed longitudinal set which removes firms that do not respond each year of the survey in which they are still operating. The longitudinal sample has 3,140 firms at baseline and 1,630 in the final year.

The primary advantage of using the imputed data is the logical imputation that fills in holes where it can be done with certainty. For example, one can logically impute a person's race if that information was previously supplied. Despite this effort, there are many holes in the data. The missing values that could not be logically imputed were split into hard and soft missing values following Stata convention. Soft missing values are imputed because they can reasonably be assumed to take a certain value. For example, the race of an owner can safely be imputed if a previous value was recorded. Hard missing values are not imputed. Some of the analysis below uses the subset of surviving firms in year eight. Among this group the regressions are affected so much by these hard missing values that the sample of unaffected observations to the sample including observations affected by hard missing values is a ratio of about 1 to 6. Many of these hard missing values are the continuous variables of interest like profit. The imputation helps but still there are large gaps in the data.

²Hightech is a KFS classification based on industry average for research and development employment.

³The NAICS replaced the Standard Industrial Classification (SIC) in 1997.

1.4.1 Outcome Variables

The majority of this study tracks profits as the primary outcome variable. This variable is used as a measure of firm success or firm well-being. The goal is to see how patenting interacts with this outcome variable to get a sense of how well patenting is converted into success. Profits are measured as a continuous variable and range over the course of the survey from -\$500,000 to over \$1,000,000.

This study also considers equity financing as a secondary outcome variable. Equity is an important outcome that captures how attractive patenting makes firms. The source of equity is external, from angel, venture, or other financial institutions, rather than equity from family, friends, owners, or employees. Equity has an important distinction from debt financing as equity financing is typically in larger amounts and the investment is analyzed for its potential to return profit rather than just for its safety.

1.4.2 Key Explanatory Variable

The key explanatory variable is a patenting dummy. Although the survey provides a count measure as well, we consider only the binary indicator. The count variable does not add to the analysis or change the results in any meaningful way. Furthermore, use of a dummy is supported by the literature, which finds that the crucial distinction is patent vs no patent, not the number of patents (Mann and Sager 2007; Lahr and Mina 2016).

The survey does not provide any information on when the patent was filed, any detailed information, or any information from the patent and trademark office. This lack of information shapes the analysis undertaken in the paper, preventing any instrumental variable regression.

1.4.3 Control Variables

In addition to the key explanatory variable, the regression models control for a number of additional factors that are likely correlated with both the outcome and key explanatory variable.

Owner Characteristics

The owner characteristics included are gender, race, education, experience within industry, how many previous startups, and credit risk. Only variables of theoretical and intuitive interest are included in the list. Gender, race, and education are categorical variables while experience and startups are continuous variables. We choose to code the credit risk variable, being discretely valued and ordinal, as categorical because it doesn't display a linear relationship with either outcome variable.

Gender is an important control variable for two reasons. Gender is an important signal that captures motivation (Burke, FitzRoy, and Nolan 2002) and firm management strategies. Gender also has demand- and supply-side implications in the equity market including risk preference (Charness and Gneezy 2012) and reliance on external sources of financing (Coleman and Robb 2009).

The survey provides an education of owners variable with ten levels, which we consolidated into four to increase its explanatory power: high school degree or less, technical/trade/vocational, some college/associates, four year college degree or more. Although we control for industry sector, an owner's education level is correlated with more granular industry indicators, attractiveness to debt and equity financiers, and likely correlated with other unobservable measures of business acumen.

The survey provides both the owners' ages and experience levels. To avoid collinearity, we use the experience level only. Experience is measured in years within the industry in which the startup was formed. Experience captures industry knowledge and access to professional networks. For those reasons experience also captures at-

tractiveness to investors.

The KFS provides the number of previous startups by entrepreneur to capture entrepreneurial experience. This is a count variable that ranges from zero to five. The survey provides an additional dummy variable asking if any of the previous were in the same industry as the current startup to refine the count measure. We use this variable as well (called "Same Ind") with three levels: First Startup; No; Yes. These variables convey another rich measure of experience and can act as a strong signal to potential investors and lenders. On the other hand, serial entrepreneurs would have the opposite effect. The Kauffman Survey does not provide any indication of the success of these previous ventures.

We have access to two measures of credit risk. The first is the D&B commercial credit rating (Credit Risk), measured as one, the highest quintile of credit ratings, to five, the lowest quintile of credit rating, with the intermediate scores representing ascending quintiles. The Kauffman Survey also provides the D&B financial stress score (FSSP) measured from one to one hundred with higher numbers representing decreasing likelihood of financial stress. We focus on the D&B credit risk variable due to high collinearity. We use this measure as a good proxy for financial risk because financial risk might correlate with other risk taking behavior.

Firm Characteristics

Control variables related to firm characteristics are financing sources, profits, assets, employees, legal status, and NAICS two-digit code. Financing, profits, assets, and employees are continuous variables. Legal status and NAICS are categorical variables.

We separate financing into external equity and other financing since external equity is a special concern. The other financing variable contains internal equity, internal debt, and external debt. The sign on both of these measures could reasonably be positive or negative. Large start-up financing amounts is correlated with firm size

but firm size does not necessarily mean profits (Cassar 2004).

Legal status (Sole Proprietorship, LLC, Subchapter S-Corporation, C-Corporation, General Partnership, and Limited Partnership) control for a number of factors and is quite important, affecting magnitude and significance of the other variables. Organizing as a corporation rather than a sole proprietorship, LLC, or partnership has implications for liability and taxes. Corporations protect owners from business liabilities at the expense of additional taxes, paperwork, and legal requirements. The structure of the business is a strong signal of preparation and experience, and also of vision of growth. We expect more successful firms to take corporation status. Due to the lack of observations, we drop general partnership and limited partnership in the regression models.

1.4.4 Variable Transformations

The continuous variables of financing, profit, and assets have extremely high variances even though most observations (5th to 95th quantile) are close to the mean. To alleviate the impact that outliers have on estimates, we take the natural logarithm. One was added to each variable, due to zeros in the data. The range of profits extended far into the negative numbers as well, so an additional step was taken. The natural log was applied to the absolute value of profit before reapplying the sign.

1.5 Analysis

1.5.1 Descriptive Statistics

Table 2.1 presents preliminary trends regarding attrition and patent ownership. Even though this survey oversamples high technology firms, the patenting rate is quite low. The rate of attrition can be observed in the change in number of firms. It is clear from the simple statistics that patenting is not a broad protection from going out of

Table 1.1: Patent Descriptives

Year	Firms	Patent Holders	Holder/Firms	Patents	Patents/Holder
2004	3140	120	3.82%	584	4.87
2005	2837	124	4.37%	608	4.9
2006	2554	122	4.78%	487	3.99
2007	2330	106	4.55%	556	5.25
2008	2092	86	4.11%	622	7.23
2009	1928	82	4.25%	386	4.71
2010	1775	72	4.06%	355	4.93
2011	1630	72	4.42%	335	4.65

Table 1.2: Finance Descriptives

Year	Firms	Debt	% Debt	Equity	% Equity	II >0	% II >0
2004	3140	193	6.15%	144	4.59%	1149	36.6%
2005	2837	264	9.31%	187	6.59%	1555	54.8%
2006	2554	290	11.4%	189	7.40%	1564	61.2%
2007	2330	307	13.2%	180	7.73%	1379	59.2%
2008	2092	303	14.5%	165	7.89%	1206	57.6%
2009	1928	307	15.9%	152	7.88%	1152	59.8%
2010	1775	310	17.5%	146	8.23%	1095	61.7%
2011	1630	292	17.9%	140	8.59%	1092	67.0%

business. The patent holders column includes those firms that acquire a first patent in a survey year. The number of patents per patent holder indicates that the sample includes firms with multiple patents. Since the ratio remains relatively stable, patent ownership does not become dominated by a couple firms.

In Table 3.2 we provide statistics regarding financing and profitability through time. Note that external financing is relatively rare in this sample of representative firms, but among financing sources, debt is more common. A large number of firms report positive profits in each year but this table does a good job showing the effect of the recession on firm health. By 2010 and 2011, however, firms seem to stabilize.

Table 2.3 below shows how patenting firms differ from non-patenting firms in terms of financing source rates. A patenting firm is much more likely to source external

Table 1.3: Patenting Firms vs Non-Patenting Firms Financing

Year	%Pat & Debt	% No Pat & Debt	%Pat & Equity	% No Pat& Equity
2004	13.3%	5.86%	19.2%	4.01%
2005	16.9%	8.96%	24.2%	5.79%
2006	25.4%	10.6%	32.8%	6.13%
2007	29.2%	12.4%	32.1%	6.56%
2008	30.2%	13.8%	30.2%	6.93%
2009	43.9%	14.7%	35.4%	6.66%
2010	47.2%	16.2%	30.6%	7.28%
2011	36.1%	17.1%	30.6%	7.57%

financing whether the source is debt or equity. In the next section we investigate if this observed relationship is due to other factors or is driven by the patent ownership.

Table 1.4 shows some of the correlations that patenting has with the outcome variables. The two survival columns indicate the percentage of firms who belong to either the patenting or non-patenting group in the year shown that operate in the next year. Year 2011 has NA values since we do not know whether a firm continues since that is the last year of the survey. The last two columns indicate the percentage of firms who report a positive profit. Most years reveal a lower survival rate and all years show a lower percentage of patenting firms with positive profits. This contrast is even more stark given the much higher percentage of patenting firms receiving external financing.

1.5.2 H_1 : The Effect of Patenting on Profits

This section opens with regression analysis using both a POLS (pooled ordinary least squares) and FE (fixed effects) approach. The two models are at odds. The POLS model suggests that patents are associated with lower profits but the FE model suggests no statistically significant association between patenting and profits. We argue that POLS model, unlike the FE model, is affected by omitted variables. The lack of significance in the FE model is perhaps due to the extreme heterogeneity found

Table 1.4: Conditional Outcomes

Year	%Survive Patent	%Survive No Patent	% $\Pi > 0$ Patent	% $\Pi > 0$ No Patent
2004	94.2%	90.2%	26.7%	37.0%
2005	88.7%	90.1%	28.2%	56.0%
2006	88.5%	91.4%	32.0%	62.7%
2007	84.9%	90.0%	42.5%	60.0%
2008	97.7%	90.9%	40.7%	58.4%
2009	91.5%	92.1%	46.3%	60.3%
2010	95.8%	91.7%	44.4%	62.4%
2011	NA*	NA*	51.4%	67.7%

* The Kauffman Survey has no follow ups after 2011 so we don't know if a firm continued operations into 2012 or ceased at the end of 2011.

in the data. To further explore the results, a Oaxaca decomposition is performed on the POLS results. This model is used to identify the variables and responses most associated with negative profits. These results are interpreted in light of the omitted variables.

The regression results are presented in Table 1.5. The first column is OLS from a cross-section from 2011, the second is OLS on the full panel with time dummies, and the third is a two-way fixed effects regression. The cross-section regression restricts the sample to those firms able to survive through the final year of the survey. The cross-section provides a view of the data without the influence of firms that drop out before the completion of the survey. The difference between the patent coefficients in columns two and three are the primary concern and discussed below.

The estimate on patent in column two, the POLS model, indicates that patents are associated with lower profits. The coefficient suggests that firms that do not patent are associated with 300% higher profits than patenting firms, all else equal. This estimate is susceptible to unobservable influences, factors correlated with patenting and captured in the residual. The estimate in the third column removes any time-invariant uncaptured factors related to the firm. This estimate is statistically significant, however, suggesting that the strong negative correlation picked up by POLS is due to

Table 1.5: Dependent Variable: Log Transform of Profits

	2011 C-S	Panel Time Dummy ⁴	Two-way FE
Patent	-3.332***	-3.453***	-.757
Ln(External Eq)	-.264***	-.323***	-.074
Ln(Other Fin)	-.126***	-.202***	-.252***
Employees	.043**	.037***	.077***
Ln(Assets)	.568***	.409***	.296***
LLC ¹	.342	.266	2.187***
S-Corp ¹	.508	-.067	2.509***
C-Corp ¹	-1.711	-1.684***	1.141
Credit Risk	-.521*	-.631***	-.162
Experience	.019	.033***	NA
Hours	.031**	.012	.020***
Emp-Owner	2.432***	3.041***	1.082***
Hours:Emp-Owner	-.014	-.015*	-.002
Gender:Male	.878*	.645***	6.563***
Same Ind:No ²	-1.023*	-.599***	NA
Same Ind:Yes ²	-.593	-.339***	NA
<Bachelors ³	-.118	-.119	.437
≥Bachelors ³	.065	.321	.940
Intercept	-2.110	-4.244***	NA
R^2	.132	.123	.021
\tilde{R}^2	.120	.122	-.194
Observations	1416	16,245	16,245
F-test	11.760***	91.231***	19.195***
Note:	*p<0.1; **p<0.05; ***p<0.01		

¹ The baseline Legal Structure is sole proprietorship.

² The baseline for Same Industry is first startup.

³ The baseline for Education is highschool.

⁴ Time dummies excluded for space considerations.

other factors.

Following Wooldridge (2010), a Hausman test confirms the fixed effects specification against a random effects model. The fixed effects model is intuitively appealing as well. The fixed effects that are removed via this modeling procedure include all firm invariant factors. These include entrepreneurial ability, business savvy, and risk preference as these traits are not captured in POLS, correlated with the right and left side of the regression equation, and are time-invariant. A panel F test confirms the use of time dummies, which are included in both the POLS and FE models. Additionally, the data likely suffers from heteroskedasticity and serial correlation, so clustered robust standard errors with the finite sample correction suggested by MacKinnon and White (1985) are used in place of the standard computations.

Oaxaca Decomposition

The Oaxaca decomposition allows one to investigate whether the patenting and non-patenting groups differ with regard to covariate response, balance, or interaction. The idea behind the Blinder-Oaxaca decomposition (Oaxaca 1973; Blinder 1973) is to investigate differences in an outcome variable between two groups. The unconditional average difference in firm profits is large since firms without patents have 454% greater profits than patenting firms. The decomposition portions out this difference into what can be explained by average endowment of the independent variables and what can be explained by average response to those independent variables.

Normally, the Oaxaca decomposition is performed on time-invariant groups, like race or gender. This is a novel application of the procedure because patent ownership is only partially time-invariant; once a patent is obtained, it remains in ownership of the firm. With the group composition changing through time, covariate balance with pre-patented firms being accounted for in the non-patenting group is potentially an issue. Otherwise, using the decomposition on time-varying groups does not present

any issues.

The Oaxaca decomposition is performed on the pooled data rather than centered data as in a fixed effects regression. This is done since interest is between the groups rather than within the firm. This opens the estimation up to influence from bias but this is a necessary trade-off to get desired group differences. Bias due to economic shocks is mitigated via year dummy variables. Furthermore, the value of a patent may be inextricable from firm/owner specific unobservable variables. That means that patents and their creator may be complementary goods, the patent losing at least some of its value if separated from its owner/creator.

This section presents the main decomposition results. See Appendix A for the full mathematical decomposition and results. While the regression finds a conditional 300% difference between patenting and non-patenting firm profits, the unconditional difference is 4.54 log units or 454%, shown in Eqn. 1.1.

A threefold decomposition (Winsborough and Dickinson 1971; Jones and Kelley 1984) portions out that difference into three sources, endowments, coefficients, and interaction. This is shown (although not derived) in Eqn. 1.2. The portion attributed to endowments is due to differences in the distribution of control variables. The portion attributed to coefficients is due to differences in coefficient values in two separate regressions, one for patent owners and one for firms without patents. The portion attributed to interaction is due to the simultaneity of both effects.

$$\begin{aligned}\Delta\bar{Y} &= \bar{Y}_{NP} - \bar{Y}_P, \\ &= 2.14 - (-2.39), \\ &= 4.54\end{aligned}\tag{1.1}$$

$$\Delta\bar{Y} = \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' \hat{\beta}_P}_{\text{endowments}} + \underbrace{\bar{\mathbf{X}}_P' (\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{coefficients}} + \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' (\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{interaction}} \quad (1.2)$$

The endowment effect accounts for a statistically significant 2.08 units of the 4.54 difference or almost half. The coefficient effect accounts for a statistically significant 3.89 units of the difference. The interaction effect accounts for -1.43 units of the difference.

The greatest contribution toward the endowment effect is equity, which contributes 2.2 of the 2.08 total, statistically significant at the 1% level. Equity is a great contributor to the coefficient effect as well, contributing 1.75 of the 3.89 total, significant at the 1% level. Equity contributes for all of the -1.43 difference attributed to the interaction effect, significant at the 1% level. The full contribution of equity is 2.52 of the 4.54 units. There are other factors at play, but the amount of equity that patenting firms receive seems to play a large part in the lower profit observed in patent owners.

1.5.3 H_2 : Adverse Selection

An important question is raised through this analysis. Why do patents attract equity investors when they are correlated with lower profits and at best show no advantage on average? The pecking order hypothesis of firm financing states that equity is the least desirable source of financing. It may be that patenting firms are adversely selected by equity firms (or rather, firm managers are adversely seeking equity) since stronger firms find more advantageous sources of financing.

We show in this section that this argument does not hold up in this data. Both patenting and non-patenting firms pursue roughly the same balance of financing sources. Furthermore, patenting firms are not (on paper) riskier assets to investment firms. They are not significantly different in terms of credit, industry experience,

education, age, and prior startup experience.

$$\begin{aligned}\Delta\bar{Y} &= \bar{Y}_{NP} - \bar{Y}_P, \\ &= .68 - 3.79, \\ &= -3.11\end{aligned}\tag{1.3}$$

Adverse selection suggests that 'good' investments are difficult to find because any owner of a 'good' firm does not want to relinquish control. They have opportunities to raise *cheaper* money since equity comes with more baggage. The only available firms for equity investors are 'bad' firms which are 'bad' because of management, experience, product, or market conditions. It is possible that the attributes associated with patent development are not well-suited to running a business. An alternative explanation is that R&D competes for resources against business growing activities. A Oaxaca decomposition will aid in investigating the role of adverse selection.

As before, the variables included in this Oaxaca decomposition are log profits, log other financing, employees, log assets, credit, and year. This decomposition also includes hours worked by owner, work experience within industry, and legal status. The threefold decomposition results are presented in Eqn. 1.4. Since $\Delta\bar{Y}$ is negative, the endowment and coefficient effects suggest that these two effects contribute to the overall difference, and that non-patenting firms received less equity financing.

$$\begin{aligned}
\Delta \bar{Y} = & \\
& \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' \hat{\beta}_P}_{\text{endowments}} = -2.73^{***}(.186) + \\
& \underbrace{\bar{\mathbf{X}}_P'(\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{coefficients}} = -2.11^{***}(.209) + \\
& \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)'(\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{interaction}} = 1.73^{***}(.166)
\end{aligned} \tag{1.4}$$

The detailed results of the decomposition are in Table 1.6. The column of values represents the mean difference in endowments, $\Delta \bar{X} = \bar{X}_{NP} - \bar{X}_P$. The variable most directly linked to risk preference is credit risk, but the results for that variable do not present any cause for patenting firms pursuing greater equity financing than non-patenting firms. The equity difference is also not caused by a lack of experience or dedication, as measured in average hours worked. The opposite seems to be true. The greater experience of patenting firms is related to some of the equity advantage.

It is interesting that other financing contributes so much to observed differences in equity financing between patenting and non-patenting firms. The endowment effect suggests the presence of the pecking order hypothesis if the firm exhausted personal and debt sources before acquiring equity financing. But, this doesn't rule out other interpretations, such as entrepreneurial orientation (Vaznyte and Andries 2019). The coefficient effect, since this informs us that the non-patenting firm's coefficient is less than patenting firms', says that if that non-patenting group were endowed with as much other financing, they would not acquire as much equity financing. This could suggest they lack the ability to attract equity but it more likely suggests there is a fundamental difference in the approach to growing a business. The difference in employees and assets seems to support this since patenting firms have more of both on average.

The difference in legal form of the firm also explains some of the equity difference. The coefficient effect on both LLC and S-Corp drive the equity difference down. Non-patenting firms that form as an LLC or S-Corp are better able to attract equity than patenting firms in those formations. That reverses for non-patenting firms that form as a C-Corp where they are less likely to attract equity compared with patenting C-Corps. The relative abundance of patenting firms in the C-Corp formation might attract equity investors to C-Corps to find patenting firms and to other formations for non-patenting firms. In terms of signal, legal formation might be the most significant.

Table 1.6: Oaxaca Decomposition on Ln(External Equity)

	$\Delta\bar{X}$	Endowment	Coefficient	Interaction
Ln(Profit)	4.54	-.77***	-.36***	.68***
Ln(Other Fin)	-3.56	-1.06***	-1.21***	.55***
Employees	-3.73	-.14**	-.06	.03
Ln(Assets)	-1.26	-.32***	-2.22***	.26***
Credit Risk	.09	-.02	-.57	.02
LLC ¹	.02	-.02	.32***	.02
S-Corp ¹	-.01	.01	.48***	-.01
C-Corp ¹	-.23	-.43***	-.26*	.20*
Experience	-1.94	.05*	.51**	-.06**
Hours	-5.13	-.08**	-.69**	.08**
Year	-.03	-.004	.47	.004
Intercept	0.0	0.0	2.42***	0.0
Observations		16,245	16,245	16,245
Note:			*p<0.1; **p<0.05; ***p<0.01	

¹ The baseline Legal Structure is sole proprietorship.

1.6 Conclusion

This article takes a thorough look at the relationship between patenting and nascent firm outcomes. The correlations present—that patents are associated with worse profits and no better survival—defy conventional wisdom that patenting is an advantage. This negative association persists in regression analysis until the use of individual

fixed effects removes the bias of unobservable factors. These fixed effects estimates, supported by diagnostic testing, imply that patenting is not causally associated with lower profits. To address this puzzle, the Oaxaca decomposition strategy allowed us to assess if patenting firms are fundamentally different from non-patenting firms in ways that cause the observed profit disadvantage.

The Oaxaca decomposition finds that the lower profits experienced by patent owners is likely related to obtaining too much equity financing. A further decomposition conducted on a model that has equity financing as the outcome finds that patent owners may over-finance across the board, putting too much pressure on young firms.

The use of a threefold Oaxaca decomposition is novel in this area of research and is almost exclusively applied to cases where group membership is permanent as in race or gender. There is nothing mathematically or intuitively preventing its adoption in a study with a self-selected treatment. The decomposition applies well when a treatment is likely to not only affect the outcome variable but a number of control variables as well.

Finally, this article finds that while patenting does not lower profits, neither are lower profits due to adverse selection. Instead, patenting and non-patenting firms have similar risk profiles and seek equity financing at similar rates. Poor performance appears to not be due to patenting, but in spite of it.

CHAPTER 2

SIGNAL VS APPROPRIATION VALUE OF PATENTS

with Michael Risch
Vice Dean, Villanova Law

2.1 Introduction

In the quest for empirical verification of the role of patents in entrepreneurial firm success, one stubborn question is the role patents play in venture funding. A small but growing body of literature has causally linked patenting with increased venture capital investment in startups, but the reason why remains hotly disputed.

The two leading explanations behind increased venture capital investment in patenting firms are exclusivity and signaling. Patents, by their very nature, provide exclusivity. The owner has the right to control use of the patented invention. As a result, a nascent firm with one or more patents may be more likely to have revenues and thus attract more funding. This is the appropriation value of patents. When firms seek funding, however, adverse selection, moral hazard and pecking order theories predict that venture-funded firms will fail more often. Patenting can be used to solve asymmetric information problems, allowing venture capitalists to choose only the best ventures, even if the patent itself provides no appropriative value. This is the signal value of patents.

These explanations are not mutually exclusive. A patent can provide both signals and extra-competitive returns, and both features may be valued by investors. Indeed, most agree that patents provide at least *some* signal, even if the magnitude is unmeasured. On the other hand, many simply assume that patents provide no appropriative value at all. Chen, Meng, and Lu (2018) and Comino and Graziano (2015), for example, each develop models in which signals are the only reason firms might invest. To date, little work has been done to untangle the two.

We add to the literature in two ways using a refined, novel dataset: the Kauffman Firm Survey of more than 3000 nascent firms. First, we provide further confirmation that patents increase venture funding. The longitudinal results allow for a strong, if not causal, inference. The Kauffman Firm Survey collected data on firm experience, resources, and other factors that might influence success. Using these data in a panel model shows that owning patents increases chances of receiving funding by about 13%, holding firm founder, risk, and industry variables constant.

Second, we use a novel decomposition method to disaggregate the signaling and appropriative value of patents in venture funding. Our method compares fixed effects and POLS (panel ordinary least squares) models. Comparing the coefficient on patent in both models allows us to tease out the relative strength of the signal and appropriation effects. This is possible because of the interpretation of the fixed effect (time-invariant firm specific) as the signal effect.

The results are initially startling: among all firms, the patenting coefficient in the fixed-effects model is statistically indistinguishable from zero, implying that there is no appropriative value to patents, and that most, if not all, of the increased likelihood of funding is due to the signaling effect. However, we find substantial heterogeneity in the data, but not where one would expect it. When analyzed by industry, the results are essentially the same. However, when analyzed by firm structure, the results reveal that one type of structure, the general partnership, is so unlikely to obtain investment

while patenting that it dominated all other types.

This article proceeds as follows. We begin with an overview of startups, IP's role in financing, and signaling v. appropriation. We then report on prior findings that tie IP and venture funding (or both) to firm outcomes. This leads to two testable hypotheses, which we outline. We then discuss our data, model and estimates used to test our hypotheses. The article concludes with a discussion of the implications.

2.2 Background and Literature Review

2.2.1 Nascent Ventures, IP, and Funding

A number of studies reference access to capital as an important growth constraint, for example Beck and Demirguc-Kunt (2006), who find that an increased access to external financing helps level the playing field among firms of different sizes. Paglia and Harjoto (2014) focus on private equity and venture capital and show that access to such sources is very restricted to small and medium sized firms and those who do attain it benefit along a number of growth metrics.

Intellectual property plays an important role in how firms overcome the adverse selection problem in a dynamic system. One of the characteristics of early stage firms is the type of intellectual property they hold. In the U.S. and just about everywhere else, companies use four primary types of IP to appropriate value. Patents grant the right to exclude others from using claimed inventions, whether others copied them or developed them independently. Copyrights protect expression, from books to music to computer software, from being copied by others. Trade secrets protect all manner of information from the prying eyes of others-primarily ex-employees. Trademarks protect brand names from potentially confusing uses.

Where VC firms are able to mitigate information asymmetries, their target firms are more likely to invest and receive investment (Lockett, Murray, and Wright 2002;

Cumming and Johan 2008). Thus, venture capitalists usually require some method of appropriation before investing. Intellectual property's appropriative features often satisfy this requirement, but other methods might be acceptable, such as first mover advantage coupled with market share. For this reason, the role of patenting in venture funding is well studied. Because trademarks offer little appropriative value for an entrepreneurial product and every company has some trade secret, this article (like many others) focuses primarily on patents.¹

Patents can reduce information asymmetry (Hottenrott, Hall, and Czarnitzki 2016; Gartner, Frid, and Alexander 2012; Long 2002). For example, one study found that only those Italian firms surpassing minimal financial constraints engaged in R&D and patenting (Scellato 2007).

Thus, IP, and particularly patent protection, has long been considered part and parcel with the entrepreneurial experience (Goel and Goktepe-Hulten 2013). But the role of patents in entrepreneurship has been contested for nearly as long. Some believe that patents are essential for the formation of nascent companies, whether the companies seek outside investment or not. Others argue that patents block entry and needlessly raise prices. Cockburn and MacGarvie (2011) highlight this tension, finding that patents hinder entry in the software industry...unless one has a patent. More general studies have found that, with a few exceptions, patents are more valuable for entrants (Hall 2004) but can also have value for established firms (Noel and Schankerman 2013).

Notwithstanding this tension in academia, entrepreneurs themselves report the desire to seek patents. Other studies have confirmed that a substantial number of startups begin with patents or other innovation (Nerkar and Shane 2003; Acs and Audretsch 1989; Goniadis and Varsakelis 2012; Shane, Locke, and Collins 2003), and our study finds the same. Startups report seeking patents for their appropriative value

¹In later work we may consider copyrights as well, though-like trademarks-many companies own copyrights but cannot use them to a market advantage (for example, user manuals).

as the primary motivation, and then for their ability to attract investment (Graham and Sichelman 2010; Rassenfosse 2012). Determining whether either or both of these value propositions holds true in practice is another story, of course. Furthermore, despite a widespread interest in patenting, most firms never seek a patent.

Based on the theoretical framework—venture capitalists seeking appropriation and information—IP should aid in attracting financing (Tondo-Kramer 2009; Fabrizi et al. 2013). Though some surveys of venture capitalists have shown that appropriability is not a key reason for investment (Macmillan, Siegel, and Narasimha 1985), other surveys have shown the opposite (Knockaert, Clarysse, and Lockett 2010; Hsu et al. 2014). Regardless of stated preferences, a small but consistent revealed preference literature has shown a link between patenting and venture investment.

2.2.2 Patenting Effects on Investment

Farre-Mensa, Hegde, and Ljungqvist (2017) study the impacts of patents on U.S. firm outcomes including survival, revenue generation, and capital formation. Through the use of instrumental variables, they find the most significant connection with the ability of patents to generate capital. Aside from this study, however, there are very few studies of United States companies that isolate the role of patents in obtaining venture funding.

There are, of course, studies that find positive results outside the U.S. Some studies find increased likelihood of investment after patenting for foreign firms. Haeussler, Harhoff, and Mueller (2014) find that select German and British firms had increased investing after filing patent applications. Engel and Keilbach (2007) get similar results with German startups. Baum and Silverman (2004) find that patents improve pre-IPO financing among Canadian firms. Conti, Thursby, and Thursby (2013) develop a model of endogenous patenting and investment and find that patents are useful signals among Israeli startups. However, foreign experience may be different; stud-

ies of European patenting have not replicated findings of increased venture funding based on patenting (Bertoni, Croce, and D’Adda 2010). While international results are illuminating and helpful, they may not be generalizable to the United States (Rassenfosse 2012; Scellato 2007).

Several studies use data from only those firms that received investment, and compare the amount. In a seminal study, Lerner (1994) associates broader patent scope with additional funding among biotech firms that had already received funding. Several other studies show similar results (Cao and Hsu 2011; Conti, Thursby, and Thursby 2013; Hoenen et al. 2014). Mann and Sager (2007) examine the differences in followup financing rounds among software and biotech firms that had already received funding. Lahr and Mina (2016) use survey data on U.S. and U.K. firms that sought financing during a two-year period to find that those patenting were more likely to receive venture capital. Guo, Lou, and Pérez-Castrillo (2015) find that patents have no effect on firm survival among firms that have already received financing. Engel (2004) obtained the opposite results, that among a subset of German firms receiving VC funding, patents had a slight positive effect on survival.

Other studies select companies at different stages of their life, or at different times, so that comparisons are not possible. Lerman (2015) studied startups over a four-year period from the Crunchbase database, finding that patenting firms were more likely to receive venture capital and that more patents led to more investment. Hottenrott, Hall, and Czarnitzki (2016) consider heterogeneous established firms. Conti, Thursby, and Rothaermel (2013) find that incubator-based firms formed over a period of years from a single university were much more likely to obtain venture capital funding upon receiving patents. Hsu David H. and Ziedonis Rosemarie H. (2013) find that patents increased financing for those semiconductor firms formed over a 25-year period that received at least one round of financing. Shane and Khurana (2003) find that management experience improves the chances that patents granted

over a 16-year period will be commercialized. Hoenen et al. (2014) find some effect in the first round of financing for 10 years of biotech startups. Munari and Toschi (2015) find that patents associated with core technologies improved investment chances for nanotechnology companies over a 21-year period. Balasubramanian and Sivadasan (2011) used census data covering 22 years. Finally, in a pecking order article Mina, Lahr, and Hughes (2013) examine a cross-section of firm financing structures and find that innovation did not change the demand for particular types of financing, but did improve the likelihood of obtaining financing.

2.2.3 Appropriation vs. Signaling

Given the apparently positive effect of patents on investment, one key area of inquiry is separating the appropriative value of patents from the signaling value. While some view signaling as a potential normative good in its own right (Long 2002; Mazzoleni and Nelson 1998; Hahn, Kim, and Kwon 2017), most seek other welfare gains from any deadweight loss associated with the right to exclude. In other words, if signaling is not a good in itself, then skeptics believe patenting that fails to provide appropriative benefit does not justify the costs of blocking entry.

In a recent literature review, Hall (2018) argues that disaggregating patent value into appropriation and signaling is a key unanswered question in the literature. Some have argued that firm level data cannot separate signaling from appropriation (Hoenig and Henkel 2015). Without this disaggregation, most tend to believe that the value must simply be a signal of the innovative prowess of the firm rather than an exclusionary benefit in its own right (Backes-Gellner and Werner 2007; Munari, Odasso, and Toschi 2011).

Farre-Mensa, Hegde, and Ljungqvist (2017) use random assignment to patent examiners as an instrumental variable to determine the effect of patenting. They find substantial increases in investment the earlier a patent is processed, regardless of

patent quality. This implies that appropriability, not quality signaling, is important because many of the patents were assigned to more lenient examiners. In other words, if even the weaker patents improved outcomes, then it was the patent itself rather than the technology behind the patent that drove the value. We agree with Hall that this study is the most useful to date. But it is not the only study to attempt to disaggregate patent value.

Haeussler, Harhoff, and Mueller (2014) find that the likelihood of investment increases from the time an application is filed and during prosecution as investors learn more about the patent. However, this can be consistent with either signaling or with the appropriative value of the patent. Cao and Hsu (2011) use instrumental variables related to industry patenting. Their study focuses only on firms receiving investment so does not allow for testing of patents as signals in obtaining initial investment or in firm outcomes for nascent firms. Similarly, Hsu David H. and Ziedonis Rosemarie H. (2013) find support for signaling among semiconductor firms, as patents had a more pronounced effect on firms in early rounds, with inexperienced management, or going IPO without prominent VCs. Surveys of small and medium sized firms point to similar results (Rassenfosse 2012). The previously mentioned studies support information asymmetry, but this study also did not examine companies with patents but no investment. Hoenen et al. (2014) found similar results among biotech firms, with new patenting having no effect on later rounds of financing (implying signaling value only). However, there could be many reasons why patents would have diminished value, such as the appropriability of early patents as well as established market or FDA position.

Hottenrott, Hall, and Czarnitzki (2016) examine established firms, and find that patents result in more investment for smaller, more financially constrained firms than larger firms. This implies that the signaling is more important than appropriability, at least for firms of this age, which is consistent with surveys (Graham and Sichelman

2010).

Conti, Thursby, and Rothaermel (2013) treats patents and early friends and family financing as endogenous in a model predicting venture capital and business angel financing. Using their endogeneity model, they find that friends/family money is more important to business angels and that patents are more important to venture capitalists. The implications for signaling are ambiguous. On the one hand, each of these may be signals in their own right, with different types of inventors preferring different signals. On the other hand, it implies that friends and family money is a sufficient signal for investment, and that patenting is further valued because of the appropriative functions. The authors conclude that patents are a signal of the "probability of success of the technology," but it is unclear that their model allows for determination that it is technology as signal or technology as appropriated that drives the investment decision.

In a related study, Rassenfosse, Palangkaraya, and Webster (2016) consider disclosure versus appropriability for patent licensing. They compare transactions with patents, patent applications, and no patent applications. They find the appropriation effect is stronger in technologies where patents are considered more important for exclusivity, like biotech. They also find that granted patents lead to better licenses than patent applications, implying that the appropriation effect is more important than the disclosure effect. This study is not directly on point, however; a direct patent license is not the same as an investment in a firm. Thus, a patent and patent application may both provide a signal of future value to investors in cases where a licensor could not obtain that value. More important, patent applications are simply patents that have not been granted yet, making it difficult to separate current from future sources of value.

Finally, some are skeptical that patents are mostly signals. Experience of the startup team is observable (Backes-Gellner and Werner 2007), and corporate disclo-

sure rules would penalize false statements about the progress of development (Hemel and Ouellette 2013). Consider, for example, Audretsch, Bönte, and Mahagaonkar (2012), which argues that appropriability is signalled through patents and prototypes only signal feasibility. Similarly, Hoenig and Henkel (2015) use conjoint analysis to find that R&D alliances and team experience were effective quality signals, but that patents were not.

2.3 Testable Hypotheses

The theoretical background and empirical landscape yields two testable hypotheses that remain mostly unanswered in the literature:

H₁: Patents as Verifiable Information: Patents have no effect on nascent firm venture funding

H₂: Patents as Signals: Patents have no appropriative value

Our approach to these hypotheses fills an important gap in the literature: isolating the signaling and appropriation value of patents.

2.4 Data and Variables

The Kauffman Firm Survey is unique micro-panel of for-profit, non-subsiary, non-inherited businesses formed in 2004. There were 4,928 firms at the start of the survey, but due to attrition, only 2,032 firms survived for the eight-year length of the survey.

The survey over-samples firms classified as high or medium-tech based on the number of employees engaged in research and development at the industry level. Roughly 13% of firms in the first sample year are considered to be hightech². Despite the oversampling, the Kauffman Survey sought to be representative of the full U.S.

²Hightech is a KFS classification based on industry average for research and development employment.

economy, and it includes sophisticated weighting data to remain representative. The survey represents 25 two-digit industry codes (industry sector) and 85 three-digit industry codes (industry subsectors) by the North American Industry Classification System (NAICS).³

Due to attrition, as well as a high non-response rate, the Kauffman Foundation provides a few versions of the data. We use the multiply imputed longitudinal set, which removes firms that drop out and reappear, leaving only the firms that respond each year of the survey while in business. The longitudinal sample has 3,140 firms at baseline and 1,630 in the final year.

The imputed data offers the advantage of logical imputation which is a procedure to fill in missing responses to questions previously answered that will not change through time. For example, the survey authors can logically impute an owner's race if that information was previously supplied. Despite this effort there are many holes in the data. The missing values that could not be logically imputed were split into hard and soft missing data following Stata convention. Soft missing values can reasonably be assumed based on question type and response but these imputations convey uncertainty through multiple imputation computations. Hard missing values, those that cannot be assumed anything with any certainty remain empty.

2.4.1 Dependent Variable

The dependent variable used is a measure of external equity financing that includes venture and angel investment received in a given year. Since this variable includes so many zero observations and very large amounts when obtained, a binary indicator of obtaining this equity is used. This transformation turns the regression models into linear probability models.

³The NAICS replaced the Standard Industrial Classification (SIC) in 1997.

2.4.2 Key Explanatory Variable

The key explanatory variable is a patenting dummy. Although the survey provides a count measure as well, we consider only the binary indicator. The count variable does not add to the analysis or change the results in any meaningful way. Furthermore, use of a binary indicator is supported by the literature, which finds that the crucial distinction is patent versus no patent, not the number of patents (Mann and Sager 2007; Lahr and Mina 2016).

The survey does not provide any information on when the patent applications were filed, any detailed information about purchased or in-house patents, or any information from the Patent and Trademark Office. This lack of information shapes the analysis undertaken in the paper, preventing any instrumental variable regression.

2.4.3 Control Variables

In addition to the key explanatory variable, the regression models control for a number of additional factors that are likely correlated with both the outcome and key explanatory variable.

Owner Characteristics

The owner characteristics are gender, race, education, experience in this industry, count of previous startups, and credit risk. The survey includes a number of additional variables, but we use only variables of theoretical and intuitive interest. Gender, race, and education are categorical variables, while experience and startups are continuous variables. We code the credit risk variable, being discretely valued and ordinal, as categorical because it does not display a linear relationship with either outcome variable.

Gender is an important control variable for two reasons because it captures motivation (Burke, FitzRoy, and Nolan 2002) and firm management strategies. Gender

also has demand- and supply-side implications in the equity market, including risk preference (Charness and Gneezy 2012) and reliance on external sources of financing (Coleman and Robb 2009).

The survey provides an education of owners variable with ten levels, which we consolidated into four to increase explanatory power: high school degree or less, technical/trade/vocational, some college/associate's, four year college degree or more. Although we control for industry sector, an owner's education level is correlated with more granular industry indicators, attractiveness to debt and equity financiers, and, likely, with other unobservable measures of business acumen.

The survey provides both the owners' ages and their experience levels. To avoid collinearity we use experience since it likely captures more of the variation in the dependent variable. Experience is measured in years within the industry in which the startup was formed. Experience not only captures industry knowledge and access to professional networks but also attractiveness to investors.

The KFS provides a the number of previous startups to capture entrepreneurial experience. This is a count variable that ranges from zero to five. The survey provides an additional dummy variable asking if any of the previous were in the same industry as the current startup to refine the count measure. We use this variable as well (called "Same Ind") with three levels: First Startup; No; Yes. These variables convey another rich measure of experience and can act as a strong signal to potential investors and lenders. Therefore the expected sign is positive. On the contrary, serial entrepreneurs would repel equity investment. The Kauffman Survey does not provide any indication of the success of these previous ventures.

We have access to two measures of credit risk. The first is the D&B commercial credit rating (Credit Risk), which ranges from one, the highest decile of credit ratings, to five, the lowest decile of credit rating, with scores of two, three, and four splitting the intermediate 80 percent. The Kauffman Survey also provides the D&B

financial stress score (FSSP) measured from one to one hundred with higher numbers representing decreasing likelihood of financial stress. We focus on the D&B credit risk variable due to high collinearity. We use this measure as a good proxy for financial risk. It is possible that financial risk correlates with other risk taking behavior.

Firm Characteristics

Control variables related to firm characteristics are other sources of financing, profits, assets, employees, legal status, and NAICS two-digit code. Other financing, profits, assets, and employees are continuous measures. Legal status and NAICS are categorical variables.

Other sources of financing, profits, assets, and employees help control the need for external equity financing. The impact of assets is likely to be negative because assets are strong signals to bank lenders. The impact of other financing, however, is ambiguous, because other financing is correlated with firm size, an important indicator of need (Cassar 2004). The pecking order hypothesis suggests that other financing should have a negative sign since equity is not desirable, but some entrepreneurs might welcome the opportunity to share the risk, increase their network, and glean insights from experienced investors.

Legal status (Sole Proprietorship, LLC, Subchapter S-Corporation, C-Corporation, General Partnership, and Limited Partnership, and Something else?) control for a number of factors and is quite important, affecting the magnitude and significance of the other variables. Organizing as a corporation rather than a sole proprietorship, LLC, or partnership has implications for liability and taxes. Corporations protect owners from business liability at the expense of additional taxes, paperwork, and legal requirements. The legal structure of the business can act as a strong signal of intent, preparation, and motivation to potential investors. Failure to organize in a particular way sends a clear signal about expected financing.

2.4.4 Variable Transformations

The continuous variables of financing, profit, and assets have extremely high variances even though most observations (5th to 95th quantile) are quite close to the mean. To alleviate the impact that outliers have on estimation, the natural logarithm was taken. One was added to each due to zeros in the data. The range of profits extended far into the negative number line as well so an additional step was taken. The natural log was applied to the absolute value of profit before reapplying the sign.

2.5 Analysis

2.5.1 Descriptive Statistics

In this section we present descriptive statistics on survival, patenting, profits, and financing through time.

Table 2.1 presents preliminary trends of attrition and patent ownership. Although this survey oversamples high technology firms, the patenting rate is still quite low. The rate of attrition can be observed in the change in number of firms. Patenting is not insurance from going out of business. The downward trend in the number of patent holders indicates that many early holders went out of business. This column also includes those firms that acquire patents throughout the survey. The number of patents per patent holders indicates that the sample includes firms with multiple patents and since the ratio remains relatively stable patent ownership does not become dominated by a couple firms.

Table 2.2 presents statistics regarding financing and profitability through time. Note that external financing is relatively rare in this sample of representative firms but among financing sources. A large number of firms report positive profits in each year, but this table does a good job showing the effect of the recession on firm health.

Table 2.1: Patent Descriptives

Year	Firms	Patent Holders	Holder/Firms	Patents	Patents/Holder
2004	3140	120	3.82%	584	4.87
2005	2837	124	4.37%	608	4.9
2006	2554	122	4.78%	487	3.99
2007	2330	106	4.55%	556	5.25
2008	2092	86	4.11%	622	7.23
2009	1928	82	4.25%	386	4.71
2010	1775	72	4.06%	355	4.93
2011	1630	72	4.42%	335	4.65

Table 2.2: Finance Descriptives

Year	Firms	Debt	% Debt	Equity	% Equity	II >0	% II >0
2004	3140	193	6.15%	144	4.59%	1149	36.6%
2005	2837	264	9.31%	187	6.59%	1555	54.8%
2006	2554	290	11.4%	189	7.40%	1564	61.2%
2007	2330	307	13.2%	180	7.73%	1379	59.2%
2008	2092	303	14.5%	165	7.89%	1206	57.6%
2009	1928	307	15.9%	152	7.88%	1152	59.8%
2010	1775	310	17.5%	146	8.23%	1095	61.7%
2011	1630	292	17.9%	140	8.59%	1092	67.0%

By 2010 and 2011, however, the number of firms stabilizes.

Table 2.3 shows how patenting firms differ from non-patenting firms in terms of financing source rates. A patenting firm is much more likely to source external financing whether the source is debt or equity. In the next section we investigate if this observed relationship is due to other factors or is driven by the patent ownership.

2.5.2 Patenting and Investment

This section shows the effect of patenting on investment in two models, a panel OLS (POLS) regression and a fixed effects (FE) regression. These models provide the framework to disaggregate observed patent value into signal and appropriative components. Signaling refers to observable firm traits transmitting unobservable in-

Table 2.3: Patenting Firms vs Non-Patenting Firms Financing

Year	%Pat & Debt	% No Pat & Debt	%Pat & Equity	% No Pat& Equity
2004	13.3%	5.86%	19.2%	4.01%
2005	16.9%	8.96%	24.2%	5.79%
2006	25.4%	10.6%	32.8%	6.13%
2007	29.2%	12.4%	32.1%	6.56%
2008	30.2%	13.8%	30.2%	6.93%
2009	43.9%	14.7%	35.4%	6.66%
2010	47.2%	16.2%	30.6%	7.28%
2011	36.1%	17.1%	30.6%	7.57%

Table 2.4: Legal Status vs Patent Ownership in 2011

Legal Status	No Patent	Patent
Sole Proprietorship	476	8
Limited Liability Company	550	20
S-Corporation	390	24
C-Corporation	90	19
General Partnership	28	1
Limited Partnership	20	0
Something Else	4	0

Table 2.5: Average Venture Capital Financing by Year and Legal Status

Year	Sole	LLC	S-Corp	C-Corp	GP	LP
2004	0	12972	915	78306	0	0
2005	0	17912	1221	117600	253	3333
2006	0	2022	16129	113862	303	0
2007	0	434	0	29856	370	0
2008	0	0	0	46431	0	0
2009	0	4819	0	53846	0	0
2010	0	3778	0	6682	0	0
2011	0	0	0	12895	862	0

formation to third parties through real or assumed correlation. For example, patents signal to investors that the patent holders are intelligent, hard working, or innovative in a way that would confer a business advantage unrelated to the true market value of the patent. The appropriative value of the patent is observed value minus signal value.

The POLS model tests the hypothesis concerning patents as verifiable information. The result of this model indicates how correlated patenting is with external equity financing and thus contains information for both signal and appropriation value. The FE model tests the hypothesis concerning the appropriative value of patenting. The dependent variable, external equity financing, represents smart money with industry and investment experience. The results presented in the main paper are for a binary indicator of equity. Appendix B presents the results for a log of equity specification.

The main assumption of the FE model is the presence of firm-specific, time-invariant, unobserved, confounding variables. The FE estimation procedure removes these influences producing an unbiased beta estimate. These confounding influences include any traits that are dually correlated with patenting and equity financing. They include intelligence, ability, or diligence and are exactly the factors that contribute to the source of the signal value of the patent. The POLS model cannot account for these influences so the patent estimate contains the signal value but the FE model does not.

The FE procedure is instructive to walk through. As long as they unobservable variables are firm-specific, time-invariant, and linearly additive, then this paper's interpretation of signal and appropriative values is correct. Notation below is borrowed from Angrist and Pischke (2009).

The POLS model estimates Eqn. 2.1 where Y_{it} represents equity financing, P_{it} is the patent variable, and \mathbf{X}_{it} represents all control variables. The true model is portrayed in Eqn. 2.2 where A_i represent unobservable time-invariant firm-specific traits.

Because they are not in the POLS model, the residual u_{it} absorb these influences, shown in Eqn. 2.3. The FE procedure, Eqn. 2.4, removes all time-invariant variables from the regression since the time average equals itself.

$$Y_{it} = b_0 + X'_{it}b_1 + pP_{it} + u_{it} \quad (2.1)$$

$$Y_{it} = \alpha + X'_{it}\beta_1 + \rho P_{it} + A'_i\gamma + \nu_{it} \quad (2.2)$$

$$u_{it} = A'_i\gamma + e_i \quad (2.3)$$

$$Y_{it} - \bar{Y}_i = (A_i - \bar{A}_i)'\gamma + (X_{it} - \bar{X}_i)'\beta_1 + \rho(P_{it} - \bar{P}_{it}) + (\nu_{it} - \bar{\nu}_i) \quad (2.4)$$

$$= (X_{it} - \bar{X}_i)'\beta_1 + \rho(P_{it} - \bar{P}_{it}) + (\nu_{it} - \bar{\nu}_i) \quad (2.5)$$

The expected impact of these unobservables on p is positive bias. The direction of this bias can be deduced from Eqn. 2.6. This implies that p is equal to the true value of patenting plus the influence of the unobservables on Y_{it} times estimates of regressions of the A_i variables on P_{it} , patenting. The bias term contains various influences but the stronger influences are assumed to be positively correlated with equity and patenting. These include intelligence, industry experience, and diligence.

$$\frac{Cov(Y_{it}, P_{it})}{Var(P_{it})} = \rho + \gamma'\delta_{AP} \quad (2.6)$$

Table 2.6 contains the linear probability model results. All standard errors are clustered on firm and are robust to heteroskedasticity and serial correlation. Column I contains the results of the pooled model with year dummy variables. Column II contains the results of the fixed effects model. The estimate in column I is the correlation between patent and external equity. Since it is a correlation it represents the sum of the effect of signal and appropriative value of the patent. Column II is the

Table 2.6: Dependent Variable: Binary Indicator of Positive Equity Financing

	I	II
Patent	.13*** (.028)	.008 (.011)
R^2	.165	.022
Observations	17,906	17,906
F-test	121.5***	17.9***
Note:	*p<0.1; **p<0.05; ***p<0.01	

causal effect of obtaining a patent on external equity financing. Column II represents the just the appropriative value of the patent since the fixed effects methodology wipes out the influence of non-patent related factors. The estimate in column I indicates a statistically significant 13% higher probability of receiving external equity with the ownership of a patent. The estimate in column II indicates that this is not caused by the patent but rather unobservable characteristics of the firm or owner.

With the evidence gathered, H_0 is rejected by the strong results in column I from both regression tables. Under H_1 , we have presented evidence that patents do contain verifiable information and affect equity financing. However, we fail to reject H_2 , which states that patents have no appropriative value. The two taken together suggest that patents mainly represent a signaling effect. To the extent that there is any appropriative value (in the eyes of investors, at least), it is subsumed by other observed and unobserved characteristics.

2.5.3 Model Diagnostics

This section verifies the model's assumptions. We are primarily concerned with the fixed effects specification and ensuring accurate inference via heteroskedasticity cluster robust standard errors. The central assumption invoked by the fixed-effects model is a special form of the zero-correlation between residual and regressor assumption of OLS. Each fixed effect, firm and time, is constructed as part of a three-way error

Table 2.7: Time Varying Coefficients

	I	II
Patent	.13*** (.028)	.006 (.011)
R^2	.17	.08
Observations	17,906	17,906
F-test	20.7***	7.5***
Note:	*p<0.1; **p<0.05; ***p<0.01	

component model. The firm effect must be constant across time for each firm and the time effect must be constant across firms for each year. Testing for the presence of such effects boils down to specification tests in the form of F or χ^2 tests.

The first test of proper specification compares a random effects model with the fixed effects specification. The random effects specification allows for correlation between the effect and the regressors. Wooldridge (2010) argues that the use of a Hausman test is appropriate for selecting between the two assumptions. The Hausman test results in a χ^2 estimate of 433.93 with 22 degrees of freedom⁵⁴ leading to a p-value of $< 2.2e - 16$. We thus reject the null in favor of the fixed-effects specification.

Even though t-tests provide statistically significant results for each of the time dummies, a panel F test is also employed. The F test compares a model with time dummies to a model without them. The F statistic is 51.49 on 7 and 13,319 degrees of freedom leading to a p-value of $< 2.2e - 16$. Since the null hypothesis is no significant difference between the models, the evidence argues for time effects. A Lagrange Multiplier test further confirms this result (Breusch-Pagan for the unbalanced panel) with a χ^2 estimate of 10739 on 1 degree of freedom and p-value of $< 2.2e - 16$.

Finally, there is evidence of heteroskedasticity identified via a Breusch-Pagan test performed on a pooled model with time dummies. This heteroskedasticity most likely persists once firm effects are included considering the vast range of firm types. To

⁴⁵Time effects are included as dummies in both models to focus on the firm effects assumption.

correct for the influence that heteroskedasticity has on inference, robust standard errors are used throughout. The standard errors are clustered on firm and employ and the finite sample correction suggested by MacKinnon and White (1985) and also implemented in Stata.

2.5.4 Heterogeneous Effects

The above estimate averages the main effect of interest over all included covariates. For example, patents in a tech-heavy industry have the same value as a patent in a tech-light industry. This is a strong assumption that is likely to be violated since patents are concentrated in certain industries and among certain entrepreneurial traits and characteristics, and they are valued differently according to different firm and owner traits. Patenting entrepreneurs are likely to be more educated, have more experience, and work in more technologically inclined industries. These examples are likely to produce great heterogeneity in patent value.

Heterogeneity Across Industry

The original specification is modified to interact two-digit NAICS with the patent indicator. This provides tests on industry-unique effects. Interestingly, this is not a dimension along which patenting has heterogeneous effects. The two-way fixed effects model does not provide any significant results related to industry or patent. The results are not provided in this article. This is likely not a degrees of freedom problem since many of the patenting industries are over-sampled.

Heterogeneity Across Legal Status

Legal status provides a useful indicator of entrepreneurial experience and intent for the company since each classification comes with different costs and benefits. Venture capitalists may be more attracted to C-corporations, as they allow different types of

stock that allow the payment of dividends. As with industry, legal status is interacted with the patent indicator. Results on the pooled and fixed effects linear probability models are presented in Table 2.8.

By interacting patent with legal status the coefficient on patent is interpreted as the effect of patenting for the reference group, sole proprietorship. The interaction terms are the impact relative to the reference group. For example, in column II, the total effect of patenting in a general partnership is .02-.7, or -68% percent. Since the other interaction terms are not statistically significant we cannot conclude that there is any difference between these groups and the reference group when patenting. Each group, like the reference group, experiences a statistically significant 2% increase in the probability of obtaining equity financing following a patent. The inclusion of all firm types in the main results was occluding this positive result.

Furthermore, C-Corporations see a large, statistically significant bump in column I, both on their own and interacting with patents. The combined effect of these results is that a) forming as a C-Corporation is a strong signal that a firm will receive financing that gets washed out in the fixed effects model, and b) patents have some appropriative value for all firms except general partnerships. Forming as a general partnership was so detrimental to receiving financing that patents were unlikely to change that, as a signal or otherwise.

To verify these results, we run the regression on subsets defined by legal status shown in Table 2.9. These results mirror the original results suggesting a pure signal effect and appropriation effect. It is worth noting that C-Corporations with patents are 25% more likely to obtain external equity than C-Corporations without patents. This result, only appearing on the pooled panel model, suggests that C-Corporation status plus patenting is a strong signal to equity investors.

Table 2.8: Patent Interacted with Legal Status on Binary Dependent

	I	II
Patent	2.2e-02 (1.8e-02)	2.0e-02* (1.2e-02)
LLC	-1.7e-03 (7.6e-03)	-4.3e-02*** (1.3e-02)
S-Corp	-2.6e-02** (1.1e-02)	-3.3e-02** (1.3e-02)
C-Corp	7.9e-02*** (2.6e-02)	1.5e-02 (2.3e-02)
GP	-2.1e-02 (3.1e-02)	-4.5e-03 (1.2e-02)
LP	2.2e-02 (5.1e-02)	5.8e-02 (6.4e-02)
Something else	-5.9e-02 (3.8e-02)	-2.5e-03 (1.8e-02)
Patent:LLC	9.5e-02* (5.4e-02)	-2.8e-02 (2.8e-02)
Patent:S-Corp	3.0e-02 (4.5e-02)	2.0e-02 (2.5e-02)
Patent:C-Corp	3.0e-01*** (7.1e-02)	3.2e-02 (3.8e-02)
Patent:GP	-1.7e-01*** (4.1e-02)	-7.0e-01** (2.9e-01)
Patent:LP	-2.0e-02 (1.2e-01)	2.7e-01 (1.8e-01)
R^2	.17	.03
Observations	17,906	17,906
F-test	110.06***	18.99***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 2.9: Patent Estimates on Subset Data by Legal Status on Binary Dependent

	I	II
C-Corp	.243*** (.0687)	.003 (.0129)
LLC	.116* (.0481) (1.2e-01)	.003 (.0217) (1.8e-01)
R^2	.17	.03
Observations	17,906	17,906
F-test	110.06***	18.99***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 2.10: Dependent Variable: Log of External Financing and Internal Equity

	I	II
Patent	1.51*** (.286)	-.21 (.225)
R^2	.11	.01
Observations	17,906	17,906
F-test	74.9***	9.6***
Note:	*p<0.1; **p<0.05; ***p<0.01	

2.5.5 Robustness Checks

There may be some concern that external equity is either too narrow or too far removed from the value of a patent to be a useful proxy. This section tries a few suitable alternatives, and the main results look remarkably similar. The first check is to expand the dependent variable to include external debt and internal equity with external equity. This accounts for the patent's making the firm more attractive to other investors as well. Internal debt still acts as a control variable because this is much less likely to be influenced by patent. The thought is that owner credit card purchases and family loans are independent of patenting.

Table 2.11: Dependent Variable: Binary of External Financing and Internal Equity

	I	II
Patent	.10*** (.018)	-.01 (.012)
R^2	.26	.06
Observations	17,906	17,906
F-test	214.6***	45.1***
Note:	*p<0.1; **p<0.05; ***p<0.01	

2.6 Conclusion

This study estimated the role of patents in obtaining financing, especially equity financing. Comparing pooled OLS and fixed effects results allows us to tease out the relative strengths of the signal and appropriation value of patents. This is important research for both investor and entrepreneur as the findings indicated an overvaluation.

The results presented in the main model suggest that patent value is pure signal value. After some heterogeneity analysis, appropriative value reveals itself for sole proprietorships, LLCs, S-Corps, and C-Corps. The results in the main model can be explained by the strong negative influence of general partnerships.

CHAPTER 3

PEER EFFECTS IN ENTREPRENEURSHIP

3.1 Introduction

The task of financing a fledgling firm is an important concern for entrepreneurs. The process of firm growth is resource intensive, involving capital, labor, materials, marketing, and the entrepreneur's time. Fledgling firms often do not have the revenue in the first few years to sustain operations and grow. Even with revenue coming in, it will not be enough to cover costs and investment. The financing must be sourced specifically for financing. Some entrepreneurs are able to self-finance in the early years but that is not a realistic option for most entrepreneurs. All firms must rely on some external source eventually.

The landscape of financing opportunities is immense for entrepreneurs. Internal sources include savings, credit cards, loans from family, and employee equity. There are also external sources, such as small business loans, angel investors, and venture capitalists. How much access an entrepreneur has to these different sources can affect entrepreneurial strategies, outlook, and firm trajectory.

The amount of financing in a given time frame and the appetite for growth, while industry contingent, are also hugely personal and depend on one's tolerance for risk exposure. Raising a large amount of money is not always beneficial, as it comes with

varying obligations and pressure, depending on the source and market conditions for the fledgling firm. An often-cited reason for the financing disparity between firms started by men and firms started by women is the appetite for risk exposure. The notion of the more risk tolerant male pervades our culture but the literature is not so conclusive. Some of the literature does indeed find this to be the case but the other half finds no gender differences in risk tolerance.

Another intuitive reason for financing differences between men and women is the role mentorship in entrepreneurship. Although women entrepreneurs are a fast growing segment of the US labor market, men have long dominated. The greater labor market share enjoyed by men affects perceptions, beliefs, and actions for both genders. Discrimination reinforces labor market outcomes, raising barriers in desirable fields and lowering them in areas that are not desirable. This explains why women are over-represented in low-growth industries while men are over-represented in high-growth industries.

In this study, I show that men and women are differently motivated by peer financing levels. I control for many of the factors claimed to drive the financing divide. These include experience, industry, credit, and motivation. Furthermore, through a fixed effects specification, I am able to remove unseen heterogeneity to argue that men are influenced by peer financing while women are not.

The remainder of this chapter is organized as follows. Section 3.2 provides a survey of the relevant literature and describes the hypothesis that I test. Section 3.3 describes the innovations used in this chapter. Section 3.4 describes the analysis and presents the main results, diagnostics, and some alternative specifications. Section 3.5 concludes the chapter with a summary of main findings.

3.2 Background

The foundations of this study lie in several veins of research. The peer effect literature concentrates on education and development. The gender differences in entrepreneurship literature considers structural differences as well as discrimination. The prospect theory literature exists in the finance space, some of which also considers gender differences. The firm financing literature also includes some discussion of gender differences. I sample from this literature to provide the necessary background.

3.2.1 Industry

The literature on gender differences in financing is crucial for this study. One of the primary differences, driving other observed differences is industry concentration. Minniti (2009) finds that female entrepreneurs are concentrated in retail and service industries.

Industry sector can help explain other observed differences between male and female entrepreneurs. These differences are explored in Coleman and Robb (2009) who find that women begin their business with less financial capital and obtain lower levels in years two and three. Scherr, Sugrue, and Ward (1993) do not address gender influence directly but acknowledges that expected income is a determinant of debt use. But, Orser, Riding, and Manley (2006) finds no difference in debt financing after controlling for industry sector among male and female entrepreneurs in Canada.

On the supply side of financing, Brush et al. (2001) explain why women entrepreneurs receive less venture capital. They attribute the gender difference to the disparity of females being employed in the venture capital industry. (Ewens and Townsend 2017) reinforce this by explaining equity financing differences as the result of discrimination. They explain the discrimination as being caused by lack of female early stage investors in the equity market.

(Minniti and Bygrave 1999) develop a model to explain entrepreneurial participation. Entrepreneurial action is the result of personal utility maximization where the benefits of entrepreneurship are weighed against the costs. Even though there might be social and personal desire to for equity in the entrepreneurial space, the implied costs of entering male-dominated markets for women still remains high. These costs include network structure, discrimination, mentorship, among others. This argument is supported by the findings of Fuentes-Fuentes et al. (2017) who find that women entrepreneurs rely on networking for business innovation rather than capital infusions.

These costs don't explain all behavior, however, as the observed industry distributions can be explained by motivation for starting a business. Burke, FitzRoy, and Nolan (2002) find that females, due to greater familial commitments and obligations, seek to balance the pecuniary and non-pecuniary benefits of self-employment, while men respond only to the pecuniary incentives. Whether these preferences are innate or driven by cultural and personal constraints is not considered.

3.2.2 Risk Preference

Another perspective on the gender balance issue is the role that differences in risk tolerance play in entrepreneurial decisions. Anecdotal evidence of gender disparities in risk preference are common and pervasive, including the sports and professions that males are attracted to, driving accident rates, and binge drinking. Despite these, the literature is divided on whether females are actually more risk averse.

Croson and Gneezy (2009) argue that the literature supports the conclusion of a more risk averse female population through a meta-analysis. In another meta-study, Charness and Gneezy (2012) find strong evidence for gender based differences in risk preference in 15 experiments on financial risk taking. However, in response to that article, Nelson (2015) performs a meta analysis of 35 articles on the subject. She finds substantial overlap in the implied distributions between male and female risk aversion,

indicating that the suggested difference is not so stark and clear. Furthermore, studies offering contextual analysis claim that covariate differences account for much of the observed risk preference difference possibly suggesting that preconceived notions of gender roles cause the difference rather than there being an inherent difference.

Fehr-Duda, Gennaro, and Schubert (2006), in an observational data study, find that women are more risk averse in financial decision making due to the curvature of women's probability weighting functions. Men and women have similar value functions. Differences in behavior arise from the evaluation of the probability of gain and loss. Women tend to overestimate small probabilities and underestimate large probabilities of gain.

3.2.3 Firm Financing

Within each industry financing is a crucial asset for nascent firms. Besides industry, financing is affected by multiple factors including start-up size, asset structure, organization type, growth orientation, and owners' characteristics (Cassar 2004). Storey (1993) suggests that obtaining bank financing is not a matter of personal characteristics but of legal structure and the use of personal savings. Whether this conclusion contains implicit gender bias is not addressed.

An entrepreneur can seek funding from a variety of sources. The pecking order hypothesis argues that these sources are ranked in terms of cost, with internal sources being the cheapest, debt sources next, and external equity sources being the costliest because this money usually comes with strict conditions and loss of autonomy. Vaznyte and Andries (2019) argue that a firm's entrepreneurial orientation, defined as the entrepreneur's propensity to engage in innovation, risk-taking, and market opportunity proactiveness, affects whether a firm follows the pecking order hypothesis and this differs based on environment and whether the firm has achieved break-even.

The literature on gender-based entrepreneurial financing is broad. Bellucci, Borisov,

and Zazzaro (2010) find that men and women pursue small business lending at the same rate but women face a tighter lending market. Neeley and Auken (2010) find no difference in female and male reliance on bootstrap financing, but results differ when slicing the sample by age, education, overdraft privilege, and change in sales. For example, more-educated female entrepreneurs rely more on bootstrap financing than more-educated males. Another survey-based study, Eddleston et al. (2016) find that gender plays an important role in determining how firms are rewarded for their characteristics, always to the female entrepreneur's detriment.

Robb and Wolken (2002) study the issue of gender-based differences in financing, through demand side factors. Using the survey of small business finances they find that the observed differences in financing are driven by owner and firm characteristics rather than supply side factors. Constandtinidis, Cornet, and Asandei (2006) find much of the same using different methods and data. They, however, took the additional step of acknowledging that industry choice is highly influenced by stereotypes and other cultural factors. Saporito, Elam, and Brush (2012) looks at both supply and demand side influences by using bank-firm pairs to investigate how perceptions matter between different gender pair combinations. Male-male pairs expressed the most satisfaction in access to credit while female-female pairs the least satisfaction.

3.2.4 Prospect Theory

This chapter will interpret the findings as either being caused by demand- or supply-side factors. The most prominent supply-side factor likely to cause differences in financing by gender is discrimination on the part of the lender. Demand-side factors include preferences and risk tolerance. One of the primary demand-side factors is based in utility theory that argues preferences can be influenced by their peers.

Classical microeconomics provides a framework for understanding risk preference. Economic agents make decisions regarding uncertain outcomes by evaluating payoffs

with respect to the objective probabilities over the outcome space and their own risk preference. In regards to a fair bet, one in which the expected payout equals the probability of winning multiplied by the net gain, risk-averse individuals do not partake but risk-loving individuals gamble.

Prospect theory, developed by Kahneman and Tversky (1979), makes the study of uncertainty more realistic by removing objective probabilities from the model. In prospect theory there are two stages of decision making under uncertainty. The first stage is framing or editing and the second stage is evaluation. Framing involves how the choice problem is presented, involving norms, habits, and expectations (Tversky and Kahneman 1986). Busenitz and Barney (1997) confirm that entrepreneurs are informed via inherent biases and heuristics in their decision making process, at least relative to managers.

Assuming that entrepreneurs are relatively well-informed regarding the industry in which they are competing, mean revenue by industry can be used as the reference point by which entrepreneurs will judge themselves. Entrepreneurial risk preference depends on which side of average industry revenue the firm finds itself. Prospect theory suggests that underperforming firms would be more likely to make riskier decisions. Firms performing above average should not exhibit risky behavior, all else equal.

3.3 Data and Variables

This chapter relies upon the Kauffman Firm Survey (KFS) for data. For a more thorough account of the data and variables used, refer to chapter 1. This section will only present the main innovations of this chapter.

3.3.1 Outcome Variable

This study considers one primary outcome, total financing. Total financing sums all financing sources for each firm and year to get a sense of how, besides revenue, the firm is supporting itself. Financing a nascent firm's operation is a crucial and difficult decision to make in a firm's early years. Financing, other than revenue, allows the firm to scale more rapidly and compete more effectively.

3.3.2 Main Explanatory Variable

The key explanatory variables are peer financing, gender of owner, and the interaction between the two. Peer financing is defined as within industry average equity financing, leaving out firm i when computed for firm i . Other financing sources, including external debt, internal debt, and internal equity, are summed into a control variable. This is computed for every firm, year, and industry, as identified by the two digit NAICS, and is represented as $\ln(\overline{Fin}_{-i})$. The interaction is the primary variable of interest while the un-interacted variables do not provide much of interest.

The gender variable is binary, but firms in the sample can and do have multiple owners. If the firm has primary owner who is a woman, that firm is identified as female-owned.

3.3.3 Descriptive Statistics

Table 3.1 presents trends regarding attrition and patent ownership. Even though this survey oversamples female-owned firms, the proportion of female owners is quite low. The rate of attrition can be observed in the change in number of firms. Female-owned firms drop out of the survey at a slightly greater rate.

Table 3.2 shows statistics regarding financing and profitability through time. External financing is relatively rare in this sample of representative firms. Debt is the

Table 3.1: Gender Ownership

Year	Firms	Female-Owned	Proportion
2004	3140	835	26.6%
2005	2837	742	26.3%
2006	2554	645	25.6%
2007	2330	577	25.3%
2008	2092	506	24.9%
2009	1928	457	24.6%
2010	1775	411	24.1%
2011	1630	375	24.1%

Table 3.2: Finance Descriptives

Year	Firms	Debt	% Debt	Equity	% Equity	$\Pi > 0$	% $\Pi > 0$
2004	3140	193	6.15%	144	4.59%	1149	36.6%
2005	2837	264	9.31%	187	6.59%	1555	54.8%
2006	2554	290	11.4%	189	7.40%	1564	61.2%
2007	2330	307	13.2%	180	7.73%	1379	59.2%
2008	2092	303	14.5%	165	7.89%	1206	57.6%
2009	1928	307	15.9%	152	7.88%	1152	59.8%
2010	1775	310	17.5%	146	8.23%	1095	61.7%
2011	1630	292	17.9%	140	8.59%	1092	67.0%

more common financing sources. A large number of firms report positive profits in each year. There appears to be a slight recession effect in years 2007 through 2009 when firms reporting positive profits dip. By 2010 and 2011, however, the effect of the recession seems to have been over.

Table 3.3 shows average financing levels by gender from all sources, internal and external. Males receive much more financing in total most years of the survey, especially pre-recession. The last two years of the survey witness a much more even distribution of total financing for the average man and woman.

Table 3.4 shows how external equity rates differ by gender. The numbers presented represent unique equity contracts for each year. Early in the survey years, males do better at acquiring external equity, but the rates by gender become quite even in later

Table 3.3: Average Total Financing by Gender

Year	Firms	Female-Owned	Average Female	Average Male
2004	3140	835	23,789	87,846
2005	2837	742	46,924	78,376
2006	2554	645	22,334	119,062
2007	2330	577	72,300	53,941
2008	2092	506	46,980	47,911
2009	1928	457	26,279	72,027
2010	1775	411	41,147	34,340
2011	1630	375	30,242	26,082

Table 3.4: External Equity by Gender

Year	Female Ext Eq	Proportion	Male Ext Eq	Proportion
2004	23	2.8%	121	5.2%
2005	14	1.9%	83	4.0%
2006	13	2.0%	57	3.0%
2007	8	1.4%	33	1.9%
2008	9	2.0%	22	1.4%
2009	6	1.3%	19	1.4%
2010	4	1.0%	15	1.2%
2011	5	1.3%	13	1.1%

survey years. This raises an interesting question, one that this chapter will not get into further. Do women entrepreneurs have more to prove, thus requiring more years of experience, before they are able to acquire financing at the same rate as males?

Table 3.5 shows the numbers and proportion of external debt financing by gender. Debt is easier to come by than equity for both genders but this table shows males still out-finance females through debt.

3.4 Estimation and Results

The single hypothesis being tested is whether there is a gender response to peer financing. Even though peer financing is the main explanatory variable, with gender, I do not test for peer effects. The estimated effect will be a sum of demand- and

Table 3.5: External Debt by Gender

Year	Female Ext Eq	Proportion	Male Ext Eq	Proportion
2004	36	4.3%	157	6.8%
2005	27	3.6%	118	5.7%
2006	21	3.3%	99	5.3%
2007	20	3.5%	79	4.6%
2008	16	3.2%	65	4.3%
2009	11	2.4%	55	3.9%
2010	8	1.9%	58	4.5%
2011	8	2.1%	33	2.8%

supply-side influences. Stated formally:

\mathbf{H}_0 : Men react more to increases in peer financing levels, as identified by industry sector.

\mathbf{H}_A : Men respond no differently than women to peer financing.

I test the null hypothesis with a two-way fixed effects model shown in Eqn. 3.1. γ_1 provides a statistical test on whether women entrepreneurs respond to peer financing, since women are the reference group. Therefore, the estimate of interest, providing a statistical test of \mathbf{H}_0 , is γ_{12} which represents the extra response to peer financing by men. This is shown in Eqn. 3.2.

$$Fin_{it} = \gamma_1 Peer_{it} + \gamma_2 Gender_{it} + \gamma_{12} Peer_{it} * Gender_{it} + \mathbf{X}_{it}\beta + \mu_i + \mu_t + \epsilon_{it} \quad (3.1)$$

$$\frac{\partial Fin_{it}}{\partial Peer_{it}} = \gamma_1 + \gamma_{12} Gender_{it} \quad (3.2)$$

An econometric concern with estimation of the effect of interest in 3.1 is endogeneity of *Peer* because certain unobservable factors influence industry selection, peer financing, and firm financing such that gender is not randomly assigned across industry. Industry selection, motivation, financing strategies, and discrimination are

already proven concerns in the literature. The two-way fixed effects model selection alleviates much of this concern since firm and time effects absorb the influence of time- and firm-constant effects.

Some unobservable confounding factors at the firm level may be time-variant, however, such as preference, entrepreneurial knowledge, and motivation. Also, market forces may not impact each industry and firm in a consistent way, such as the fact that the recession hit manufacturing harder than finance. To alleviate these concerns, a wide range of control variables are included to remove firm and industry variation, that may be correlated with our primary explanatory variables, from the residual.

Despite these precautions, endogeneity is still an issue for of the above reasons not captured via specification. Since the ultimate goal is the observation of a phenomenon, this is not an issue here. The goal is to see if the process that affects male investment also affects female investment.

The use of a group average to predict an individual's outcome traces back to the social interaction work of Manski (1993). The identification of a peer influence requires at least the removal of each individual's covariate value when calculating every individual's peer variable (Kline and Tamer 2011). This strategy is deployed within a panel data fixed effects framework.

Table 3.6 presents the truncated results of the two-way fixed effects model. All reported standard errors are robust to heteroskedasticity and serial correlation and are clustered at the firm level. The coefficient on peer financing indicates that peer financing has no statistically significant effect on own financing when gender is female. The coefficient on male confirms no statistically significant difference between the genders when peer financing is zero. The coefficient on the interaction term is statistically significant at the 1% level. The nature of the gender dummy means that additional peer financing does have an influence on male own financing. The magnitude's interpretation is that for every percent increase in peer financing, male

own financing increases by .15%, an inelastic but significant result. The max amount of financing raised each year ranges from 10 million to 120 million. At the upper threshold of what these firms are raising, .15% is quite economically significant.

Table 3.6: Dependent Variable: Log of Total Financing

	Estimate	S.E.	p-value
Peer	-.05	(.04)	-
Male	-2.79	(2.79)	-
Peer:Male	.15	(.05)	***
...			
Observations			17,895
R ²			.06
F-test			19.1***
Note:	*p<0.1; **p<0.05; ***p<0.01		

3.4.1 Diagnostics and Robustness

Causal identification of the effect of interest rests, in part, on the appropriateness of the two-way fixed effects model. The two-way fixed effects model is the correct specification and tests against alternative specifications conclude this. A second barrier to identification is the connection between peer financing and self-financing. This connection is a bit trickier to reason with but a few alternative specifications are presented make a case for this formulation.

While the random effects model does not produce significantly different results for the coefficients of interest, a Hausman test is conducted. The null hypothesis, that fixed and random effects are not significantly different, is conclusively rejected with a χ^2 of 318.9 on 52 degrees of freedom providing a p-value of $2.22e - 16$. Following Wooldridge (2010), a rejection of the null hypothesis is statistical support for the fixed effects model.

The Hausman test confirms only the inclusion of the firm fixed effect. An F test is performed to test the importance of including the time effect. The results

are conclusive that the time dummies improve the model with an F statistic of 76.4 with 7 and 14703 degrees of freedom providing a p-value of $2.22e - 16$. Without the time effect, the coefficient for females is positive and significant and the coefficient on males is stronger, positive, and significant. This means that time is a significant confounding factor that needs to be addressed.

In Table 3.7, I present an alternate specification that interacts time dummies with all control variables. This specification addresses the influence of time-variant firm effects so long as that influence impacts the control variables. Again, robust standard errors are reported with an 'Arellano' estimator that controls for heteroskedasticity and serial correlation, grouped at the firm level.

Table 3.7: Time Interaction with all Control Variables

	Estimate	S.E.	p-value
Peer	-.002	(.04)	-
Male	-2.92	(2.17)	-
Peer:Male	.09	(.05)	*
...			
Observations			17,895
R ²			.16
F-test			12.6***
Note:	*p<0.1; **p<0.05; ***p<0.01		

The male and interaction coefficients remain nearly the same. There is still strong evidence that males respond to increased average financing by their industry cohort. The effect is slightly weaker in this specification. This possibly suggests that some of the response is due to unobservable influences.

Serial Correlation is not found via a Breusch-Godfrey test but the panel is short and testing is likely to be unreliable. Serial correlation is intuitively likely to be present so standard errors are employed robust to this issue. Heteroskedasticity is found via a Breusch-Pagan test on a pooled model with time fixed effects. The Arellano estimator of the robust covariance matrix is employed to control for both serial correlation and

heteroskedasticity. Results change little when using an Arellano estimator vs a White estimator. The inclusion that makes a large difference is clustering on the firm.

3.4.2 Heterogeneous Effects

One concern of the above analysis is that the male and female distributions of the control variables are not equal. Males and females are not present in the same concentrations within each industry or each legal status, for example. The results of the linear models above absorb these differences to give an average over the other observed factors. This means that the effect may be weaker or stronger if one looks at different slices of the data.

This section explores specifications related to these heterogeneous effects. Each model presents a different control variable interacted with the with gender, other financing interaction. As discussed in the literature, legal status is a predictive factor for multiple outcomes due to its relation to owner motivation, experience, and intent. The first alternative specification interacts legal status with the gender peer financing interaction. The results in Table 3.8 allow us to see if the main results hold in each legal status or if it is constrained.

The main result in this specification, Peer:Male, is not significant but this is just for sole proprietorships because of the third interaction with legal status. Peer interacted with just legal status allows us to see if females behave differently within each status. Female entrepreneurs are more responsive to peer financing within the LLC and GP legal structures, significant at the 1% level, and S-Corporations at the 10% level. But, males are an extra .26 percentage points more responsive in S-Corporations, significant at the 10% level. Males finance an extra .46% for every 1% increase in peer financing within C-Corporations.

There is evidence of heterogeneity by legal status. Given the size of the dataset and the number of observations per group, statistical power is weak. There are not

enough observations to speculate about what causes these results.

Table 3.8: Legal Status Interaction

	Estimate	S.E.	significance
Peer ¹	-.237	(.031)	
Male ¹	-3.321	(2.706)	
Peer:Male ¹	.011	(.039)	
Peer:LLC	.335	(.090)	***
Peer:S-Corp	.230	(.118)	*
Peer:C-Corp	.232	(.184)	
Peer:GP	.909	(.260)	***
Peer:LP	.641	(.984)	
Peer:Male:LLC	.028	(.113)	
Peer:Male:S-Corp	.259	(.141)	*
Peer:Male:C-Corp	.459	(.248)	*
Peer:Male:GP	-.332	(.314)	
Peer:Male:LP	.200	(1.037)	
...			
Observations			17,871 ³
R ²			.06
F-test			16.77***
Note:	*p<0.1; **p<0.05; ***p<0.01		

¹The reference legal structure is sole proprietorship.

²Legal Status and Legal:Male are omitted.

³Legal Status 'Something else?' was removed.

Heterogeneity within and between industry results is presented in Table 3.9. Only the coefficients of interest are presented to conserve space. The three-way interactions suggest that men in the identified industries respond strongly to peer financing. These industries drive the main results.

Table 3.9: Industry Interaction

	Estimate	S.E.	significance
Peer ¹	.360	(.481)	
Male ¹	9.143	(6.566)	
Peer:Male ¹	-.533	(.523)	
Peer:54	-1.265	(.523)	**
Peer:Male:42	1.598	(.764)	**
Peer:Male:51	.951	(.555)	*
Peer:Male:54	.985	(.570)	*
Peer:Male:72	1.196	(.598)	**
... ²			
Observations			17,895
R ²			.06
F-test			9.38***
Note:			*p<0.1; **p<0.05; ***p<0.01

¹The reference industry is code 11.

²Industry and Industry:Male are omitted.

Industry 21, 22, 55, 91, 92 omitted due to not enough women representation.

3.5 Discussion and Conclusion

The causal interpretation presented above is susceptible to criticism still. Following the argument presented in Angrist and Pischke (2009) starting with equation 4.6.9 there is a thorny identification issue present whenever one regresses a variable on that variable's group average. The issue, not addressed by individual and time effects, is that Peer Financing and Individual Financing are both susceptible to group-level shocks. These shocks are captured by the residual, creating a classic case of endogeneity. This would arise in this study if, for example, a news-generating company, like a Facebook, either attracts or scares away investors to the whole industry sector. The issue is less egregious when the group average is calculated for each individual leaving out that individual's covariate value, but still present.

This problem does not invalidate the results but it prevents further analysis. Gender peer influences cannot be decomposed into demand and supply factors. This

group-level endogeneity is a supply-side influence that cannot be addressed with these data. This study cannot say further how to interpret the results, just that there is strong support for male-female differences in financing. I cannot indicate that the results are due to preferences, risk preference, or discrimination, unfortunately.

This chapter makes an important contribution to the literature, however. It demonstrates that males are responsive to peer financing, as defined by the total amount of financing within industry sector. A one percent increase in peer financing causes a .15 percent increase in firm financing for males. This effect is not present among female entrepreneurs. These findings are robust to heteroskedasticity, serial correlation, firm and time effects. The source of financing also does not influence the nature of this result.

Due to the nature of the data, this finding cannot be explained as a demand-side or supply-side phenomenon. Further research is needed to address this shortcoming. Demand-side studies are inconclusive whether male and female run firms differ in risk aversion or innate preferences. Supply-side studies, however, do point to the presence of gender discrimination in the lending markets. Decomposing the influence of demand- and supply-side effects would be a natural next step.

CHAPTER 4

NETWORK ANALYSIS OF THE S&P 500

4.1 Introduction

All investors want better performance out of their portfolio. The two sides to managing performance are risk and return. Hedging strategies require the careful selection of securities to manage risk and return. For example, a pairs trading strategy involves selecting two securities that trend together. If one of those securities deviates from that trend the investor can use that information to buy one long and sell the other short in anticipation of a return to equilibrium. This type of strategy can both mitigate risk and increase returns over the buy and hold strategy but only with properly selected securities.

The selection of pairs is crucial. One approach is to select two securities from a group known to trend together, for example, two stocks out of the same industry, or two currencies from countries that are strong trading partners. Another method involves searching over the constituents of a particular index or market. This search can be exhaustive, testing all pairs within of the group.

Two search methods, one seeking high correlation pairs and the other incorporating Granger causality or cointegration, have gained traction over the past decade. The more recent literature tends to use cointegration more but there has not been

a study comparing the two methods. On a practical level for any strategy that requires frequent recalibration, a correlation approach is much faster than some of the multi-step regression approaches out there.

The two above procedures for finding pairs can be used to construct threshold network graphs. Network analysis is increasingly being applied to financial networks. Applications of this technology to finance include studies on global structures, market stability, market sentiment on either side of market corrections, and the aforementioned trading literature.

Network analysis involves visualization and formal metrics on the connections between nodes. For both the visual and statistical component of network graphs, the number of nodes and connections, and how the nodes cluster or fail to cluster into cliques is important. One aspect of financial market networks that is important is how strongly the constituents cluster within or across industry lines.

This chapter makes the following contributions to the literature. This is the first study to compare the two main techniques of financial network formation. The two main techniques are the previously mentioned correlation and a regression based approach. The regression approach either tests for cointegration in the residuals of regressions involving pairs of stocks or those that test for Granger causality. The different approaches lead to different network models.

The second contribution is demonstrating how the resultant network graphs can be used as data in various ways. I use the network graphs in a classification scheme to test how well industries cluster by correlation and cointegration. This aspect of the chapter tests the conventional wisdom that pairs are best selected from within industries.

A third, albeit less important, contribution is the collection of the literature found below. Articles in this area have been conducted by two disparate groups that, for the most part, ignore the other side of the literature. The two groups have either

an economics and finance background or a physics and engineering background. The two strands of this very related literature make no reference to each other.

The layout of this study is as follows. In Section 4.2 the important literature is highlighted. In Section 4.3 the data, data processing, network formation, and analysis are discussed. Section 4.4 discusses the results. Section 4.5 presents a brief but important simulation as critique of the literature. Section 4.6 contains concluding remarks.

4.2 Background

4.2.1 Previous Studies

A financial market network graph visualizes how firms in a given market, the nodes, relate to other nodes via being linked or not, via a correlation or other algorithm. The first study to approach U.S. financial markets via network graph analysis, which launched a number of papers from the physics group is Mantegna (1999). In contrast with later research out of the physics and engineering strand, this paper seeks to gain some intuitive, economic understanding of the structure of the market through network analysis. The authors approach this through the use of the minimal spanning tree, which is a subset of edges of a connected edge-weighted undirected graph to connect all nodes without any cycles and minimum possible edge weight. In effect, it identifies the central and peripheral stocks to a market or index.

Boginski, Butenko, and Pardalos (2003), creates the method that many later studies will adopt. The authors develop a threshold network of 6546 US financial instruments using daily return data with 500 consecutive trading days in 2000-2002. Two instruments, nodes, share an edge in the threshold graph if the correlation of daily returns exceeds a chosen correlation value. The statistics that they present express market connectivity for a sliding threshold value from negative one to one.

A number of studies mirror the approach taken by Boginski et al. (2003, 2005, 2006). Huang, Zhuang, and Yao (2009) applies this approach to the Chinese market. Namaki et al. (2011) looks at the Iranian market. Their approach is modified slightly. They first eliminate the market mode by subtracting the first principal component from the correlation matrix before network analysis. Vizgunov et al. (2014) looks at the Russian market. Their innovation is to include trading volume which they find is predictive of the structure of maximum cliques, a clique that cannot be extended by including one more adjacent vertex.

Like Namaki et al. (2011), Wang et al. (2013) utilize a detrended statistic to form a network graph. The detrending procedure is local as opposed to the eigenvector approach of Namaki et al. (2011), meaning that Wang et al. (2013) detrend each pair via OLS and compute the Pearson coefficient on the residuals. This local method is appealing in spirit due to its similarity to cointegration.

Nobi et al. (2014) go a step further than the above static views by investigating the interplay between economic conditions and network structure. They are interested in the global vs local (South Korean) network similarities and differences over a period of time spanning the great recession by following indexes returns and volatilities. Although different from this paper in the time and global aspects of the paper, their paper shows how networks can provide comparative analysis.

The literature on non-stationary time series is important for understanding the data preprocessing of the correlation studies and the approach taken by the cointegration studies. Yule (1926) is the earliest article laying out the difficulties involved in the statistical analysis of non-stationary time series. Granger and Newbold (1974) continues this work by specifically analyzing the prevalence of spurious associations among economic and financial series. Nelson and Plosser (1982) argues that macroeconomic time series are difference stationary. Fama (1965) argues that stocks follow a random walk model thus making them non-stationary. Dickey and Fuller (1979)

demonstrates that the distribution of estimators is non-standard in the presence of a unit root invalidating t-tests. The thrust of these papers in the current context is to show a separate approach to understanding comovements among non-stationary series.

Tong, Chen, and Buckle (2015) uses global market network visualization to answer how integrated the global markets are. Using both threshold networks and minimum spanning trees, they identify an increasing level of global integration and that the Americas, Europe, and Asia cluster continentally. They do not find an Africa cluster as individual African countries have stronger links with areas outside the continent, particularly Europe.

A few studies have used cointegration to form the connections in a market network as this article does. Singh, Roca, and Li (2018) use cointegration to form the weighted and unweighted links among national markets on a global market network. Like the other market network studies, they use log daily close price returns, but on local stock indexes. The authors form the linkages via cointegration coefficients, the larger the coefficient the greater the connections. Their study searches for global market linkages via stock indexes, for which they find some evidence. Tang et al. (2019), takes this approach a step further by employing directed graphs, where links are arrows suggesting the direction of influence, formed from Granger causality tests and lead/lag models among global indexes.

Wen et al. (2018) performs an interesting two-step filter to find links in the Chinese market. The first step filters stocks to only include those over a certain Pearson correlation threshold. The second step filters again by Augmented Dickey Fuller Test between pairs from the first filter. Network links are defined by the Pearson correlation between stocks that were found to be cointegrated, at an unstated level of significance. Gao et al. (2018) create a similar cointegration-based network. Their procedures does not model edge weight as correlation but provides a binary indicator

of connection if cointegration is detected. Their multi-step procedure passes each pair through a Johansen test for cointegration and accepts every pair for the network at the 5% level.

Ji et al. (2018) uses directed acyclic causal networks to find link, among financial assets, including Bitcoin. They find Bitcoin to be an isolated asset when viewed contemporaneously but having a lagged connection with other assets, especially during declining periods.

Bu, Tang, and Wu (2019) uses a causal network approach to find that comovement generally increases with time. Their definition of causal is Granger causality. Stock x Granger causes stock y if lagged values of x provide statistically significant information in predicting y beyond what lagged values of y provide. They assign a directed edge between two nodes if Granger causality exists in that pair. They use their constructed network to argue that sentiment based factors explain the excess comovement exceeding what fundamentals would predict.

As the use of network analysis on financial markets grows, so will the number of applications. This review has focused on graph construction through daily stock returns where links are found through either a correlation or cointegration approach. Other studies are turning to networks constructed from other metrics. Li et al. (2018), for example, construct a threshold liquidity network to analyze downturns in the Chinese market. This area of research is expanding rapidly.

4.2.2 Theoretical Perspective

This chapter contrasts networks created via correlation and networks created by cointegration. Pearson correlation measures the linear relationship represented in Equation 4.1 and provides a good representation of two stocks' tendency to move in the same direction. Only one paper on this subject, (Li et al. 2018) using correlation as the underlying construct deviated from Pearson in favor of Spearman. The Pearson

statistic takes on values between -1 and 1, inclusive. While most articles ignore negative correlation, understanding which stocks move in opposite directions provides just as much information as those that tend to move together. This is perhaps more true than ever given market trends that are making stocks in indexes more correlated (Wurgler 2010; Turner and Sushko 2018).

$$\rho = \frac{N \sum xy - \sum x \sum y}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}} \quad (4.1)$$

Cointegration is a more precise relationship between two series. Two variables¹¹, X and Y , are cointegrated if each is integrated of order d and a linear combination, $aX + bY$, is integrated of order less than d . A variable is integrated of order d if the d th difference makes that variable stationary. The simplest cointegrated relationship is a series that is linearly related to a random walk.

$$y_{1t} = \beta y_{2t} + u_{1t}, \text{ where } u_{1t} \sim I(0) \quad (4.2)$$

$$y_{2t} = y_{2t-1} + u_{2t}, \text{ where } u_{2t} \sim I(0) \quad (4.3)$$

In practice, cointegrated relationships are often identified via intuition or theory, with statistics reinforcing the notion. In financial markets, this often leads to a search restricted by industry sectors. One classic example is that of Pepsi and Coke which may have short term idiosyncratic movements away from each other but are driven by the same market forces, causing them to follow the same stochastic trend in the long-run. The Engle-Granger method is intuitively appealing, in a two variable system, regressing one on the other and testing the residual series for stationarity. It employs

¹¹Cointegrated relationships are obviously not restricted to pairs but this study restricts itself to just considering pairwise relationships. Pairwise analysis is already a $\mathcal{O}(n^2)$ problem.

the augmented Dickey-Fuller test, with null of the presence of unit root against the alternative of stationarity.

Both correlation and cointegration suggest comovement between two series. With the recent advances of computing power and desire to gain additional insight on complex financial markets, more analysis is turning toward complex networks as a tool. A complex network graph represents nodes (individuals, firms, stocks, etc.) and their connections/edges. A threshold value on the underlying statistic defines whether two nodes share an edge. For example, only stocks that have a correlation above .7 share an edge. When the network is formed from cointegration or Granger causality, a standard p-value threshold, like .05 or .01, is chosen.

These graphs are described via a number of statistics on connectivity and clustering. This article calculates edge density, node degree, correlation coefficient, the number of connected components, and the parameter the degree distribution. These statistics describe the graph's distribution.

Edge density is defined as $d = \frac{2m}{n(n-1)}$, where m is the number of edges and n is the number of nodes. This metric takes on values of zero, for a network with no edges, and one, for a complete network, where complete means every node is connected with every other node. As threshold value increases in a correlation graph, edge density decreases from one to zero.

A node's degree is the number of edges adjacent to that node. The average node degree is the sum of each node's degree divided by the number of nodes. For an increasing threshold value in a correlation graph, the edge degree will decrease, indicating that the network is becoming less connected.

The number of connected components in a graph is the number of subgraphs within which every node is connected to every other node in the subgraph via a path and is not connected via a path to any node outside the subgraph. In other words, it is a count of the number of islands, or disconnected segments, in the graph. As

threshold value increases in a correlation graph, the number of connected components increase from one, before the threshold is so high that no edges remain.

Finally, node distribution is a summary measure. Past literature has found the distribution of networks follows the power law. A network's degree distribution follows the power if the fraction $P(x)$ of nodes having x connections follows

$$P(k) \propto Cx^{-\alpha} \tag{4.4}$$

where α is a parameter with values typically in the range, $2 < \alpha < 3$ (Clauset, Shalizi, and Newman 2007). Distributions that follow the power law occur frequently in the real world. Markets in which there are a few highly connected stocks but many stocks with a low number of connections are a good candidate for being described by a power law because of the distribution's right skewness. These highly connected nodes in the right tail are under-predicted by the normal distribution.

4.3 Data and Methodology

4.3.1 Data and Variables

The S&P 500 constituent data consists of daily close price over 251 trading days from March 23, 2018 through March 22, 2019. The count of stocks included is 482 due to data availability over the period examined. The time frame is arbitrary but captures an interesting period in US financial markets. This period is interesting because of the high volatility observed. Studies on stock movement trends find an increasing correlation through time (Forbes and Rigobon 2002; Bu, Tang, and Wu 2019). Since this study is not concerned with trends, one year is enough.

4.3.2 Network Formation

Network formation and analysis is broken into three main steps. The first involves data preparation. The second takes the prepared data to form a network graph. The third analyzes the network graphs.

Data preparation is slightly different for the correlation and cointegration approaches. The correlation approach requires that all series are made stationary first through the calculation of real daily returns on daily close price. Other papers in this area use log returns but I do not see the benefit.

The cointegration approach does not require stationary series so daily close prices are left untransformed. For each unique pair of the 482 series, the Engle-Granger two-step method (Engle and Granger 1987) is used to test for cointegration. Just like the correlation approach, there are 482 by 482 elements in the output matrix to fill. A regression is run on each pair and the residuals from each are retained to test for stationarity via an augmented Dickey Fuller test with intercept only and 16 lags following the suggestion of Schwert (2002)⁵². Normally, statistical tests provide binary decisions, over accepting or rejecting the null hypothesis, using pre-chosen significance levels. From each paired regression, the p-value is saved so multiple graphs can be formed with various threshold values. As the p-value decreases, the graph loses edges and becomes more sparse.

The result of step one, via either method, is a square matrix with the number of rows and columns equal to 482. The correlation approach produces a symmetric matrix so only the lower triangular matrix is preserved. The cointegration approach is not symmetric but the upper and lower triangular matrices contain similar information so the the lower triangular is preserved while the upper triangular matrix is discarded. The main diagonal is also not needed because that represents a stock's relationship with itself.

⁵²This article is cited with various dates of working paper versions.

The resulting matrix is called the adjacency matrix. Adjacency matrices define the graph structure of nodes and edges. Each row and column represents a node. The values in row i and column j are used to define edge length. A high absolute value correlation produces a short edge length. A low p-value also produces a short edge length. The plotting algorithm used is a type of force-directed approach, specifically the Fruchterman-Reingold force-directed algorithm, implemented in Python. This approach tries to achieve relative closeness based on edge weights and as few edge crossings as possible for a clear display. Nodes will be more centered in the graph for a greater number of high-weight connections and toward the edges for fewer.

The network produced from the above adjacency matrices is fully connected, meaning that every node connects to every other node since there were no perfect zero correlations or perfect one p-values. Threshold values are chosen to prune the weaker connections, leaving only the stronger connections. As the threshold value changes, the network structure changes as well.

Step three analyzes the each type of network using network graph statistics. To conclude, each network is run through a predictive machine learning algorithm. The algorithm takes x-y coordinates from the graph as data to predict industry sector. Support vector machine (SVM) and k-nearest neighbors classification algorithms are considered for this problem. The SVM approach uses a radial basis function kernel, a standard approach in the field. The KNN approach uses a k of one given the number of nodes to edges. For both algorithms a k-folds cross validation approach is used to improve model fit.

4.4 Empirical Results

4.4.1 Step One - Data Preparation

Figure 4.1 shows the stacked correlation density plots of correlations. Spearman is displayed for robustness since there was some concern that non-spherical paired relationships might skew Pearson. A center of mass around .3 confirms a positive amount of market comovement. The mean amount of correlation for every stock is positive. The correlation network literature uses Pearson exclusively.

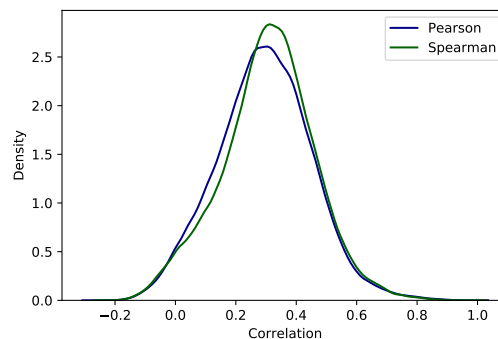


Figure 4.1: Kernel density plots of Pearson and Spearman Correlation

The distribution of p-values from the Engle-Granger approach reveals a much different structure. Figure 4.2 looks much more like a uniform distribution compared to the near-normality of the correlation distributions. The large number of pairs with low p-values indicates that the graph will remain quite connected even as the threshold value moves towards zero. Compare that to the density of Pearson correlations. As the correlation grows above the mean, the number of pairs decreases rapidly.

4.4.2 Step Two - Network Graphs

Figure 4.3 displays a network graph produced by correlation values without imposing any threshold. The fully connected graph contains 115,921 edges connecting every node to every other node. Only by imposing thresholds to restrict edges to only the

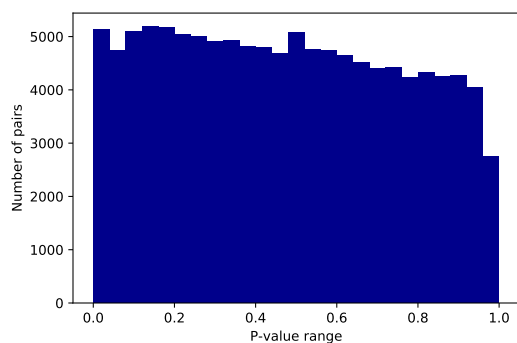


Figure 4.2: A histogram of Augmented Dickey Fuller p-values on the residuals of pairwise regressions.

stronger connections does market structure reveal itself. Correlation and cointegration approaches produce extremely different graphs. As the threshold value grows, the correlation approach is distinguished by segmentation, largely by industry, as seen in Figure 4.4. The cointegration approach produces no such industry segmentation.

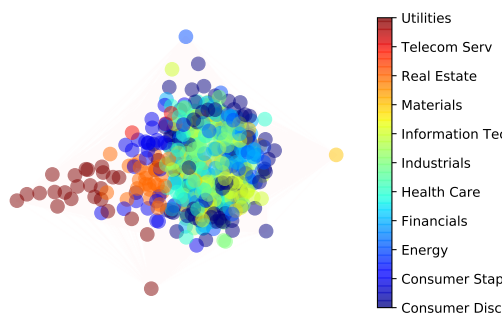


Figure 4.3: A spring network of Pearson correlations for the S&P 500.

Figure 4.5 shows a correlation graph with .7 threshold and a cointegration graph with .975 (.025 p-value) threshold side by side. With a much lower threshold value, the correlation approach produces a graph with 249 nodes and 912 edges while the higher threshold cointegration approach produces a graph with 480 nodes and 1,762 edges. No nodes have been pruned from this approach. One of the driving factors creating such different graphs is that imposing a positive threshold on the correlation graph removes all negatively correlated stocks leaving only strongly positive

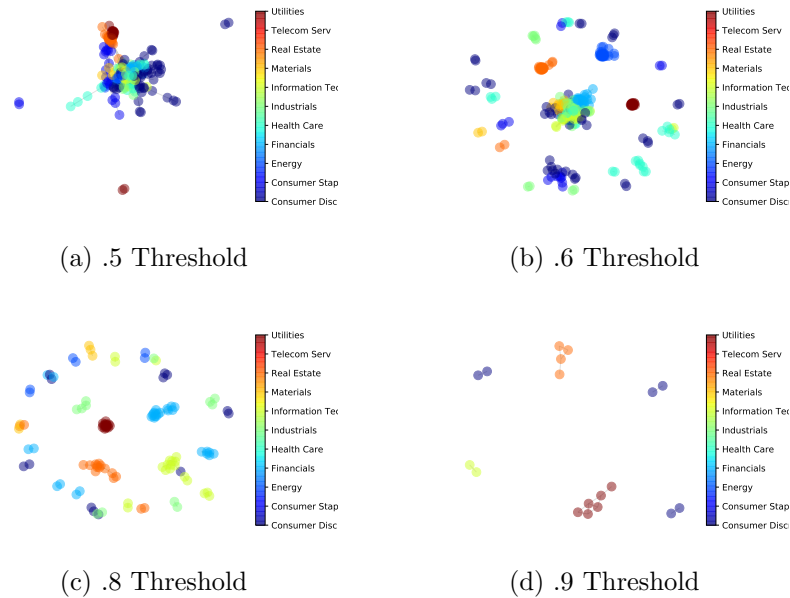


Figure 4.4: Positive correlation network graphs at four threshold values.

correlations. Increasing the threshold on the cointegration graph does not remove negatively correlated stocks. Positively correlated stocks are much more likely to come from the same industry, forming highly connected clusters explaining the high edge to node ratio observed. The regression-based approaches in the literature (cointegration, Granger causality) do not make this distinction, but has important consequences for analysis and use in trading systems.

The hedge strategy for positively rather than negatively correlated stocks will be opposite from each other. Furthermore, drawing a network graph without distinguishing negative from positive correlations reveals very little, since the graph remains nearly complete for most thresholds, including very high values.

The cointegrated graph, rather than resembling a positive threshold correlation graph, more resembles a correlation graph with a negative threshold value, one that removes any correlations greater than that level. Figure 4.6 shows an example of such a negative threshold correlation graph with threshold value -0.1 . The circular layout is a natural consequence of having more pairs and fewer clusters, as the plotting algorithm gives centrality to nodes with more connections. Here, most nodes have about

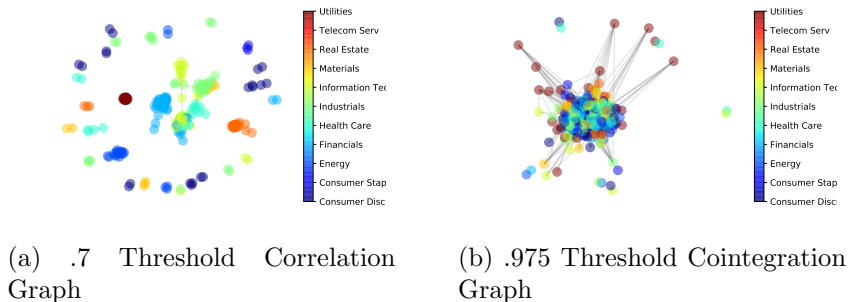


Figure 4.5: Correlation and Cointegration Approach Graphs

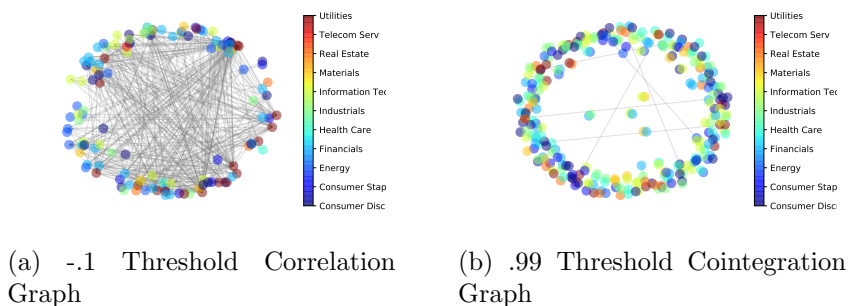


Figure 4.6: Negative Correlation and Cointegration Approach Graphs

the same number of connections. The number of nodes in the negative correlation graph is 129 to 398 edges while in the cointegration graph with a .99 threshold is 280 nodes to 140 edges.

Edge Density

Figures 4.7 shows how edge density changes in positive correlation and negative correlation network graphs as the threshold value moves. Figure 4.8 is the same graph for a cointegration-formed network. The positive and cointegration approaches look similar with much slower edge density decay in the cointegration graph³³. These two graphs show a network becoming less complete as the threshold value increases. The negative correlation shows that for a range of threshold values, the graph is nearly complete. This is significant and part of the reason why the cointegration graph re-

³³The spike at the right side of the positive correlation graph should be disregarded as is caused by Google's two tickers in the S&P.

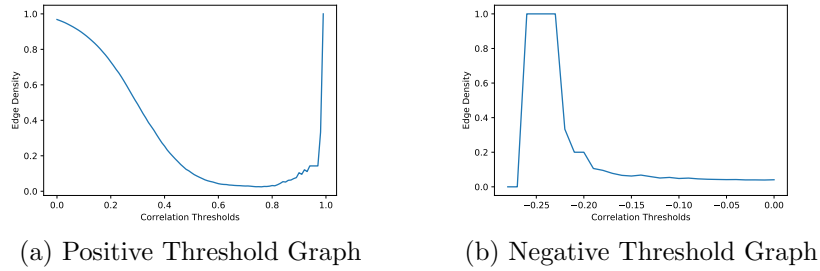


Figure 4.7: Edge density as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.

mains much *more* complete as threshold increases. Negative correlation in the S&P 500 exists between stocks from different industries. Graphs formed from negative correlations tend not to cluster because of these cross-market connections.

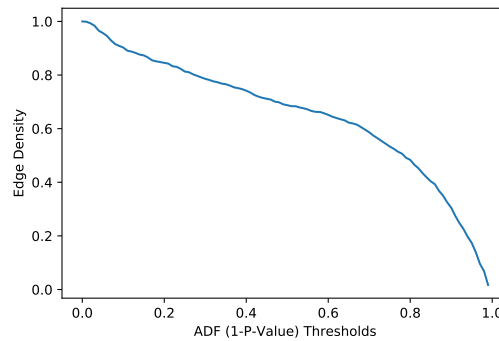


Figure 4.8: Cointegration Graph: Edge density as the threshold value changes on the cointegration-formed network.

Node Degree

Figures 4.9 and 4.10 show how mean node degree changes in the correlation and cointegration network graphs, respectively. These graphs depict the average number of adjacent edges a node claims. Mirroring node density, the positive and cointegration graphs look similar, but the cointegration graph decays more slowly. This is due to a number of negatively correlated stocks not found in the positive correlation graph but found at high thresholds in the cointegration graph.

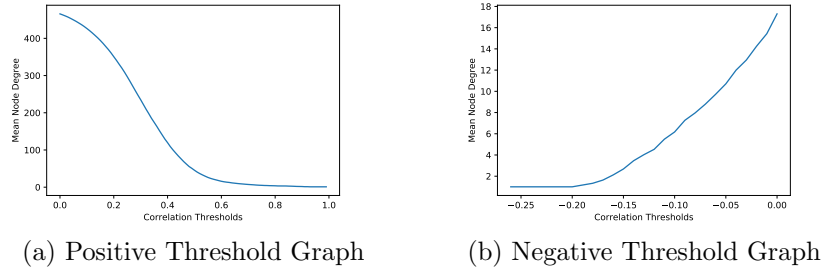


Figure 4.9: Node degree as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.

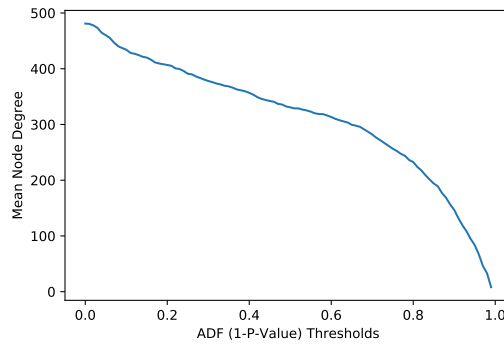


Figure 4.10: Cointegration Graph: Mean node degree as the threshold value changes on the cointegration-formed network.

Clustering Coefficient

Figures 4.11 and 4.12 show how the mean clustering coefficient changes in the correlation and cointegration network graphs as threshold changes. The positive correlation graph shows the transition from one large cluster, a value of one, to a graph that becomes more segmented. The cointegration graph looks very similar. The negative correlation graph becomes more clustered as the threshold increases.

Connected Components

Figures 4.13 and 4.14 show the component number as threshold values increase in the correlation and cointegration networks, respectively. The connected components number is a count of how many unconnected clusters are in the graph. With low threshold values, all nodes are connected and the number of connected components is

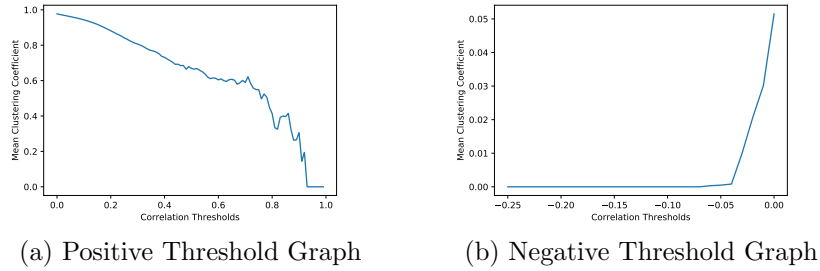


Figure 4.11: Clustering coefficient as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.

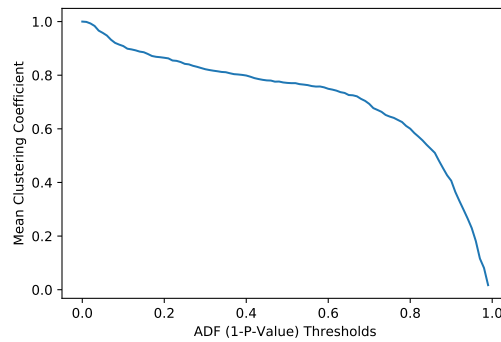


Figure 4.12: Mean cluster coefficient as the threshold value changes on the cointegration-formed network.

one. In the positive correlation graph, at a threshold in the .6 to .8 range, there are around 30 connected components. This number is more than the number of industries which means that industry sectors are not wholly descriptive of the market.

An important note is that cointegration graph remains completely connected as nodes are stripped away with an increasing threshold. The fact that the cointegration graph does not segment like the correlation graphs makes it less useful to study market structure.

Power Law Distribution

This section estimates the parameters of the power law distribution and assesses that distributions fit to this data. Figure 4.15 shows histograms of node density with graph insets for four threshold values of positive correlation graphs. The alpha measure falls

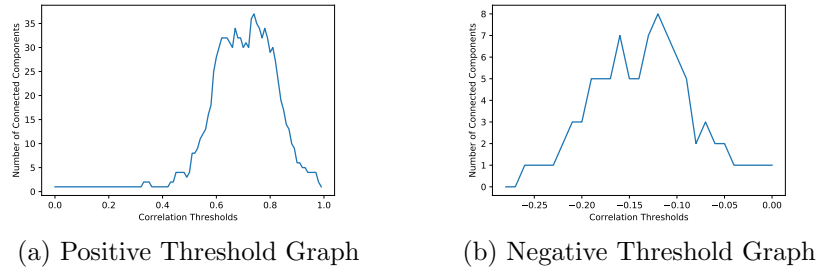


Figure 4.13: Number of connected components as the threshold value increases in absolute value on the positive (left) and negative (right) correlation-formed networks.

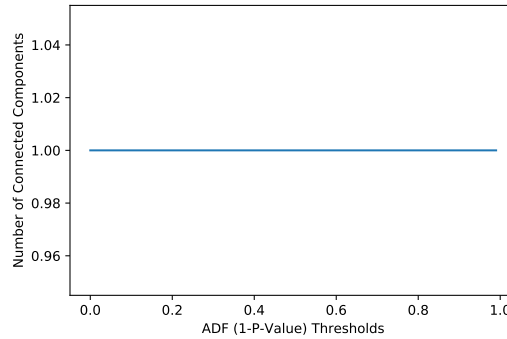


Figure 4.14: The number of connected components as the threshold value changes on the cointegration-formed network.

in between two to three for the mid-threshold values. This indicates a good fit by the power law distribution.

4.4.3 Step Three - Machine Learning Classification

Figure 4.16 portrays the confusion matrix heatmap generated by a $k=1$ KNN algorithm applied on a correlation graph with a .7 cutoff. Rows in the matrix represent actual classification of industry and columns represent what the trained algorithm predicted. High numbers in the main diagonal correspond to high prediction accuracy. The KNN algorithm perfectly predicts six of nine industries. The worst predicted industry, Consumer Discretionary, still has a two thirds successful prediction rate.

Figure 4.17 shows the confusion matrix heatmap generated by a $k=1$ KNN algorithm applied on a cointegration network with .005 ADF p-value cutoff. This

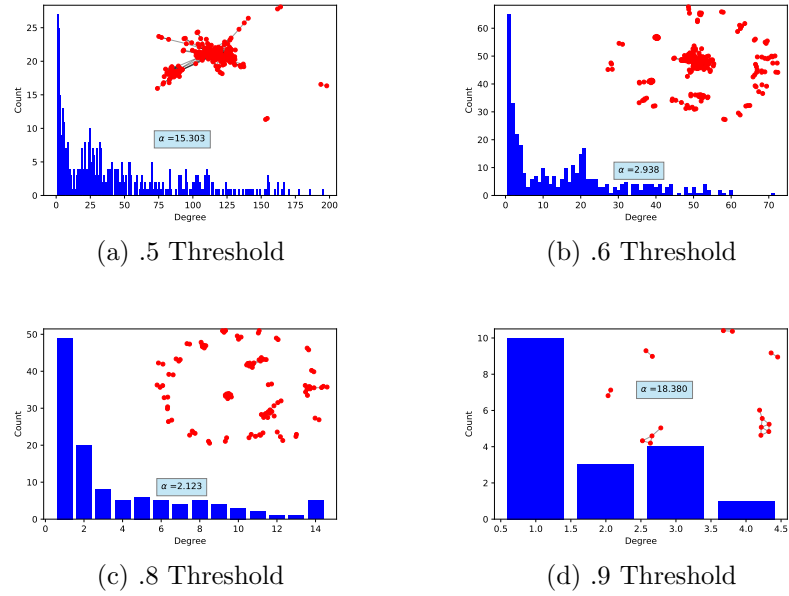


Figure 4.15: Positive correlation network graph histograms of degree with graph insets.

approach is far less able to predict industry than the correlation approach. This is because cointegration is not limited to cases where like follows like. The cointegration space includes far more exotic stationary-residual producing pairs.

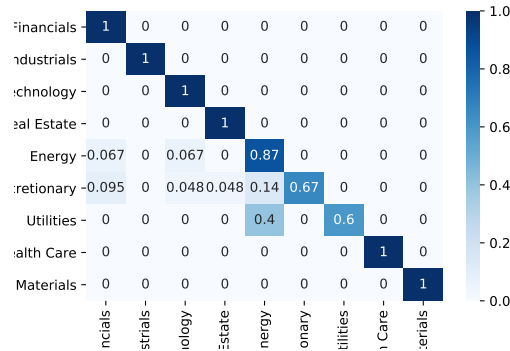


Figure 4.16: KNN confusion matrix heatmap for correlation network.

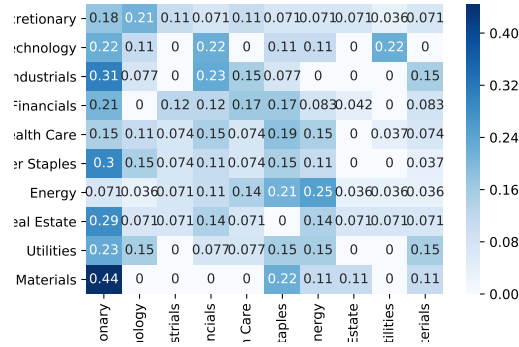


Figure 4.17: KNN confusion matrix heatmap for cointegration network.

4.5 Simulation

This section⁴⁴ simulates two pairs of cointegrated relationships based on Equations 4.2 and 4.3 (repeated below). The same underlying error processes are used in each run. The only difference between the two y_{1t} runs is β which is $.5$ and $-.5$, respectively in each run. The result is shown in Figure 4.18. The two runs of y_{1t} have correlations with y_{2t} of $.95$ and $-.95$, respectively. Therefore, the y_{1t} series generated with $\beta = .5$ would be included in a correlation approach but $\beta = -.5$ would not. Both series would be included in most of the cointegration approaches discussed above.

$$y_{1t} = \beta y_{2t} + u_{1t}, \text{ where } u_{1t} \sim I(0)$$

$$y_{2t} = y_{2t-1} + u_{2t}, \text{ where } u_{2t} \sim I(0)$$

⁴⁴This section is meant more as a proof of concept to demonstrate what the other papers are missing regarding cointegration. This is not meant to be rigorous.

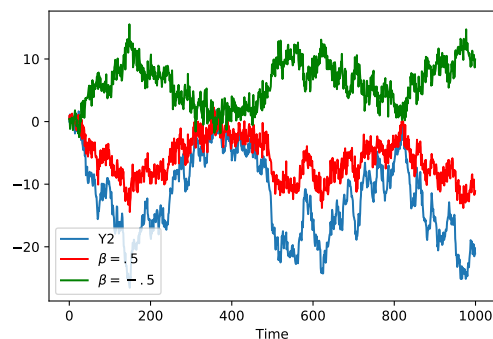


Figure 4.18: Simulated Cointegrated Series.

The point of this exercise is to show that both procedures in the literature to form networks are missing something. The correlation approach, through setting purely positive thresholds, misses cointegrated series with negative correlations. The literature using a cointegration or regression-based approach have not addressed any differences with a correlation approach. The cointegration approach includes a much larger set of relationships including negative correlations. Without accounting for this, analysis and trading strategies based on a cointegration approach will be misleading.

4.6 Conclusion

The threshold market graph literature is mostly split on the adoption of a correlation or a regression approach. The correlation approach is to form adjacency matrices from Pearson correlations between pairs of returns. The regression approach includes the Engle-Granger cointegration method, Granger causality, and the Johansen cointegration approach. This article is the first to compare the two broad categories of approaches. The network structure created differs substantially in each so this difference needs to be better understood rather than discarded.

The regression approaches allows for positive or negative relationships compared with the correlation approach which contains edges between stocks that move together

(the literature only considers positive threshold graphs). This has important implications for the uncovered market structure and any applications derived thereafter. A common application of this literature is making the pairs trading literature more robust by investigating network strategies to uncover suitable pairs. A co-moving pair requires a different hedging strategy than a repelling pair.

This chapter is the only paper to consider negative threshold correlation graph. The result of this approach is completely different from positive threshold graphs. The cointegration approach produces a graph that is a mix of negative and positive correlation threshold approaches. Since negative correlation graphs tend to connect industries while positive correlation graphs tend to cluster nodes by industry, the cointegration approach does not produce graphs that create subgraphs. Cointegration graphs remain very connected as the threshold pushes to high levels. The implication of this is that underlying market structure is not revealed as well as with correlation graphs.

Finally, this article employs machine learning and simulation for proof of concept robustness checks. Machine learning, via k-nearest neighbors, proves that the cointegration approach does not reveal industry structure. Whereas the correlation approach produces a graph structure where nearest neighbors is very accurate at predicting industry. The simulation approach proves that cointegrated relationships can just as easily be negatively correlated.

BIBLIOGRAPHY

- Acs, Zoltan J. and David B. Audretsch (1989). “Small-Firm Entry in US Manufacturing”. In: *Economica* 56, pp. 255–265.
- Alperovych, Yan and Georges Hübner (Oct. 2013). “Incremental impact of venture capital financing”. en. In: *Small Business Economics* 41.3, pp. 651–666.
- Angrist, Joshua D. and Jorn-Steffen Pischke (2009). *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press.
- Arvanitis, Spyros and Tobias Stucki (Apr. 2014). “The Impact of Venture Capital on the Persistence of Innovation Activities of Start-Ups”. In: *Small Business Economics* 42.4, pp. 849–870.
- Audretsch, David B., Werner Bönte, and Prashanth Mahagaonkar (Oct. 2012). “Financial signaling by innovative nascent ventures: The relevance of patents and prototypes”. In: *Research Policy* 41, pp. 1407–1421.
- Backes-Gellner, Uschi and Arndt Werner (June 2007). “Entrepreneurial Signaling via Education: A Success Factor in Innovative Start-Ups”. en. In: *Small Business Economics* 29.1-2, pp. 173–190.
- Balasubramanian, Natarajan and Jagadeesh Sivadasan (Feb. 2011). “What Happens When Firms Patent? New Evidence from U.S. Economic Census Data”. In: *Review of Economics and Statistics* 93, pp. 126–146.
- Baum, Joel A. C. and Brian S. Silverman (May 2004). “Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups”. In: *Journal of Business Venturing*. Evolutionary approaches to entrepreneurship: Honoring Howard Aldrich 19, pp. 411–436.
- Beck, Thorsten and Asli Demirguc-Kunt (2006). “Small and medium-size enterprises: Access to finance as a growth constraint”. In: *Journal of Banking and Finance* 30.11, 2931–2943.
- Bellucci, Andrea, Alexander Borisov, and Alberto Zazzaro (2010). “Does gender matter in bank-firm relationships? Evidence from small business lending”. In: *Journal of Banking and Finance*.
- Berg-Utby, Terje, Roger Sorheim, and L. Oystein Widding (Jan. 2007). “Venture Capital Funds: Do They Meet the Expectations of Portfolio Firms?” In: *Venture Capital* 9, pp. 23–41.

- Bertoni, Fabio, Annalisa Croce, and Diego D’Adda (Oct. 2010). “Venture Capital Investments and Patenting Activity of High-Tech Start-Ups: A Micro-econometric Firm-Level Analysis”. In: *Venture Capital* 12.4, pp. 307–326.
- Bhattacharya, Utpal, Alexander Borisov, and Xiaoyun Yu (Apr. 2015). “Firm Mortality and Natal Financial Care”. In: *Journal of Financial and Quantitative Analysis* 50. Firm Mortality and Natal Financial Care - Volume 50 Issue 1-2 - Utpal Bhattacharya, Alexander Borisov, Xiaoyun Yu, pp. 61–88.
- Blinder, Alan S. (1973). “Wage Discrimination: Reduced Form and Structural Estimates”. In: *The Journal of Human Resources*.
- BLS (2016). *Business establishment age*.
- Boginski, Vladimir, Sergiy Butenko, and Panos M. Pardalos (2003). “Innovations in Financial and Economic Networks”. In: ed. by Anna Nagurney. UK: Edward Elgar Publishing. Chap. Network approach for the Russian stock market, 9–45.
- Boyer, Tristan and Régis Blazy (Apr. 2014). “Born to be alive? The survival of innovative and non-innovative French micro-start-ups”. en. In: *Small Business Economics* 42.4, pp. 669–683.
- Brown, Nerissa C. and Michael D. Kimbrough (Sept. 2011). “Intangible Investment and the Importance of Firm-Specific Factors in the Determination of Earnings”. In: *Review of Accounting Studies* 16, pp. 539–573.
- Brush, Candida, Nancy Carter, Elizabeth Gatewood, Patricia Greene, and Myra Hart (2001). “The Diana Project: Women business owners and equity capital: The myths dispelled”. In: *Kauffman center for Entrepreneurial Leadership*.
- Bu, Hui, Wenjin Tang, and Junjie Wu (2019). “Time-varying comovement and changes in comovement structure in the Chinese stock market: A causal network method”. In: *Economic Modelling*.
- Buddelmeyer, Hielke, Paul H. Jensen, and Elizabeth Webster (Apr. 2010). “Innovation and the determinants of company survival”. en. In: *Oxford Economic Papers* 62.2, pp. 261–285.
- Burke, Andrew E, Felix R FitzRoy, and Michael A Nolan (2002). “Self-employment Wealth and Job Creation: The Roles of Gender, Non-pecuniary Motivation and Entrepreneurial Ability”. In: *Small Business Economics* 19.3, 255–270.
- Busenitz, Lowell W. and Jay B. Barney (1997). “Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making”. In: *Journal of Business Venturing* 12.1.

- Cao, Jerry and Po-Hsuan Hsu (June 2011). “The Informational Role of Patents in Venture Capital Financing”. SSRN Scholarly Paper. Rochester, NY.
- Cassar, Gavin (2004). “The financing of business start-ups”. In: *Journal of Business Venturing*.
- Cefis, Elena and Orietta Marsili (Aug. 2011). “Born to flip. Exit decisions of entrepreneurial firms in high-tech and low-tech industries”. en. In: *Journal of Evolutionary Economics* 21.3, pp. 473–498.
- Charness, Gary and Uri Gneezy (2012). “Strong Evidence for Gender Differences in Risk Taking”. In: *Journal of Economic Behavior & Organization* 83.1, 50–58.
- Chemmanur, Thomas J., Karthik Krishnan, and Debarshi K. Nandy (2011). “How Does Venture Capital Financing Improve Efficiency in Private Firms? A Look Beneath the Surface”. In: *The Review of Financial Studies* 24, pp. 4037–4090.
- Chen, Shi, Wei Meng, and Haitian Lu (Apr. 2018). “Patent as a Quality Signal in Entrepreneurial Finance: A Look beneath the Surface”. In: *Asia-Pacific Journal of Financial Studies* 47.2, pp. 280–305.
- Chou, Ting-Kai, Jia-Chi Cheng, and Chin-Chen Chien (May 2013). “How useful is venture capital prestige? Evidence from IPO survivability”. In: *Small Business Economics* 40, pp. 843–863.
- Clauset, Aaron, Cosma Rohilla Shalizi, and M.E.J. Newman (2007). “Power-law distributions in empirical data”. In: *arXiv:0706.1062v1 [physics.data-an]*.
- Cockburn, Iain M. and Megan J. MacGarvie (May 2011). “Entry and Patenting in the Software Industry”. In: *Management Science* 57, pp. 915–933.
- Cole, Rebel A. and Tatyana Sokolyk (June 2018). “Debt financing, survival, and growth of start-up firms”. In: *Journal of Corporate Finance* 50, pp. 609–625.
- Coleman, Susan and Alicia Robb (2009). “A comparison of new firm financing by gender: evidence from the Kauffman Firm Survey data”. In: *Small Business Economics*.
- Colombelli, Alessandra, Jackie Krafft, and Marco Vivarelli (Aug. 2016). “To be born is not enough: the key role of innovative start-ups”. In: *Small Business Economics; Dordrecht* 47, pp. 277–291.
- Colombo, Massimo G. and Samuele Murtinu (Feb. 2017). “Venture Capital Investments in Europe and Portfolio Firms’ Economic Performance: Independent Ver-

- sus Corporate Investors”. In: *Journal of Economics & Management Strategy* 26, pp. 35–66.
- Comino, Stefano and Clara Graziano (Nov. 2015). “How many patents does it take to signal innovation quality?” In: *International Journal of Industrial Organization* 43, pp. 66–79.
- Constandtinidis, Christina, Annie Cornet, and Simona Asandei (2006). “Financing of women-owned ventures: The impact of gender and other owner- and firm-related variables”. In: *Venture Capital* 8:02, pp. 133–157.
- Conti, Annamaria, Jerry Thursby, and Marie Thursby (2013). “Patents as Signals for Startup Financing”. In: *Journal of Industrial Economics* 61, pp. 592–622.
- Conti, Annamaria, Marie Thursby, and Frank T. Rothaermel (2013). “Show Me the Right Stuff: Signals for High-Tech Startups”. In: *Journal of Economics and Management Strategy* 22, pp. 341–364.
- Croson, Rachel and Uri Gneezy (2009). “Gender Differences in Preferences”. In: *Journal of Economic Literature* 47.2, 448–474.
- Cumming, Douglas and Sofia Johan (July 2008). “Information Asymmetries, Agency Costs and Venture Capital Exit Outcomes”. In: *Venture Capital* 10, pp. 197–231.
- Czarnitzki, Dirk and Kornelius Kraft (May 2010). “On the Profitability of Innovative Assets”. In: *Applied Economics* 42, pp. 1941–1953.
- Czarnitzki, Dirk and Andrew A. Toole (Feb. 2011). “Patent Protection, Market Uncertainty, and R&D Investment”. In: *Review of Economics and Statistics* 93, pp. 147–159.
- Davidsson, Per and Scott R. Gordon (Nov. 2012). “Panel studies of new venture creation: a methods-focused review and suggestions for future research”. In: *Small Business Economics* 39, pp. 853–876.
- Dickey, David A. and Wayne A. Fuller (1979). “Distribution of the estimators for autoregressive time series with a unit root”. In: *Journal of the American Statistical Association*.
- Eddleston, Kimberly A., Jamie J. Ladge, Cheryl Mitteness, and Lakshmi Balachandra (2016). “Do you see what I see? Signaling effects of gender and firm characteristics on financing entrepreneurship”. In: *Entrepreneurial Theory and Practice*.

- Engel, Dirk (Sept. 2004). “The Performance of Venture-Backed Firms: The Effect of Venture Capital Company Characteristics”. In: *Industry and Innovation; Sydney* 11, pp. 249–263.
- Engel, Dirk and Max Keilbach (Mar. 2007). “Firm-level implications of early stage venture capital investment — An empirical investigation”. In: *Journal of Empirical Finance* 14, pp. 150–167.
- Engle, Robert F. and C.W.J. Granger (1987). “Cointegration and error correction: Representation, estimation, and testing”. In: *Econometrica*.
- Ewens, Michael and Richard R. Townsend (2017). “Can Access to Capital Explain the Entrepreneurship Gender Gap?” <https://ssrn.com/abstract=2953011>.
- Fabrizi, Simona, Steffen Lippert, Pehr-Johan Norback, and Lars Persson (Sept. 2013). “Venture Capitalists and the Patenting of Innovations”. In: *Journal of Industrial Economics* 61, pp. 623–659.
- Fama, Eugene F. (1965). “The behavior of stock market prices”. In: *The Journal of Business*.
- Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist (Mar. 2017). *What is a Patent Worth? Evidence from the U.S. Patent “Lottery”*. Working Paper.
- Fehr-Duda, Helga, Mauele De Gennaro, and Renate Schubert (2006). “Gender, financial risk, and probability weights”. In: *Theory and Decision*.
- Forbes, Kristin J. and Roberto Rigobon (2002). “No contagion, only interdependence: Measuring stock market comovements”. In: *The Journal of Finance*.
- Frid, Casey J (2009). “Acquiring Financial Resources to Form New Ventures: Pecking Order Theory and the Emerging Firm”. In: *Frontier of Entrepreneurship Research* 29, pp. 1–16.
- Fuentes-Fuentes, María del Mar, Matilde Ruiz-Arroyo, Ana Maria Bojica, and Friederike Welter (2017). “Innovativeness and business relationships in women-owned firms: The role of gender stereotypes”. In: *Canadian Journal of Administrative Sciences*.
- Gao, Xiangyun, Shupe Huang, Xiaoqi Sun, Xiaoqing Hao, and Feng An (2018). “Modelling cointegration and Granger causality network to detect long-term equilibrium and diffusion paths in the financial system”. In: *Royal Society Open Science*.
- Gartner, William B., Casey J. Frid, and John C. Alexander (Oct. 2012). “Financing the Emerging Firm”. In: *Small Business Economics* 39, pp. 745–761.

- Goel, Rajeev K. and Devrim Goktepe-Hulten (Aug. 2013). “Nascent Entrepreneurship and Inventive Activity: A Somewhat New Perspective”. In: *Journal of Technology Transfer* 38, pp. 471–485.
- Goniadis, Iraklis and Nikos C. Varsakelis (Aug. 2012). “Factors affecting patentees to start-up new venture: evidence from Greece”. In: *International Journal of Entrepreneurial Behavior & Research* 18, pp. 544–558.
- Gorman, Michael and William A. Sahlman (July 1989). “What do venture capitalists do?” en. In: *Journal of Business Venturing* 4.4, pp. 231–248.
- Graham, Stuart J. H. and Ted M. Sichelman (2010). “Patenting by Entrepreneurs: An Empirical Study”. In: *Michigan Telecommunications and Technology Law Review* 117, pp. 111–180.
- Granger, C.W.J. and P. Newbold (1974). “Spurious regressions in econometrics”. In: *Journal of Econometrics*.
- Guo, Bing, Yun Lou, and David Pérez-Castrillo (June 2015). “Investment, Duration, and Exit Strategies for Corporate and Independent Venture Capital-Backed Start-Ups”. In: *Journal of Economics & Management Strategy* 24, pp. 415–455.
- Gurmu, Shiferaw and Fidel Perez-Sebastian (2008). “Patents, R&D and Lag Effects: Evidence from Flexible Methods for Count Panel Data on Manufacturing Firms”. In: *Empirical Economics* 35, pp. 507–526.
- Haeussler, Carolin, Dietmar Harhoff, and Elisabeth Mueller (Oct. 2014). “How patenting informs VC investors – The case of biotechnology”. In: *Research Policy* 43, pp. 1286–1298.
- Hahn, Guangsug, Kwanho Kim, and Joon Yeop Kwon (Feb. 2017). “Startup Financing with Patent Signaling under Ambiguity”. In: *Asia-Pacific Journal of Financial Studies* 46, pp. 32–63.
- Hall, Bronwyn and Josh Lerner (Sept. 2009). *The Financing of R&D and Innovation*. Tech. rep. Cambridge, MA: National Bureau of Economic Research.
- Hall, Bronwyn H. (Dec. 2004). “Exploring the Patent Explosion”. In: *The Journal of Technology Transfer* 30.
- Hall, Bronwyn H (2018). “Is there a role for patents in the financing of new innovative firms?” In: *NBER Working Paper No. 24370*, p. 39.
- Hechavarría, Diana M., Charles H. Matthews, and Paul D. Reynolds (Jan. 2016). “Does start-up financing influence start-up speed? Evidence from the panel study

- of entrepreneurial dynamics”. en. In: *Small Business Economics* 46.1, pp. 137–167.
- Hellmann, Thomas and Manju Puri (2000). “The Interaction between Product Market and Financing Strategy: The Role of Venture Capital”. In: *Review of Financial Studies* 13.
- Hemel, Daniel Jacob and Lisa Larrimore Ouellette (2013). “Beyond the Patents-Prizes Debate”. en. In: *Texas L. Rev.* 92, pp. 303–382.
- Hirukawa Masayuki and Ueda Masako (Oct. 2011). “Venture capital and innovation: which is first?” In: *Pacific Economic Review* 16, pp. 421–465.
- Hoenen, Sebastian, Christos Kolympiris, Wilfred Schoenmakers, and Nicholas Kalaitzandonakes (July 2014). “The diminishing signaling value of patents between early rounds of venture capital financing”. In: *Research Policy* 43, pp. 956–989.
- Hoenig, Daniel and Joachim Henkel (June 2015). “Quality signals? The role of patents, alliances, and team experience in venture capital financing”. In: *Research Policy* 44, pp. 1049–1064.
- Hottenrott, Hanna, Bronwyn H. Hall, and Dirk Czarnitzki (2016). “Patents as Quality Signals? The Implications for Financing Constraints on R&D”. In: *Economics of Innovation and New Technology* 25.
- Hsu, Dan K., J. Michael Haynie, Sharon A. Simmons, and Alexander McKelvie (Jan. 2014). “What Matters, Matters Differently: A Conjoint Analysis of the Decision Policies of Angel and Venture Capital Investors”. In: *Venture Capital* 16, pp. 1–25.
- Hsu David H. and Ziedonis Rosemarie H. (Jan. 2013). “Resources as Dual Sources of Advantage: Implications for Valuing Entrepreneurial-Firm Patents”. In: *Strategic Management Journal* 34, pp. 761–781.
- Huang, Wei-Qiang, Xin-Tian Zhuang, and Shuang Yao (2009). “A network analysis of the Chinese stock market”. In: *Physica A*.
- Huyghebaert, Nancy and Linda M. Van de Gucht (Jan. 2007). “The Determinants of Financial Structure: New Insights from Business Start-ups”. In: *European Financial Management* 13, pp. 101–133.
- Ivanov, Vladimir I. and Fei Xie (Mar. 2010). “Do Corporate Venture Capitalists Add Value to Start-Up Firms? Evidence from IPOs and Acquisitions of VC-Backed Companies”. In: *Financial Management* 39, pp. 129–152.

- Jain, Bharat A. and Omesh Kini (Dec. 2000). “Does the Presence of Venture Capitalists Improve the Survival Profile of IPO Firms?” en. In: *Journal of Business Finance & Accounting* 27.9&10, pp. 1139–1183.
- Ji, Qiang, Elie Bouri, Rangan Gupta, and David Roubaud (2018). “Network causality structures among Bitcoin and other financial assets: A directed acyclic graph approach”. In: *The Quarterly Review of Economics and Finance*.
- Jones, F.L. and Jonathan Kelley (1984). “Decomposing differences between groups: A cautionary note on measuring discrimination”. In: *Sociological Methods and Research*.
- Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision Under Risk”. In: *Econometrica* 47.2.
- Kline, Brendan and Elie Tamer (2011). “Some interpretation of the linear-in-means model of social interactions”. Presented at AEA annual meetings 2012.
- Knockaert, Mirjam, Bart Clarysse, and Andy Lockett (Oct. 2010). “Are Technology VC Investors a Distinct Species on the Investment Market?” In: *Venture Capital* 12, pp. 267–283.
- Kolympiris, Christos and Nicholas Kalaitzandonakes (July 2013). “The Geographic Extent of Venture Capital Externalities on Innovation”. In: *Venture Capital* 15, pp. 199–236.
- Kortum, Samuel and Josh Lerner (2000). “Assessing the Contribution of Venture Capital to Innovation”. In: *The RAND Journal of Economics* 31, pp. 674–692.
- Lahr, Henry and Andrea Mina (Feb. 2016). “Venture capital investments and the technological performance of portfolio firms”. In: *Research Policy* 45, pp. 303–318.
- Lerman, Celia (2015). “Patent Strategies of Technology Startups: An Empirical Study”. In:
- Lerner, Joshua (1994). “The Importance of Patent Scope: An Empirical Analysis”. In: *RAND Journal of Economics* 25, pp. 319–333.
- A new identification method of the market downturns in Chinese A-shares using liquidity network* (2018). 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA). Shanghai, China: IEEE.

- Liu, Sheen, Peter Woodlock, Howard Qi, and Yan Alice Xie (Dec. 2006). “Cash Reserve and Venture Business Survival Probability”. In: *The Journal of Entrepreneurial Finance & Business Ventures; Greenwich* 11, pp. 123–136.
- Lockett, Andy, Gordon Murray, and Mike Wright (Aug. 2002). “Do UK venture capitalists still have a bias against investment in new technology firms”. In: *Research Policy* 31.
- Long, Clarisa (2002). “Patent Signals”. In: *The University of Chicago Law Review* 69, pp. 625–679.
- MacKinnon, James G. and Halbert White (1985). “Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties”. In: *Journal of Econometrics*, pp. 305–325.
- Macmillan, Ian C., Robin Siegel, and P. N. Subba Narasimha (Dec. 1985). “Criteria used by venture capitalists to evaluate new venture proposals”. In: *Journal of Business Venturing* 1, pp. 119–128.
- Manigart, Sophie, Katleen Baeyens, and Wim Van Hyfte (Apr. 2002). “The Survival of Venture Capital Backed Companies”. In: *Venture Capital* 4, pp. 103–124.
- Mann, Ronald J. and Thomas W. Sager (Mar. 2007). “Patents, venture capital, and software start-ups”. In: *Research Policy* 36, pp. 193–208.
- Manski, Charles F. (1993). “Identification of endogenous social effects: The reflection problem”. In: *The Review of Economic Studies* 60, pp. 531–542.
- Mantegna, R.N. (1999). “Hierarchical structure in financial markets”. In: *The European Physical Journal B*.
- Mazzoleni, Roberto and Richard R Nelson (July 1998). “The benefits and costs of strong patent protection: a contribution to the current debate”. In: *Research Policy* 27, pp. 273–284.
- Megginson, William L. and Kathleen A. Weiss (July 1991). “Venture Capitalist Certification in Initial Public Offerings”. In: *The Journal of Finance* 46, pp. 879–903.
- Michaelas, Nicos, Francis Chittenden, and Panikkos Poutziouris (Mar. 1999). “Financial Policy and Capital Structure Choice in U.K. SMEs: Empirical Evidence from Company Panel Data”. In: *Small Business Economics* 12.2, pp. 113–130.

- Mina, Andrea, Henry Lahr, and Alan Hughes (Aug. 2013). “The demand and supply of external finance for innovative firms”. en. In: *Industrial and Corporate Change* 22.4, pp. 869–901.
- Minniti, Maria (2009). *Gender issues in entrepreneurship*. Now Publishing.
- Minniti, Maria and William Bygrave (1999). “The microfoundations of entrepreneurship”. In:
- Moskowitz, Tobias J. and Annette Vissing-Jørgensen (2002). “The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?” In: *The American Economic Review* 92, pp. 745–778.
- Munari, Federico, Maria Cristina Odasso, and Laura Toschi (2011). “Patent-backed Finance”. In: *The Economic Valuation of Patents Methods and Applications*. Cheltenham, UK: Edward Elgar Publishing.
- Munari, Federico and Laura Toschi (Sept. 2015). “Do patents affect VC financing? Empirical evidence from the nanotechnology sector”. In: *International Entrepreneurship and Management Journal* 11, pp. 623–644.
- Myers, Stewart C. (July 1984). “The Capital Structure Puzzle”. en. In: *The Journal of Finance* 39.3, pp. 574–592.
- Myers, Stewart C. and Nicholas S. Majluf (June 1984). “Corporate financing and investment decisions when firms have information that investors do not have”. In: *Journal of Financial Economics* 13.
- Namaki, A., A.H. Shirazi, R. Raei, and G.R. Jafari (2011). “Network analysis of a financial market based on genuine correlation and threshold method”. In: *Physica A*.
- Neeley, Lynn and Howard Van Auken (2010). “Differences between female and male entrepreneurs’ use of bootstrap financing”. In: *Journal of Developmental Entrepreneurship*.
- Nelson, Charles R. and Charles I. Plosser (1982). “Trends and random walks in macroeconomic time series”. In: *Journal of Monetary Economics*.
- Nelson, Julie A. (2015). “Are Women Really More Risk-Averse than Men?” In: *Journal of Economic Surveys* 29.3, 566–585.
- Nerkar, Atul and Scott Shane (Nov. 2003). “When do start-ups that exploit patented academic knowledge survive?” In: *International Journal of Industrial Organization* 21, pp. 1391–1410.

- Nobi, Ashadun, Sungmin Lee, Doo Hwan Kim, and Jae Woo Lee (2014). “Correlation and network topologies in global and local stock indices”. In: *Physics Letters A*.
- Noel, Michael and Mark Schankerman (Sept. 2013). “Strategic Patenting and Software Innovation”. In: *The Journal of Industrial Economics* 61, pp. 481–520.
- Oaxaca, Ronald (1973). “Male-Female Wage Differentials in Urban Labor Markets”. In: *International Economic Review*.
- Orser, Barbara J., Allan L. Riding, and Kathryn Manley (2006). “Women entrepreneurs and financial capital”. In: *Entrepreneurship Theory and Practice*.
- Ou, Charles and George W. Haynes (Oct. 2006). “Acquisition of Additional Equity Capital by Small Firms – Findings from the National Survey of Small Business Finances”. In: *Small Business Economics* 27.2-3, pp. 157–168.
- Paglia, John K. and Maretno A. Harjoto (2014). “The effects of private equity and venture capital on sales and employment growth in small and medium-sized businesses”. In: *Journal of Banking and Finance* 47, 177–197.
- Parker, Simon C. and Yacine Belghitar (2006). “What Happens to Nascent Entrepreneurs? An Econometric Analysis of the PSED”. In: *Small Business Economics*.
- Paul, Stuart, Geoff Whittam, and Janette Wyper (Feb. 2007a). “The pecking order hypothesis: does it apply to start-up firms?” In: *Journal of Small Business and Enterprise Development* 14.1, pp. 8–21.
- (Feb. 2007b). “The pecking order hypothesis: does it apply to start-up firms?” In: *Journal of Small Business and Enterprise Development* 14.1, pp. 8–21.
- Peneder, Michael (Apr. 2010). “The Impact of Venture Capital on Innovation Behaviour and Firm Growth”. In: *Venture Capital* 12, pp. 83–107.
- Popov, Alexander and Peter Roosenboom (July 2012). “Venture capital and patented innovation: evidence from Europe”. en. In: *Economic Policy* 27.71, pp. 447–482.
- Puri, Manju and Rebecca Zarutskie (Dec. 2012). “On the Life Cycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms”. In: *The Journal of Finance* 67, pp. 2247–2293.
- Ragozzino, Roberto and Dane P. Blevins (Sept. 2016). “Venture-Backed Firms: How Does Venture Capital Involvement Affect Their Likelihood of Going Public or Being Acquired?” In: *Entrepreneurship Theory and Practice* 40, pp. 991–1016.

- Rassenfosse, Gaetan de (Sept. 2012). “How SMEs Exploit Their Intellectual Property Assets: Evidence from Survey Data”. In: *Small Business Economics* 39, pp. 437–452.
- Rassenfosse, Gaétan de, Alfons Palangkaraya, and Elizabeth Webster (Sept. 2016). “Why do patents facilitate trade in technology? Testing the disclosure and appropriation effects”. In: *Research Policy* 45.7, pp. 1326–1336.
- Robb, Alicia and David Robinson (2010). “The Capital Structure Decisions of New Firms”. In: *NATIONAL BUREAU OF ECONOMIC RESEARCH*.
- Robb, Alicia and John Wolken (2002). “Firm, owner, and financing characteristics: Differences between female- and male-owned small businesses”. In: *Board of Governors of the Federal Reserve System (US)* 18. Finance and Economics Discussion Series.
- Rocca, Maurizio La, Tiziana La Rocca, and Alfio Cariola (July 2011). “Capital Structure Decisions During a Firm’s Life Cycle”. en. In: *Small Business Economics* 37.1, pp. 107–130.
- Saparito, Patrick, Amanda Elam, and Candida Brush (2012). “Bank-firm relationships: Do perceptions vary by gender?” In: *Entrepreneurial Theory and Practice*.
- Scellato, Giuseppe (Jan. 2007). “Patents, Firm Size and Financial Constraints: An Empirical Analysis for a Panel of Italian Manufacturing Firms”. In: *Cambridge Journal of Economics* 31, pp. 55–76.
- Scherr, Frederick C., Timothy F. Sgrue, and Janice B. Ward (1993). “Financing the small firm start-up: Determinants of debt use”. In: *The Journal of Entrepreneurial Finance* 3, pp. 17–36.
- Schwert, G. William (2002). “Tests for unit roots: A Monte Carlo investigation”. In: *Journal of Business and Economic Statistics*.
- Shane, Scott and Rakesh Khurana (June 2003). “Bringing Individuals Back In: The Effects of Career Experience on New Firm Founding”. In: *Industrial and Corporate Change* 12, pp. 519–543.
- Shane, Scott, Edwin A. Locke, and Christopher J. Collins (June 2003). “Entrepreneurial motivation”. In: *Human Resource Management Review* 13, pp. 257–279.
- Singh, Vikkram, Eduardo Roca, and Bin Li (2018). “Cointegration networks in stock markets”. In: *Applied Economics Letters*.

- Åstebro, Thomas and Irwin Bernhardt (July 2003). “Start-up financing, owner characteristics, and survival”. en. In: *Journal of Economics and Business* 55.4, pp. 303–319.
- Storey, D.J. (1993). “New Firm Growth and Bank Financing”. In: *Small Business Economics* 6, pp. 139–150.
- Stucki, Tobias (Feb. 2014). “Success of start-up firms: the role of financial constraints”. In: *Industrial and Corporate Change* 23. test the impact of financial constraints on the success of start-ups, the persistence of this impact, pp. 25–64.
- Tang, Yong, Jason Jie Xiong, Yong Luo, and Yi-Cheng Zhang (2019). “How do the global stock markets influence one another? Evidence from finance big data and Granger causality directed network”. In: *International Journal of Electronic Commerce*.
- Tian, Xuan (July 2011). “The causes and consequences of venture capital stage financing”. In: *Journal of Financial Economics* 101, pp. 132–159.
- Tondo-Kramer, Kyle (2009). “Increasing Access to Startup Financing through Intellectual Property Secularization Comment”. eng. In: *John Marshall Journal of Computer and Information Law* 27, bi–646.
- Tong, Chen, Jing Chen, and Mike J. Buckle (2015). “A network visualization approach and global stock market integration”. In: *International Journal of Finance and Economics*.
- Triguero, Angela, David Corcoles, and Maria C. Cuerva (Dec. 2014). “Persistence of Innovation and Firm’s Growth: Evidence from a Panel of SME and Large Spanish Manufacturing Firms”. In: *Small Business Economics* 43, pp. 787–804.
- Tsvetkova, Alexandra, Jean-Claude Thill, and Deborah Strumsky (Oct. 2014). “Metropolitan innovation, firm size, and business survival in a high-tech industry”. en. In: *Small Business Economics* 43.3, pp. 661–676.
- Turner, Grant and Vladyslav Sushko (2018). “What risks do exchange-traded funds pose”. In: *Financial Stability Review*.
- Tversky, Amos and Daniel Kahneman (1986). “Rational Choice And The Framing Of Decisions”. In: *The Journal of Business* 59.4.
- Useche, Diego (Oct. 2014). “Are patents signals for the IPO market? An EU–US comparison for the software industry”. In: *Research Policy* 43, pp. 1299–1311.

- Vanacker, Tom R. and Sophie Manigart (July 2010). “Pecking order and debt capacity considerations for high-growth companies seeking financing”. In: *Small Business Economics* 35, pp. 53–69.
- Vaznyte, Egle and Petra Andries (2019). “Entrepreneurial orientation and start-ups’ external financing”. In: *Journal of Business Venturing*.
- Vizgunov, A., B. Goldengorin, V. Kalyagin, A. Koldanov, P. Koldanov, and P.M. Pardalos (2014). “Network approach for the Russian stock market”. In: *Computational Management Science*.
- Wang, Gang-Jin, Chi Xie, Shou Chen, Jiao-Jiao Yang, and Ming-Yan Yang (2013). “Random matrix theory analysis of cross-correlations in the US stock market: Evidence from Pearson’s correlation coefficient and detrended cross-correlation coefficient”. In: *Physica A*.
- Wen, Danyan, Chaoqun Ma, Gang-Jin Wang, and Senzhang Wang (2018). “Investigating the features of pairs trading strategy: A network perspective on the Chinese stock market”. In: *Physica A*.
- Winsborough, H.H. and Peter Dickinson (1971). “Components of negro-white income differences”. In: *Proceedings of the American Statistical Association*.
- Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.
- Wurgler, Jeffrey (2010). *On the economic consequences of index-linked investing*. Tech. rep. 16376. National Bureau of Economic Research.
- Yule, G. Udny (1926). “Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time series”. In: *Journal of the Royal Statistical Society*.

APPENDIX A

THREEFOLD OAXACA DECOMPOSITION

As mentioned above, the threefold Oaxaca decomposition is performed on the pooled model. This section presents the mathematical decomposition and the full estimation results. See Winsborough and Dickinson 1971; Jones and Kelley 1984 for a more detailed reference.

$$\hat{Y}_{NP} = \hat{\beta}_{NP} X_{NP} \quad (5)$$

$$\hat{Y}_P = \hat{\beta}_P X_P \quad (6)$$

The separate group coefficient estimates in the regression above allow one to derive a number of useful decompositions to explain the difference in average outcome, shown in Eqn. 7. This suggests that firms without a patent experience a difference in log profits of 4.54. A threefold decomposition (Winsborough and Dickinson 1971; Jones and Kelley 1984) portions out that difference into three sources, endowments, coefficients, and interaction. This is shown (although not derived) in Eqn. 8.

$$\begin{aligned} \Delta \bar{Y} &= \bar{Y}_{NP} - \bar{Y}_P, \\ &= 2.14 - (-2.39), \\ &= 4.54 \end{aligned} \quad (7)$$

$$\Delta \bar{Y} = \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' \hat{\beta}_P}_{\text{endowments}} + \underbrace{\bar{\mathbf{X}}_P' (\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{coefficients}} + \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' (\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{interaction}} \quad (8)$$

The first term in Eqn. 8 represents the portion of $\Delta \bar{Y}$ that can be explained by differences in average endowments of the independent variables. For example, were the non-patenting group more experienced and the profit responds positively to experience, then part of the observed difference in profit would be attributable to the non-patenting group's extra endowment of experience. The second term represents the portion of $\Delta \bar{Y}$ that can be explained by differences in response (coefficients). For example, were the non-patenting group more rewarded for experience, endowment held constant between the groups, then a portion of the observed difference in profit would be attributable to the extra return in experience. The third term is the interaction between the two first effects. Differences in endowments and coefficients are

observed simultaneously. That simultaneous presence is captured here. For example, if there were no difference in the endowment of experience, the interaction effect would be zero. Likewise, if there were no difference the coefficient on experience, then the interaction effect would also be zero.

Eqn. 7, $\Delta\bar{Y}$, portions out among the three effects as shown in Eqn. 9. This suggests that the endowment effect explains about 42% of the observed difference in profit, the coefficient effect explains about 87%, and the interaction effect explains about -29% of the difference.

$$\begin{aligned}
\Delta\bar{Y} = & \\
& \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' \hat{\beta}_P}_{\text{endowments}} = 2.08^{***}(.274) + \\
& \underbrace{\bar{\mathbf{X}}_P' (\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{coefficients}} = 3.89^{***}(.432) + \\
& \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' (\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{interaction}} = -1.43^{***}(.273)
\end{aligned} \tag{9}$$

Differences in endowment are shown in Eqn. 10. These differences feed the endowment effect, a further breakdown of which is presented below in Eqn. 11 which demonstrates how the total endowment effect is portioned among the individual terms in the model. Equity, Employees, and Assets play the most significant roles. Among them Equity is by far the most significant in magnitude. This suggests that the endowment effect is mostly comprised of the endowment differences in equity financing. The patenting group has 3.11 greater share of mean equity. The 2.20 (of 2.08 total endowment effect) statistically significant endowment effect of equity says that the non-patenting group greatly benefits from having less equity which is negatively related to profits.

$$\begin{aligned}
\Delta\bar{\mathbf{X}} = \bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P = & \\
\Delta X_{Equity} = & -3.11, \\
\Delta X_{Other} = & -3.56, \\
\Delta X_{Employees} = & -3.73, \\
\Delta X_{Assets} = & -1.26, \\
\Delta X_{Credit} = & .09 \\
\Delta X_{Year} = & -.03
\end{aligned} \tag{10}$$

$$\begin{aligned}
& \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)' \hat{\beta}_P}_{\text{endowments}} = \\
& (\Delta X_{Equity}) \hat{\beta}_{Equity,P} = 2.20^{***}(.203) + \\
& (\Delta X_{Other}) \hat{\beta}_{Other,P} = .57^{**}(.253) + \\
& (\Delta X_{Employees}) \hat{\beta}_{Employees,P} = -.17^*(.087) + \\
& (\Delta X_{Assets}) \hat{\beta}_{Assets,P} = -.47^{***}(.160) + \\
& (\Delta X_{Credit}) \hat{\beta}_{Credit,P} = -.01(.038) + \\
& (\Delta X_{Year}) \hat{\beta}_{Year,P} = -.03(.060) + \\
& (\Delta X_{Intercept}) \hat{\beta}_{0,P} = 0.0(0.0)
\end{aligned} \tag{11}$$

The column of differences in coefficients is presented in Eqn. 12. Eqn. 13 presents the components of the full coefficients effect with corresponding p-value and standard errors. The largest differences in magnitude and significance are the intercept and equity, identifying larger values in the non-patenting groups. The difference in intercept is interpreted as an unexplained difference between the two groups and accounts for most of the coefficients effect. The difference in equity, on the other hand, suggests that patent owning firms are less efficient with equity financing than non-patent owning firms. Whether this is due to management or to different equity structure imposed by the investor is unclear from this data.

$$\begin{aligned}
\Delta \hat{\beta} &= \hat{\beta}_{NP} - \hat{\beta}_P = \\
& \Delta \hat{\beta}_{Equity} = .46^{***}(.003), \\
& \Delta \hat{\beta}_{Other} = -.07^{***}(.003), \\
& \Delta \hat{\beta}_{Employees} = .01^{***}(.001), \\
& \Delta \hat{\beta}_{Assets} = .12^{***}(.004), \\
& \Delta \hat{\beta}_{Credit} = -.62^{***}(.014), \\
& \Delta \hat{\beta}_{Year} = -.30^{***}(.101), \\
& \Delta \hat{\beta}_0 = 4.28^{***}(.064)
\end{aligned} \tag{12}$$

$$\begin{aligned}
& \underbrace{\bar{\mathbf{X}}_P'(\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{coefficients}} = \\
& \bar{X}_{Equity,NP}(\Delta\hat{\beta}_{Equity}) = 1.75^{***}(.219)+ \\
& \bar{X}_{Other,NP}(\Delta\hat{\beta}_{Other}) = -.52(.560)+ \\
& \bar{X}_{Employees,NP}(\Delta\hat{\beta}_{Employees}) = .06(.151)+ \\
& \bar{X}_{Assets,NP}(\Delta\hat{\beta}_{Assets}) = .1.33(1.226)+ \\
& \bar{X}_{Credit,NP}(\Delta\hat{\beta}_{Credit}) = -1.83*(1.183)+ \\
& \bar{X}_{Year,NP}(\Delta\hat{\beta}_{Year}) = -1.19*(.673)+ \\
& \bar{X}_{Intercept}(\Delta\hat{\beta}_0) = 4.28^{***}(1.923)
\end{aligned} \tag{13}$$

The components of the interaction effect are presented in Eqn. 14. This effect accounts for the simultaneous presence of both the endowment and coefficient effects. The only significant result is on Equity and mitigates the total contribution of Equity on the full 4.54 Log Profit difference to 2.52 (2.20+1.75-1.43). Both groups react negatively to equity financing, however the patent owning group more so. With each dollar of equity, the forces driving this observed effect cause the patenting and non-patenting groups' profits to further diverge.

$$\begin{aligned}
& \underbrace{(\bar{\mathbf{X}}_{NP} - \bar{\mathbf{X}}_P)'(\hat{\beta}_{NP} - \hat{\beta}_P)}_{\text{interaction}} = \\
& (\Delta X_{Equity})(\Delta\hat{\beta}_{Equity}) = -1.43^{***}(.187)+ \\
& (\Delta X_{Other})(\Delta\hat{\beta}_{Other}) = .24(.254)+ \\
& (\Delta X_{Employees})(\Delta\hat{\beta}_{Employees}) = -.04(.084)+ \\
& (\Delta X_{Assets})(\Delta\hat{\beta}_{Assets}) = -.15(.138)+ \\
& (\Delta X_{Credit})(\Delta\hat{\beta}_{Credit}) = -.05(.042)+ \\
& (\Delta X_{Year})(\Delta\hat{\beta}_{Year}) = -.01(.024)+ \\
& (\Delta X_{Intercept})(\Delta\hat{\beta}_0) = 0.0(0.0)
\end{aligned} \tag{14}$$

It is important to remember that Equity does not account for the full observed difference so there are other fundamental differences at play. The intercept accounts for 4.28 of the 4.54 Log Profit difference meaning that unobservable (or unaccounted by this model) variables are a large component of what is observed in $\Delta\bar{Y}$.

APPENDIX B

LOG EQUITY RESULTS

Table A1 contains the continuous outcome variable results. As above, all standard errors are clustered on firm and are robust to heteroskedasticity and serial correlation and employ the same small sample correction as Stata. As above, column I is a correlation and represents the sum of signal and appropriative effects. Column II is the causal effect of obtaining a patent on external equity and represents just the appropriative value. Patents ownership is associated with a statistically significant 184% increase in external equity financing. Column II indicates that this is not caused by the patent but rather unobservable characteristics of the firm or owner.

Table A1: Dependent Variable: Log of External Equity Financing

	I	II
Patent	1.84*** (.131)	.12 (.369)
R^2	.19	.03
Observations	17,906	17,906
F-test	143.9***	23.1***
Note:	*p<0.1; **p<0.05; ***p<0.01	