**Title**
Essays on Venture Capital Financing and Entrepreneurship



**Author**
Lu, Sizhu

**Publication Date**
2022

Peer reviewed|Thesis/dissertation

# Essays on Venture Capital Financing and Entrepreneurship

by

Sizhu Lu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Amir Kermani, Chair
Professor Nancy Wallace
Professor Peng Ding

Summer 2022

Essays on Venture Capital Financing and Entrepreneurship

# Abstract

Essays on Venture Capital Financing and Entrepreneurship

by

Sizhu Lu

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Amir Kermani, Chair

Entrepreneurship and innovation are believed to be the driving forces of the US economy. Many new startups rely on venture investors to acquire adequate capital and make it possible to start and run their businesses. In addition to capital, early-round investors such as venture capitalists also provide expertise, knowledge of management and product markets, and essential resources to the entrepreneurs, thus playing a crucial role in their growth. This dissertation aims to understand several essential questions on venture capital financing, entrepreneurship, and other venture investments.

The first chapter evaluates the network effects in venture capital financing in the US market and provides possible channels. We conduct two empirical studies and estimate the positive network effects with rigorous addresses on the endogeneity problem. First, we simultaneously model the endogenous network formation and outcome models with unobserved confounding variables incorporated in both models. The estimation result shows a positive network effect of 0.127 when the outcome of interest is the performance of the venture capitalists, measured by the success rate defined as the ratio between the number of successful deals over the total number of deals this venture capitalist ever made. The result indicates that for a venture capitalist in the network, a one percentage point increase in the average success rate of its peers will lead to a 0.127 percent increase in its success rate. We further decompose the observed VC network into two parts, one within the same industry sector and one across different sectors, and estimate their heterogeneous network effects. We find that peer effects in both networks are significantly positive, indicating not only does the sharing of information within an industry matter, but the sharing of resources and expertise also adds value to the connections across different industries. Our second study uses the quasi-experimental design, the generalized difference-in-differences method, to more intuitively illustrate the network effects by estimating the treatment effect of a venture capitalist's extremely successful event (an IPO) on the future performance of their connected venture capitalists. We find that once an IPO occurs in a venture capitalist's portfolio, more startups in their peers' portfolios

will receive new rounds of funding and achieve successful exits, and fewer startups will go bankrupt. In addition to these empirical studies, we build a theoretical model on the decision-making processes of venture capitalists to better understand the network effects and demonstrate the benefits and costs of syndication compared with standalone investments. We also test several model implications using our data. The syndicated deals have a 9.65% higher success rate, a 15.3% lower bankruptcy rate, and a 12.3% higher internal rate of return compared with standalone deals, aligning with the empirical implications of our theoretical model.

In the second chapter, we study the widespread common ownership in venture capital financing. Common ownership describes the phenomenon that competitors in the same industry share the same investors. Venture capitalists strategically build and actively manage their portfolio of startups, leading to the possibility that they invest simultaneously in multiple startups in the same industry and share information and ideas among them. While seemingly beneficial for startups to be commonly owned, the "horse race" investment strategy may hurt the startups at the same time. We evaluate the effect of common ownership on startups utilizing a matched-pair design as our identification strategy. Our results show that when a successful financing event occurs to a startup, its peers in the same common ownership pool have a 1.31% higher probability of getting new rounds of funding within 180 days, 2.4% higher within 365 days, and 3.37% higher within 730 days, compared with those that are not commonly owned. Moreover, the effects of common ownership are heterogeneous across industries.

Although conventional wisdom regards equity as the pivotal financing vehicle for new firms, in recent years we have observed unexpectedly active debt financing in the early-round startup financing market, considering relatively low rates of return and extremely high risks. The third chapter builds a theoretical model to study signaling in venture debt and capital and explains the seemingly puzzling existence of venture debt. We find that startups use venture debt as a good signal for financing. The cost of due diligence for venture capital firms is sufficiently high that they prefer to utilize this signal instead of investigating. We document and test four empirical implications of our model. First, startups with venture debt can get next-round funding faster than those without venture debt due to the signaling effect of venture debt. On average, a startup with venture debt reaches the next round of funding about a hundred days faster than those without venture debt. Second, startups with venture debt generally have significantly better long-term performance measured by their exit status. They have a lower probability of failure and a higher probability of going public. Third, conditional on the startups getting their next round of funding, those with venture debt have worse long-run performance compared to a counterfactual world without venture debt. Finally, the signaling effect is more substantial when venture capital investors suffer from more severe asymmetric information problems. Our empirical results align with the model predictions and are robust to various specifications.

To my parents.

# Contents

**Contents** ii

**List of Figures** iv

**List of Tables** v

**1 Network Effects in Venture Capital Financing** 1
- 1.1 Introduction . . . 1
- 1.2 Data and the construction of networks . . . 4
- 1.3 Stylized facts: centrality and performance . . . 6
- 1.4 Empirical study I: network formation and outcome models . . . 17
- 1.5 Empirical study II: generalized difference-in-differences method . . . 29
- 1.6 Model . . . 36
- 1.7 Conclusion . . . 48

**2 Common Ownership in Venture Capital Financing** 53
- 2.1 Introduction . . . 53
- 2.2 Data and the time-varying common ownership pool . . . 55
- 2.3 Stylized facts . . . 58
- 2.4 Identification and estimation . . . 62
- 2.5 Conclusion . . . 70

**3 Signaling in Venture Debt and Capital** 73
- 3.1 Introduction . . . 73
- 3.2 Model . . . 78
- 3.3 Data and empirical method . . . 84
- 3.4 Empirical results . . . 87
- 3.5 Robustness testing . . . 95
- 3.6 Conclusion . . . 101

**Bibliography** 107

**A Appendix to Chapter 1** 111

A.1 Proofs . . . . . . . . . . 111
A.2 Supplementary regression results . . . . . . . . . . 116

**B Appendix to Chapter 2** **129**
B.1 Supplementary regression results . . . . . . . . . . 129

**C Appendix to Chapter 3** **137**
C.1 Proofs . . . . . . . . . . 137
C.2 List of good venture capital firms . . . . . . . . . . 142

# List of Figures

1.1 A snapshot of the undirected network of a subsample of VCs . . . . . . . . . . . 8
1.2 Density function for posterior distribution: outcome model with exogenous networks 26
1.3 Density function for posterior distribution: network formation model with endogenous network . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
1.4 Density function for posterior distribution: outcome model with endogenous networks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
1.5 Decomposition of VC portfolios . . . . . . . . . . . . . . . . . . . . . . . . . 36
1.6 Trend of Gini coefficient of centrality . . . . . . . . . . . . . . . . . . . . . . 50
1.7 Trend of top 1% proportion of centrality . . . . . . . . . . . . . . . . . . . . 51

2.1 Pitchbook industry map . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
2.2 A sketch of common ownership definition . . . . . . . . . . . . . . . . . . . . 58
2.3 A snapshot of the time-varying common ownership pool . . . . . . . . . . . . 59
2.4 Cumulative incidence functions . . . . . . . . . . . . . . . . . . . . . . . . . 63

A.1 Expected returns as functions of VC's ability . . . . . . . . . . . . . . . . . . 113

# List of Tables

1.1 Summary statistics for outcomes and features of VCs . . . 7
1.2 Summary statistics for VC centrality . . . 11
1.3 Correlation coefficients between different centrality measures (undirected graph) 11
1.4 Correlation coefficients between different centrality measures (directed graph) . 12
1.5 Performance and degree centrality . . . 13
1.6 Performance and closeness centrality . . . 14
1.7 Performance and eigenvector centrality . . . 15
1.8 Performance and betweenness centrality . . . 16
1.9 Summary statistics for posterior distribution: outcome model with exogenous networks . . . 25
1.10 Summary statistics for posterior distribution: network model with endogenous networks . . . 27
1.11 Summary statistics for posterior distribution: outcome model with endogenous networks . . . 29
1.12 Heterogeneous network effects . . . 31
1.13 Main effects of IPO on future performance . . . 33
1.14 Peer effects of IPO on future performance . . . 35
1.15 Peer effects of IPO on future performance: common portfolio . . . 37
1.16 Peer effects of IPO on future performance: non-overlapping portfolio . . . 38
1.17 Success exit indicator and syndication indicator . . . 45
1.18 Failure indicator and syndication indicator . . . 46
1.19 Rate of return and syndication indicator . . . 47
1.20 Syndication and past centrality . . . 49

2.1 Performance and common ownership indicator . . . 61
2.2 Effects of common ownership on long-term performance . . . 67
2.3 Effects of common ownership on short-term financing behavior . . . 68
2.4 Heterogeneous effects on financing behavior across industries . . . 69
2.5 Subgroup effects on short-term financing behavior . . . 71
2.6 Effects on financing performance when a negative event occurs . . . 72

3.1 Summary statistics for funding rounds and startups . . . 88
3.2 Test of signaling effect of venture debt . . . 89

3.3 Effects of venture debt on long-term performance of startups . . . . . . . . . . . 92
3.4 Effects of venture debt on conditional long-term performance of startups . . . . 94
3.5 Intensity of signaling effects on severity of asymmetric information problem . . . 96
3.6 Effects of good venture capital investors on long-term performance . . . . . . . . 98
3.7 Effects of good venture capital investors on conditional long-term performance . 99
3.8 Effects of venture debt on long-term performance in the technology industry . . 100
3.9 Effects of venture debt on conditional long-term performance in the technology industry . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 101
3.10 Effects of venture debt on long-term performance of startups founded during 2001 - 2016 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102
3.11 Effects of venture debt on conditional long-term performance of startups founded during 2001 - 2016 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103
3.12 Intensity of signaling effects on severity of asymmetric information problem (experienced investors defined as total # of investment rounds $\geq$ p95) . . . . . . . 104
3.13 Intensity of signaling effects on severity of asymmetric information problem (experienced investors defined as total # of companies invested in $\geq$ p90) . . . . . 105
3.14 Intensity of signaling effects on severity of asymmetric information problem (experienced investors defined as total # of companies invested in $\geq$ p95) . . . . . 106

A.1 Main effects of IPO on future performance (2 years) . . . . . . . . . . . . . . . . 117
A.2 Main effects of IPO on future performance (3 years) . . . . . . . . . . . . . . . . 118
A.3 Peer effects of IPO on future performance (2 years) . . . . . . . . . . . . . . . . 119
A.4 Peer effects of IPO on future performance (3 years) . . . . . . . . . . . . . . . . 120
A.5 Peer effects of IPO on future performance: common portfolio (2 years) . . . . . 121
A.6 Peer effects of IPO on future performance: common portfolio (3 years) . . . . . 122
A.7 Peer effects of IPO on future performance: non-overlapping portfolio (2 years) . 123
A.8 Peer effects of IPO on future performance: non-overlapping portfolio (3 years) . 124
A.9 Peer effects of IPO on future performance (1 year, weighted regressions) . . . . . 126
A.10 Peer effects of IPO on future performance (2 years, weighted regressions) . . . . 127
A.11 Peer effects of IPO on future performance (3 years, weighted regressions) . . . . 128

B.1 Performance and common ownership indicator (use # of early rounds to define experienced investors) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 130
B.2 Subgroup effects on short-term financing behavior (ratio > p50) . . . . . . . . . 131
B.3 Effects of common ownership on long-term performance (1-1 matching) . . . . . 132
B.4 Effects of common ownership on short-term financing behavior (1-1 matching) . 133
B.5 Subgroup effects on short-term financing behavior (ratio > p75, 1-1 matching) . 134
B.6 Subgroup effects on short-term financing behavior (ratio > p50, 1-1 matching) . 135
B.7 Effects on financing performance when a negative event happens (1-1 matching) 136

# Acknowledgments

I am extremely grateful to my dissertation committee members, Amir Kermani, Nancy Wallace and Peng Ding, for their support and guidance. I am deeply indebted to Professor Amir Kermani for his invaluable guidance, support, and patience. He guided me through the rigorous process of doing research. The ideas of the first two chapters in this dissertation are also primarily inspired by conservations with him. I am always encouraged and touched by his deep passion for research and much love for his students. He is far beyond the best advisor I could ever imagine. I could not have undertaken this journey without the tremendous help from Professor Nancy Wallace. She has spent countless amounts of time advising my research projects and providing me with many helpful suggestions. This dissertation would not have been possible without her help. Professor Peng Ding guided me into the wonderful world of causal inference and taught me how to be a good thinker and researcher. I have learned so much from his guidance. I am also grateful to Professors Richard Stanton, Hoai-Luu Nguyen, Ulrike Malmendier and Christina Romer for their help on many of my projects.

I am also very fortunate to have many brilliant colleagues and officemates. I am very thankful to Can Huang, Yunbo Liu, Dayin Zhang, Xiao Yin, Mohammad Rezaei, Konhee Chang, Jaeyeon Lee, Jun Peng, Yao Zhao, and also too many others to list. I always learn much from our conservation, both intellectually and emotionally.

In addition, I would like to extend my sincere thank to the support from Fisher Center for Real Estate and Urban Economics. I am thankful to Thomas Chappelear for giving me exceptional editing help for my dissertation. I am also very grateful to Melissa Hacker and Lisa Marie Sanders for their amazing administrative assistance.

Finally, I also want to thank my dear friends Yue, Shihui, Sijia, and Chen for being there for me during many ups and downs in the past years. Thank my parents and boyfriend for their unconditional love and support. Thank you for always believing in me and supporting my decisions. Special thanks to my grandfather. I miss you so much.

# Chapter 1

# Network Effects in Venture Capital Financing

## 1.1 Introduction

Entrepreneurship and innovation are believed to be driving forces in the US economy, especially in recent years. Venture capitalists (VCs) not only act as investors who provide capital to startups, but also provide expertise, knowledge of management and product markets, and essential resources such as connections with other investors and startups, thus playing a crucial role in their growth. It is well documented in the literature that VC investment has been highly correlated with the rapid growth of many successful companies (Gompers and Lerner, 2001; Gompers et al., 2005; Lerner and Nanda, 2020).

Additionally, in the past few decades, the cooperation and syndication of VCs has become widespread (Lerner, 1994; Brander et al., 2002; Hochberg et al., 2007; Hochberg et al., 2010). VCs benefit from cooperation and connection with others for various reasons. Lerner (1994) argues that VCs are willing to share great investment opportunities with other VCs, hoping that those that benefit from these opportunities will share similar opportunities with them in the future. VCs also enjoy the diversification and risk-sharing advantages of syndication. Another proposed reason for VCs to be in favor of cooperation and co-investment is that they are able to pool the (imperfect) observed signals from startups to make investment decisions. As the general models of organizational design in Sah and Stiglitz (1986) suggest, satisfying the evaluation of two separate observed signals between individuals with veto power helps with the decision-making process. VCs can therefore achieve better screening and selection performance when they observe the signals received by other VCs. Additionally, different VCs are likely to have different expertise and skills. Therefore, VCs are likely to add more value by cooperating with each other in the growth process of the startups. Brander et al. (2002) use data on a small sample of Canadian VCs to show empirical evidence on this point: the syndicated investments have significantly higher rates of return than the standalone deals in their sample.

Given all possible gains from cooperation, it is natural to ask the question of whether VCs benefit from building connections and networking with others, and whether more connections improve their investment performance. Past literature provides empirical evidence of widespread co-investment by showing the strong correlation between VCs' centrality in the network and their performance. It also proposes possible reasons why VCs can benefit from building better connections and playing a more central role in the network. Castilla (2003) first studies this question by comparing the network structure of multiple VC firms in Silicon Valley and Route 128 in Massachusetts, using the large difference in their density to explain the greater success of Silicon Valley. The paper finds that VC firms in Silicon Valley are more densely connected compared with those in Route 128, and associates the higher growth and success rate of Silicon Valley with its higher network density. Hochberg et al. (2007) use past-deal information to construct a VC network model, and document that VCs with better networks have significantly better fund performance measured by successful exit rates. In a subsequent project, Hochberg et al. (2010) also argue that the incumbent VCs in the market cooperate in order to increase the entry cost of new VCs, thus reducing new entries. These incumbent VCs benefit from lower entry by paying lower prices for their VC investments.

This project aims to learn, estimate and understand the network between VCs in the US market, using the large-volume Pitchbook database on US startups and investors, including their features, financing history, and features of investments made by the VCs. We first construct a VC network model using past investment information and compute various measures of centrality and performance. We observe the positive correlation between network centrality and performance in the US VC market, after controlling for several observed features and fixed effects. To further explore this reduced form stylized fact, we conduct two sets of empirical studies, each of which addresses the endogeneity problem of the VCs' network formation and performance.

Our first empirical study resolves the problem in the past literature that the reduced form regressions are far from rigorously identified network effects due to endogeneity challenges. The observed VC network is extremely likely to be endogenously formed due to the strategic connection-making processes of the VCs. Analyses relying on problematic identification assumptions—such as that the observed network is as good as randomly assigned conditioning on observed features—are less than satisfactory due to problems such as homophily bias and contextual confounding problems (Manski, 1993; VanderWeele and An, 2013; Ogburn, 2018). VCs strategically choose which deal to syndicate and whom to cooperate with. It is highly probable that there are unobserved confounding variables that affect both the formation of connections and the performance of the VCs. For instance, VCs that share similar preferences for specific types of startups are more likely to co-invest in them. The preferences of VCs also matter greatly to their investment performance, leading to the unobserved confounding issue that preference is not directly measurable using our available data. In order to take the endogenous network formation process of VCs into consideration, we follow the econometric model suggested in Goldsmith-Pinkham and Imbens (2013) to model the network formation and the outcome simultaneously. By explicitly modeling these two, we are able to incorporate unobserved confounding variables in both models and estimate them as well as all other parameters of interest using Bayesian estimation procedures. Results from this empirical study show significantly positive network effects in our VC network after controlling for these unobserved confounding features. To better understand the mechanism through which the network is influential, we further decompose the network into two parts—connections within the same industry and across different industries—and estimate the heterogeneous network effects of these two networks. By separately estimating these two network effects and observing significantly positive effects in both cases, we are able to conclude that the positive peer effects are driven not only by the sharing of information within the same industry but also by adding more value to the startups through the sharing of expertise and cooperation across different industries.

Our second empirical study more intuitively illustrates the network effects by estimating the treatment effect of a VC's extremely successful event (an IPO) on the future performance of its own investment portfolio and its peers' portfolios. Using a generalized difference-in-differences method (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Goodman-Bacon, 2021), we find that once an IPO occurs in a VC's portfolio, the future performance of this VC significantly improves in their ability to raise funding, achieve successful exits, and avoid bankruptcy. Moreover, its peers also experience a significant improvement in these measures. When we further decompose the peers' portfolios into two sets, one commonly invested in by the successful VC, and one non-overlapping, we find that VCs not only benefit directly from the better performance of the overlapping investment but also improve the performance of their own non-overlapping portfolio when they are connected to peers that have IPOs.

Finally, we provide a theoretical model of VC investments to demonstrate the benefits and costs of syndication versus standalone deals and to better understand the positive network effects estimated in both empirical studies. By modeling the decision-making process of the VCs, we explicitly specify the two main reasons why VCs benefit from making connections and co-investing with each other: (1) They learn from each other's signals, so that conditioning on making an investment, the success rates of the syndicated deals are higher than the standalone deals; (2) VCs have more added value by cooperating with each other and sharing their expertise, leading to a higher rate of return of syndicated deals than standalone ones. However, cooperation also has a cost to a VC, as their stake in a startup is diluted with syndication. High-ability VCs receive more accurate signals from startups and suffer from less severe dilution problems when syndicating with others. The model concludes that whether syndication or standalone investment is the optimal choice for a VC depends on its ability. We then test the model predictions using our empirical data. The syndicated deals have a 9.65% higher success rate and a 15.3% lower bankruptcy rate compared with the standalone deals and have a 12.3% higher internal rate of return (IRR). We also discuss the reason that past centrality predicts the decision of a VC to choose between standalone investment and syndication, and provides theoretical explanations of empirical facts such as the increasing trend towards centrality inequality.

Our contribution to the literature is fourfold. First, this is the first paper in the VC literature that estimates the network effects by simultaneously modeling the network formation and the outcome. We directly tackle the endogeneity problems such as contextual bias and the homophily bias in the VC network. Second, the results on heterogeneous network effects shed light on the channel and mechanism of the network effects. Not only does the flow of information within the network of a single industry matter, but the sharing of expertise and resources through networks across different industries significantly improves VC performance as well. Thirdly, the quasi-experimental generalized difference-in-differences design provides intuitive empirical results for where the positive network effects come from. We are the first to study these types of peer effects in VC networks. Finally, we build a theory that illustrates the channels through which the network effects occur. This theory models the benefits and costs of syndication versus standalone investments and solves the optimal choice problem of VCs as a function of their features in their decision-making process.

The rest of this paper is organized as follows: Section 1.2 introduces the Pitchbook database we use for all the empirical studies in this paper and the construction of our VC network. Section 1.3 documents the significantly positive correlation between the centrality of a VC in the network and its investment performance and discusses in detail the endogeneity problem arguing why the reduced form results are not enough to prove positive network effects in the VC network. Section 1.4 tackles the endogeneity problem by simultaneously modeling the endogenous network formation and the outcome model. It also estimates the heterogeneous peer effects of the network both within the same industry and across different industries. Section 1.5 conducts a generalized difference-in-differences study to estimate and understand the positive network effects in a more intuitive way. Section 1.6 builds a theoretical model studying the investment decisions of VCs and where the positive network effects come from. Section 1.7 concludes.

## 1.2 Data and the construction of networks

### Data

The data we use in this study is from the Pitchbook database. Pitchbook is a Software as a Service (SaaS) company that delivers data, research, and technology covering private capital markets, including venture capital, private equity, and mergers and acquisitions transactions. Pitchbook uses machine learning and natural language processing technology to review publicly available resources and summarize the information in their database. Specifically, we use the information on entrepreneurial companies, their deals and the deal investors, and

VCs in our analyses. In this paper, we restrict our sample to the startups and VCs with headquarters in the US.

### Features of interest

Pitchbook provides various features of VCs, among which we are interested in the foundation year, the assets under management (in millions, AUM), the number of investment professionals, and the location (state) of the headquarters of the VCs. We augment the features of VCs with additional features calculated using the information on deals, e.g., the preferred industry sector and the Herfindahl index (HHI). The preferred industry sector is simply defined as the sector that each VC invested most deals in, and the Herfindahl index, which measures the concentration level of a VC's investment portfolio, is calculated using the formula $\mathrm{HHI} = \sum_{k=1}^{K} s_k^2$, where $s_k$ is the share of deals in industry $k$, and $K$ is the total number of industry sectors.

### Outcome of interest

The outcome of interest is the performance of the VCs. Successful VC-backed startups usually go public (ipo) or seek acquisition by large companies (acq). Therefore, two commonly used measures of VC performance (Gompers and Lerner, 1997; Hochberg et al., 2007) are: (1) the successful exit rate, which is defined as the proportion of successful exits among the total number of exits in the portfolio of all companies a VC invested in; (2) the success rate, which is defined as the ratio of the number of successful startups over the total number of startups a VC invested in. We define

$$\begin{aligned} \mathrm{success\_rate\_exit} &= \frac{\mathrm{ipo} + \mathrm{acq}}{\mathrm{ipo} + \mathrm{acq} + \mathrm{closed}}, \\ \mathrm{success\_rate} &= \frac{\mathrm{ipo} + \mathrm{acq}}{\mathrm{total\ \#\ of\ startups\ invested}}. \end{aligned}$$

### Summary statistics

Our final sample consists of 1,987 VCs. Table 1.1 reports the summary statistics of our sample. The average age of our sampled VCs is about 13 years. Nearly half of the VCs in our sample prefer investing in the Information Technology industry, and about 20 percent prefer the Healthcare industry. The mean of the HHI calculated at the industry sector level is 0.59, indicating the investment portfolios are not very concentrated in one industry sector. On average, each VC invests in about 11 startups in its entire history, while very few have successful exits, with only 0.17 startups having IPOs and 2.04 being acquired. About 1.45 startups go bankrupt and/or closed. The average successful exit rate is nearly 0.45, and the average success rate is 0.2. It is worth noting that there are a lot of missing values in the AUM feature, and fewer missing values in features such as the number of investment professions and found year. We calculate the logarithm of AUM due to its heavy right tail.

The summary statistics we report here are for the cross-sectional data. We also construct their time-varying version for some of the features calculated using deal information, e.g., preferred industry and HHI, and use them in the second empirical study.

### Construction of the network

To learn the network effect of the VCs, we construct the observed network among them using the Pitchbook data. Let $(V, W)$ denote the graph, where $V$ denotes the vertices and $W$ denotes the edges between nodes. In our network, the vertices $V$ are all VCs in the sample and $W_{ij}$ indicates the relationship between two VCs $i$ and $j$. Define $W_{ij} = 1$ if they are connected and $W_{ij} = 0$ otherwise. Based on the investment information using all deals made during 2007 - 2021 in the Pitchbook database, we construct two types of VC networks: undirected and directed networks. We restrict the deals to equity rounds. For the undirected graph, we define the edge between two VCs $i$ and $j$ is connected if and only if both VCs ever simultaneously invested in the same startup. For the directed graph, we define the edge from VC $i$ to VC $j$ as 1 if and only if they ever co-invested a deal in which $i$ is the leader and $j$ is the follower. Note that there are no connections between the followers if a deal has multiple followers. Figure 1.1 shows the observed undirected network of a small subsample of VCs. Each circle in the figure stands for a venture capitalist, with name on the circle. The size of the circles are proportional to their degree centrality, and the edges between VCs stand for their co-investment.

We build the networks using our sample data. In the undirected network, there are 1,987 nodes and 10,475 total connections. This is a very sparse network with a density of 0.005, which means only a tiny proportion of the potential connections are actually built. In the directed network, there are only 758 unisolated nodes and 838 edges among them. The number of connections in the directed graph is much less due to two facts: (1) We only consider VCs that co-invest in the startup in the same deal as connected when building the directed graph, while in the undirected graph, all VCs that ever simultaneously invested in the same startup are considered as connected, whether they invested in the same deal or different ones; (2) In the directed graph, there are no connections between followers of the same deal, while in the undirected graph all types of co-investment are treated as connected. The directed network is also very sparse, with a density of 0.0014. In the following sections, we use the graphs built in this way to measure the network effects of VC performance. In addition to the cross-sectional networks, we also build dynamic networks using time-varying deal information.

## 1.3 Stylized facts: centrality and performance

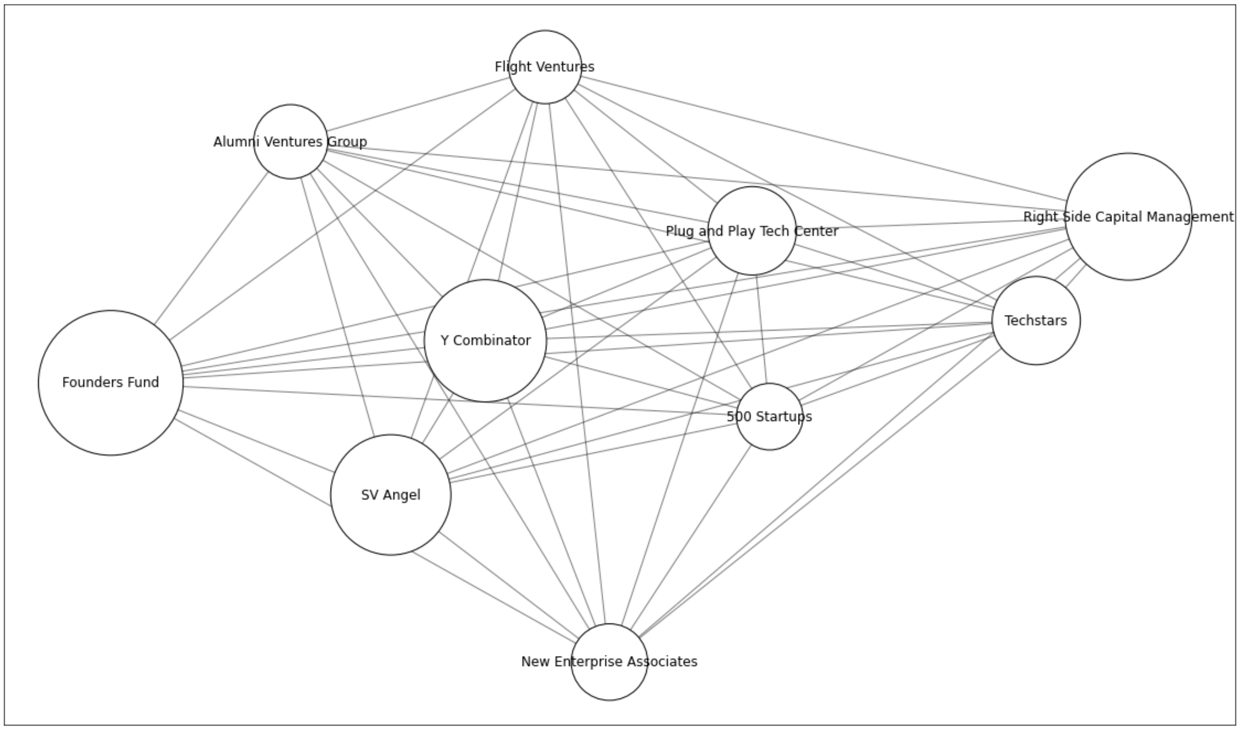
Past literature (Castilla, 2003; Hochberg et al., 2007) finds that VCs with better networks have significantly better fund performance. This motivates the research question on estimat-

Table 1.1: Summary statistics for outcomes and features of VCs

| | count | mean | std | p25 | median | p75 |
|---|---|---|---|---|---|---|
| AUM | 499 | 7459.3603 | 139357.7637 | 20 | 51.5 | 197.75 |
| log(AUM) | 499 | 4.1630 | 1.9845 | 2.9957 | 3.9416 | 5.2870 |
| # of investment professionals | 1628 | 3.4478 | 11.2973 | 1 | 2 | 4 |
| foundation year | 1556 | 2009.3380 | 11.7703 | 2006 | 2012 | 2016 |
| Business Products and Services | 1987 | 0.1022 | 0.3029 | 0 | 0 | 0 |
| Consumer Products and Services | 1987 | 0.1676 | 0.3736 | 0 | 0 | 0 |
| Energy | 1987 | 0.0156 | 0.1240 | 0 | 0 | 0 |
| Financial Services | 1987 | 0.0292 | 0.1684 | 0 | 0 | 0 |
| Healthcare | 1987 | 0.2043 | 0.4033 | 0 | 0 | 0 |
| Information Technology | 1987 | 0.4690 | 0.4992 | 0 | 0 | 1 |
| Materials and Resources | 1987 | 0.0121 | 0.1093 | 0 | 0 | 0 |
| hhi_industry_sector | 1987 | 0.5861 | 0.2639 | 0.3750 | 0.5 | 0.8253 |
| hhi_industry_group | 1987 | 0.4922 | 0.2829 | 0.2663 | 0.4028 | 0.625 |
| hhi_industry_code | 1987 | 0.3521 | 0.286 | 0.1429 | 0.25 | 0.5 |
| # of invested companies | 1987 | 10.6965 | 21.628 | 3 | 6 | 12 |
| # of ipo | 1987 | 0.1726 | 0.7753 | 0 | 0 | 0 |
| # of acq | 1987 | 2.0352 | 4.7599 | 0 | 1 | 2 |
| # of closed | 1987 | 1.4529 | 4.1403 | 0 | 0 | 1 |
| success_rate_exit | 1987 | 0.4497 | 0.4219 | 0 | 0.5 | 1 |
| success_rate | 1987 | 0.2 | 0.2465 | 0 | 0.1250 | 0.3333 |

*Notes:* This table summarizes several features of the sample VCs. log(AUM) is the logarithm of total assets under management (AUM) due to the heavy tail and high skewness of AUM. Variables hhi_industry_* are the Herfindahl index of each VC calculated from the deal history, where $* \in \{\text{sector, group, code}\}$ stands for three levels of industrial classification in the Pitchbook data. # of ipo, acq, and closed count the total number of startups that went IPO, were acquired, and closed in the entire portfolio history of each VC. success_rate_exit takes the ratio of total successful exits (ipo + acq) over the total number of exits (ipo + acq + closed), while in success_rate the denominator is the total number of startups in the portfolio instead.

Figure 1.1: A snapshot of the undirected network of a subsample of VCs




*Notes:* This figure illustrates an example of the undirected network of a subsample of VCs. Each circle stands for a venture capitalist, with name on the circle. The size of the circles are proportional to their degree centrality, and the edges between VCs stand for their co-investment.

ing and understanding the network effects on VC performance. In this section, we explore the correlation between the centrality of a VC and its performance using Pitchbook data and document some stylized facts.

## Measures of centrality

Following the literature on graphical models, we calculate several different measures of the centrality of VCs in our network. We first introduce their definitions and then provide some summary statistics for these measures.

### Degree centrality

The degree centrality is the most intuitive measure of centrality. The degree centrality of a node $v$ is simply defined as the fraction of nodes it is connected to, i.e.,

$$C_D(v) = \frac{\deg(v)}{|V| - 1}$$

where $\deg(v)$ is the total number of connections $v$ has and $|V| - 1$ is the largest number of possible connections in the graph (in our network without self connections or loops). It measures a node's immediate exposure to information flow through the network. In our VC network, the higher degree centrality a VC has, the more peers it is connected to, giving it a more central role in the network.

For the directed graph, we calculate both indegree centrality and outdegree centrality. Indegree centrality measures the proportion of nodes that point inward at the node $v$, and outdegree centrality measures the proportion of nodes that $v$ is pointing at. In our network, high indegree centrality means the VC acts as a frequent follower of other VCs, while high outdegree centrality indicates the VC often lead other followers in the deals.

### Closeness centrality

The closeness centrality is the average length of the shortest path between a node $v$ and all other nodes in the graph. The larger the closeness centrality, the closer a VC is to all other VCs in terms of the shortest path. Here we use the harmonic closeness centrality which has the form of

$$C_H(v) = \sum_{s \neq v} \frac{1}{d(s, v)}$$

where $d(s, v)$ is the length of the shortest path between nodes $s$ and $v$, and is set to be $\infty$ if the two nodes are disconnected.

### Betweenness centrality

The betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes,

$$C_B(v) = \sum_{s \neq t \neq v \in V} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

where $\sigma_{st}$ is the total number of shortest paths from node $s$ to node $t$, and $\sigma_{st}(v)$ is the number of those shortest paths that pass through $v$. The centrality takes the average over all distinct pairs $(s, t)$ in the network. This measures the "control" of a VC on the communication between other VCs. The larger betweenness centrality a VC has, the more likely it is to participate in the communication and information sharing of other VCs.

### Eigenvector centrality

Eigenvector centrality measures the transitive influence of a node in a network. It takes the centrality of nodes connected with node $v$ into account when calculating $v$'s centrality. Connections with high-scoring nodes contribute more to the eigenvector centrality compared with those low-scoring nodes. A high eigenvector score means that a node is connected to many nodes who themselves have high scores. Formally, the eigenvector centrality for node $v$ is the $v$-th element of the vector $x$ defined by the equation $\mathbf{W}x = \lambda x$ where $\mathbf{W}$ is the adjacency matrix of the graph with eigenvalue $\lambda$. The $(i, j)$-th item of the adjacency matrix $\mathbf{W}$, $W_{ij}$, is 1 if nodes $i$ and $j$ are connected and $W_{ij} = 0$ otherwise. In our network, a VC with high eigenvector centrality means this VC is connected with many other influential VCs.

**Summary statistics of centrality measures**

We calculate different measures of the centrality of all VCs in our sample and report the summary statistics in Table 1.2. All except for closeness centrality are positively skewed, indicating that a large number of VCs have relatively small centrality, while a small proportion of them take very central roles in the network. We also calculate the correlation matrix between these measures and summarize in Tables 1.3 and 1.4. Table 1.3 shows that all measures in the undirected graph are positively correlated, with an almost perfect correlation between degree centrality and eigenvector centrality. Degree centrality is also strongly correlated with closeness and betweenness centrality. However, betweenness centrality and closeness centrality are much less correlated. This indicates that different centrality measures are capturing different aspects of VCs' roles in the network. The high correlation between degree and eigenvector centrality shows that VCs with central roles in the network are more likely to connect with each other, instead of connecting with other peripheral VCs. The low correlation between closeness and betweenness centrality indicates that the VCs that are close to others are not necessarily those who lie on the shortest path of others. The correlation between indegree and outdegree centrality is reported in Table 1.4. Unsurprisingly, the correlation is relatively low, indicating that leaders in the directed network are not frequent followers.

## Centrality and performance

In this subsection, we further explore whether more central VCs perform better. For each of the centrality measures, we regress the success rate on centrality, controlling for the observed features of VCs, industry fixed effects, found year fixed effects, and headquarter-state fixed effects. Again, the performance of each VC is measured by the total number of successful (ipo and/or acq) companies invested over the total size of the portfolio. The higher the rate, the better the investment portfolio of the VC performs. Tables 1.5 – 1.8 summarize the results. In each table, we report three columns that vary in terms of control variables and the fixed effects included.

Table 1.5 shows the significant positive correlation between degree centrality and performance (the first column), even after controlling for the industry, found year, and location fixed effects (the second column). The third column further controls for a bunch of other crucial features that are highly predictive of performance, including the logarithm of total assets

Table 1.2: Summary statistics for VC centrality

| | count | mean | std | p25 | median | p75 |
|---|---|---|---|---|---|---|
| degree_centrality | 1987 | 0.0053 | 0.0079 | 0.0010 | 0.0030 | 0.0065 |
| closeness_centrality | 1987 | 0.2908 | 0.0438 | 0.2632 | 0.2894 | 0.3265 |
| betweenness_centrality | 1987 | 0.0013 | 0.0072 | 0.0000 | 0.0002 | 0.0011 |
| eigenvector_centrality | 1987 | 0.0120 | 0.0190 | 0.0013 | 0.0048 | 0.0156 |
| indegree_centrality | 758 | 0.0015 | 0.0017 | 0.0013 | 0.0013 | 0.0013 |
| outdegree_centrality | 758 | 0.0015 | 0.0029 | 0 | 0 | 0.0013 |

*Notes:* This table summarizes the statistics for several measures of VC centrality in the network. Centrality measures in the first four rows are calculated from the undirected VC network while the last two centrality measures are calculated from the directed VC network.

Table 1.3: Correlation coefficients between different centrality measures (undirected graph)

| | degree | closeness | betweenness | eigenvector |
|---|---|---|---|---|
| degree | 1.0000 | 0.6357 | 0.7775 | 0.9460 |
| closeness | 0.6357 | 1.0000 | 0.2742 | 0.7344 |
| betweenness | 0.7775 | 0.2742 | 1.0000 | 0.6366 |
| eigenvector | 0.9460 | 0.7344 | 0.6366 | 1.0000 |

*Notes:* This table presents the correlation matrix of the four centrality measures calculated from the undirected graph. Each off-diagonal item summarizes the correlation coefficient between two centrality measures in the row and the column.

Table 1.4: Correlation coefficients between different centrality measures (directed graph)

| | indegree_centrality | outdegree_centrality |
|---|---|---|
| indegree_centrality | 1.0000 | 0.1647 |
| outdegree_centrality | 0.1647 | 1.0000 |

*Notes:* This table presents the correlation matrix of the indegree centrality and outdegree centrality calculated from the directed graph. The off-diagonal number stands for the correlation coefficient between these two centrality measures.

under management, the total number of investment professionals, the HHI, and the total number of deals that the VC ever made. Notice that controlling for the observed features and fixed effects significantly decreases the total number of observations, leading to a decrease in the statistical power of our test on correlation, especially in the third column. This is due to the large number of missing values in these features in the Pitchbook data. Despite the small sample size, the reduced form results suggest a significantly positive correlation between degree centrality and performance. Holding other control variables and fixed effects constant, the success rate of a VC is approximately 0.5% higher when one more additional connection is made. (The number $9.848 * 1/1986 = 0.5\%$ is calculated accounting for the normalizing constant $|V| - 1 = 1986$ in our definition of degree centrality.)

Table 1.6 shows similar reduced form implications. High closeness centrality is significantly positively correlated with improved performance whether we control for the fixed effects and the features or not. More concretely, holding the fixed effects and control variables constant, one standard deviation improvement in closeness centrality is associated with an approximately 8.1% higher success rate. The correlation of performance of eigenvector centrality is only significant when the sample size is relatively large, as shown in Table 1.7. After controlling for the features with high numbers of missing values, the statistical power is not strong enough. Finally, Table 1.8 suggests that performance is almost unrelated to betweenness centrality.

In summary, we observe that better VC performance is highly correlated with degree and closeness centrality, while the correlation with betweenness centrality is very loose. This indicates that not all centrality measures are relevant in the VC network. For instance, a VC's number of connections matters much more than how close it is to all other VCs.

Table 1.5: Performance and degree centrality

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | success_rate | success_rate | success_rate |
| degree centrality | 9.242*** | 7.385*** | 9.848** |
| | (2.821) | (1.212) | (4.195) |
| log(AUM) | | | 0.00727 |
| | | | (0.0115) |
| # of investment professionals | | | 0.00809 |
| | | | (0.00500) |
| hhi industry sector | | | -0.0952 |
| | | | (0.0852) |
| total # of deals | | | -0.00269** |
| | | | (0.00131) |
| Constant | 0.401*** | 0.437*** | 0.477*** |
| | (0.0175) | (0.0121) | (0.0681) |
| industry FE | | ✓ | ✓ |
| found year FE | | ✓ | ✓ |
| location FE | | ✓ | ✓ |
| Observations | 1,987 | 1,535 | 434 |
| R-squared | 0.030 | 0.183 | 0.303 |

*Notes:* This table reports the reduced form correlation between VC's performance and degree centrality. The outcome variable is the success rate, which is defined as the ratio of the total number of successful exits over the total number of startups in the portfolio for each VC. Column (1) reports the results of regressing the outcome variable on degree centrality. Column (2) reports the same regression with industry fixed effects, found year fixed effects, and location fixed effects incorporated. Column (3) adds more control variables into the previous regression in addition to the fixed effects. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.6: Performance and closeness centrality

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | success_rate | success_rate | success_rate |
| closeness centrality | 1.513*** | 1.706*** | 1.838*** |
| | (0.202) | (0.249) | (0.572) |
| log(AUM) | | | 0.00781 |
| | | | (0.0113) |
| # of investment professionals | | | 0.00746 |
| | | | (0.00497) |
| hhi industry sector | | | -0.0321 |
| | | | (0.0893) |
| total # of deals | | | -0.000604 |
| | | | (0.000567) |
| Constant | 0.00968 | -0.0205 | -0.0668 |
| | (0.0609) | (0.0737) | (0.194) |
| industry FE | | ✓ | ✓ |
| found year FE | | ✓ | ✓ |
| location FE | | ✓ | ✓ |
| Observations | 1,987 | 1,535 | 434 |
| R-squared | 0.025 | 0.188 | 0.313 |

*Notes:* This table reports the reduced form correlation between VC's performance and closeness centrality. The outcome variable is the success rate, which is defined as the ratio of the total number of successful exits over the total number of startups in the portfolio for each VC. Column (1) reports the results of regressing the outcome variable on closeness centrality. Column (2) reports the same regression with industry fixed effects, found year fixed effects, and location fixed effects incorporated. Column (3) adds more control variables into the previous regression in addition to the fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.7: Performance and eigenvector centrality

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | success_rate | success_rate | success_rate |
| eigenvector centrality | 3.489*** | 2.941*** | 2.018 |
| | (0.642) | (0.523) | (1.349) |
| log(AUM) | | | 0.00895 |
| | | | (0.0115) |
| # of investment professionals | | | 0.00873* |
| | | | (0.00501) |
| hhi industry sector | | | -0.114 |
| | | | (0.0856) |
| total # of deals | | | -0.000957 |
| | | | (0.000897) |
| Constant | 0.408*** | 0.441*** | 0.492*** |
| | (0.0126) | (0.0120) | (0.0687) |
| industry FE | | ✓ | ✓ |
| found year FE | | ✓ | ✓ |
| location FE | | ✓ | ✓ |
| Observations | 1,987 | 1,535 | 434 |
| R-squared | 0.025 | 0.180 | 0.297 |

*Notes:* This table reports the reduced form correlation between VC's performance and eigenvector centrality. The outcome variable is the success rate, which is defined as the ratio of the total number of successful exits over the total number of startups in the portfolio for each VC. Column (1) reports the results of regressing the outcome variable on eigenvector centrality. Column (2) reports the same regression with industry fixed effects, found year fixed effects, and location fixed effects incorporated. Column (3) adds more control variables into the previous regression in addition to the fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.8: Performance and betweenness centrality

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | success_rate | success_rate | success_rate |
| betweenness centrality | 3.832 | 2.238* | -7.649 |
| | (3.108) | (1.246) | (5.248) |
| log(AUM) | | | 0.0101 |
| | | | (0.0114) |
| # of investment professionals | | | 0.00837* |
| | | | (0.00502) |
| hhi industry sector | | | -0.117 |
| | | | (0.0854) |
| total # of deals | | | 0.00309 |
| | | | (0.00210) |
| Constant | 0.445*** | 0.476*** | 0.482*** |
| | (0.0101) | (0.0102) | (0.0713) |
| industry FE | | ✓ | ✓ |
| found year FE | | ✓ | ✓ |
| location FE | | ✓ | ✓ |
| Observations | 1,987 | 1,535 | 434 |
| R-squared | 0.004 | 0.164 | 0.297 |

*Notes:* This table reports the reduced form correlation between VC's performance and betweenness centrality. The outcome variable is the success rate, which is defined as the ratio of the total number of successful exits over the total number of startups in the portfolio for each VC. Column (1) reports the results of regressing the outcome variable on betweenness centrality. Column (2) reports the same regression with industry fixed effects, found year fixed effects, and location fixed effects incorporated. Column (3) adds more control variables into the previous regression in addition to the fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

### The endogeneity problem

Empirically, VCs playing more central roles in the network have better performance. This observed stylized fact suggests positive network effects in our VC network. However, these reduced-form regression results are far from rigorously identified network effects. First, the observed VC network is not exogenously given, nor are the connections made randomly. This leads to the homophily bias problem (Manski, 1993). VCs strategically choose which deals to syndicate and whom to cooperate with. With an extremely high probability, there are some unobserved confounding variables that affect both the formation of connections and the performance of the VCs. For example, the ability and preferences of the VCs are not directly observed in our data. The VCs that share similar preferences in some specific types of startups are more likely to co-invest in startups since they favor the same types. The preference of VCs may also affect their performance. Therefore, in this example, preference acts as an unobserved confounding variable and is not taken into consideration in our reduced form regressions. Second, connected VCs share many common contextual factors. Failing to control for them leads to contextual confounding problems (VanderWeele and An, 2013; Ogburn, 2018). To tackle the endogeneity problems, we conduct two empirical studies in the following two sections utilizing different identification strategies and estimation approaches to estimate the network effects.

## 1.4 Empirical study I: network formation and outcome models

In this first empirical study, we aim to have a better understanding of the network effects (peer effects/spillover effects) in the VC financing market. As discussed in the last section, it faces two main challenges in identification. The first is the high probability of the endogeneity of this network—there are possibly unobserved features that affect both the formation of the network and our outcome of interest. The second is that since most of the VCs are connected (either directly or indirectly) in the VC setting, we only observe network structure data from one single network instead of a large number of exchangeable (independent) networks, as in some friendship data. Therefore, the standard asymptotic arguments do not work in this setting. We solve these two challenges by imposing some structure on the formation of the network and utilizing a novel modeling and estimation methodology proposed in Goldsmith-Pinkham and Imbens (2013). Our outcome of interest is the performance of the VCs, measured by the ratio between the number of successful deals over the total number of deals a VC ever made.

We first review the construction of the graph and introduce some new notations. Throughout the empirical studies, we focus on the undirected graph as it better describes the relationship between VCs since we consider the followers as connected with each other. Also, VCs that co-invest in a startup have the chance to cooperate and get connected, and the connection is not restricted to those who invest in the same single deal. In our graph $(V, W)$, $V$ are the nodes and $W$ are the edges between nodes. We define the edge between two VCs $i$ and $j$, $W_{ij}$, as 1 if both VCs ever simultaneously invested in the same startup and 0 otherwise. We construct the network based on all historical financing information using deals in the Pitchbook database.

Consider the network of $n$ VCs. Throughout the paper, we use $\mathbf{W}$, an $n \times n$ matrix to represent the network among them, with the $(i, j)$-th entry of $\mathbf{W}$:

$$W_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected,} \\ 0 & \text{otherwise.} \end{cases}$$

The diagonal elements of $\mathbf{W}$ are set to zero. $\mathbf{W}$ is called the adjacency matrix in the literature. Under our setting, the links are not directed so that the adjacency matrix $\mathbf{W}$ is symmetric. For each VC $i$, we denote the total number of VCs $i$ is connected with as $M_i$, i.e., $M_i = \sum_{j=1}^{n} W_{ij}$. We also introduce the row-normalized adjacency matrix $\mathbf{G}$ with the $(i, j)$-th entry of $\mathbf{G}$:

$$G_{ij} = \begin{cases} W_{ij}/M_i & \text{if } M_i > 0, \\ 0 & \text{otherwise.} \end{cases}$$

## Manski's linear-in-means model

To study whether interference effects exist among VCs, we would like to explore how VC $i$'s outcome is affected by its peers' outcomes and covariates. Let $Y_i$ denote the outcome of interest of VC $i$, $X_i$ denote some observable covariates such as the total amount under management, the industry of its primary interest, the location of the VC's headquarters, etc. In this subsection, we start by introducing Manski's linear-in-means model (Manski, 1993). Note that the following original linear-in-means model does not deal with the endogenous network formation problem. We use it only as a starting point instead of our final solution.

Consider the following outcome model

$$Y_i = \beta_0 + \beta_x X_i + \beta_{\bar{x}} \bar{X}_{(i)} + \beta_{\bar{y}} \bar{Y}_{(i)} + \eta_i, \tag{1.1}$$

where

$$\bar{Y}_{(i)} = \frac{1}{M_i} \sum_{j=1}^{n} W_{ij} Y_j = \sum_{j=1}^{n} G_{ij} Y_j,$$
$$\bar{X}_{(i)} = \frac{1}{M_i} \sum_{j=1}^{n} W_{ij} X_j = \sum_{j=1}^{n} G_{ij} X_j,$$

are the average outcome and covariates of the connected VCs, respectively. The main parameter of interest is $\beta_{\bar{y}}$, the network effect of connected VCs' performance on a VC's own performance, which is defined as the endogenous peer effect following terminologies in Manski (1993) and Goldsmith-Pinkham and Imbens (2013). This measures how the outcome of a VC is endogenously affected by the average outcomes of all VCs it is connected with, i.e., how much better it will perform if its connected VCs' performances are directly improved. Another parameter of interest is the exogenous peer effect $\beta_{\bar{x}}$, the network effect of connected VCs' features on a VC's own performance. The other parameters $\beta_x$ provide the effect of a VC's features on its own performance.

Rewriting Manski's linear-in-means model equation (1.1) into matrix form, we have

$$\mathbf{Y} = \beta_0 \mathbf{1}_n + \beta_x \mathbf{X} + \beta_{\bar{x}} \mathbf{G}\mathbf{X} + \beta_{\bar{y}} \mathbf{G}\mathbf{Y} + \eta. \tag{1.2}$$

As shown in Goldsmith-Pinkham and Imbens (2013), under the identification assumptions that $\eta = (\eta_1, \ldots, \eta_n)^{\mathrm{T}} \perp\!\!\!\perp \mathbf{X}, \mathbf{W}$ and $\eta \mid \mathbf{X}, \mathbf{W} \sim N\left(0, \sigma_\eta^2 \mathbf{I}\right)$, the coefficients $(\beta_0, \beta_x, \beta_{\bar{x}}, \beta_{\bar{y}})$ are identified.

**Remark.** *Conditioning on the network structure* $\boldsymbol{W}$ *is equivalent to conditioning on the normalized adjacency matrix* $\boldsymbol{G}$ *since* $\boldsymbol{G}$ *is a deterministic function of* $\boldsymbol{W}$*.*

Despite the fact that the parameters are identified, estimating the coefficients using linear regression is generally invalid, as the residual in the linear model $\eta$ is not independent of the right-hand-side variable $\mathbf{GY}$. To consistently estimate the coefficients, we follow the methodology provided in Goldsmith-Pinkham and Imbens (2013) and utilize a Bayesian estimation approach. We describe our detailed estimation procedure and empirical settings in later subsections.

This model implicitly assumes that the given network structure is exogenous by imposing distributional assumptions on $\eta$. However, as discussed above, it is possible that there are some unobserved confounding variables that affect both the formation of connections and the outcome. This is extremely likely to be the case in our VC network setting. For example, ability and preference are possible unobserved confounding variables in this naive Manski's linear-in-means model. To deal with this type of endogeneity problem, we can consider instrumenting the endogenous error term $\eta$ with some exogenous instrumental variables of the network and run IV regressions. However, exogenous IVs for a network structure are always very hard to find in practice (Hsieh and Lee, 2016). Instead, we deal with the endogeneity problem by directly modeling the network formation process and including the endogeneity part in the outcome model in the following subsection.

## Network effects with endogeneity

To estimate the network effects under the endogenous VC network setting, we simultaneously analyze the network formation model and the outcome model, incorporating the unobserved confounding variables in both models. We start with modeling the network formation.

### Network formation model

Let $w_{ij}$ denote the probability that VC $i$ is willing to build a connection with VC $j$. The probability of each directed link $w_{ij}$ is

$$\text{pr}\left\{w_{ij}=1 \mid C_i, C_j, C_{ij}, \xi_i, \xi_j; \gamma, \delta\right\} = \text{sigmoid}\left\{\gamma_0 + C_i\gamma_1 + C_j\gamma_2 + C_{ij}\gamma_3 + \delta\left|\xi_i - \xi_j\right|\right\}, \tag{1.3}$$

where $C_i$ is a vector of observed individual-specific features of the VCs nested in the covariates $X_i$, $C_{ij}$ is a vector of observed dyad-specific features between $i$ and $j$ derived from $X_i$ and $X_j$, and $\xi_i$ is a vector of unobserved features that affect both the network formation probability and the outcome. In our VC network model, the individual-specific features $C_i$ can be their total assets under management, age, preferred industry, and so on. The dyad-specific features $C_{ij}$ measure the similarity of two VCs, such as whether the two VCs have a primary interest in the same industry, the geographic distance between their headquarters, etc. These measures of similarity come into the network formation model due to the homophilic nature of social networks (Manski, 1993). Connections are usually more likely to form between individuals sharing similar features. For instance, VCs preferring startups in the same industry are more likely to co-invest in ventures in this industry. Investors geographically close to each other are also more likely to form connections due to lower communication costs. In addition to these observed features, we include the unobserved feature $\xi_i$ into our formation model. This contributes to the homophily bias in the previous linear-in-means model. The similarity in unobserved features also matters in the decision-making process of forming a connection or not. We can understand the distance $|\xi_i - \xi_j|$ as a measure of similar preferences or abilities between two VCs.

With the unobserved feature incorporated in the network formation model, we make the following conditional independence assumption.

**Assumption 1.1.** *Assume that given the exogenous variables $\boldsymbol{C} = \{C_i, C_{ij} \mid i,j \in \{1,2,\ldots,n\}\}$ and the unobserved variable $\xi$, each link of the network $\boldsymbol{W}$ is conditionally independent of other links.*

Suppose that a connection $W_{ij}$ is made if and only if both the two VCs are willing to build the connection, i.e., $W_{ij} = w_{ij} \cdot w_{ji}$, under Assumption 1.1, we can model the joint probability of the network $\mathbf{W}$ conditional on the observed features $\mathbf{C}$ and the unobserved features $\xi$ as

$$\mathbb{P}\left\{\mathbf{W} \mid \mathbf{C}, \xi; \gamma, \delta\right\} = \prod_{i \neq j} \mathbb{P}\left\{w_{ij} \mid C_i, C_j, C_{ij}, \xi_i, \xi_j; \gamma, \delta\right\}. \tag{1.4}$$

### Outcome model with endogeneity

Next, we consider the outcome model under endogenous network formation. Taking the endogeneity problem into consideration, we are no longer able to assume the $\eta_i$'s in Equation

(1.1) are independent of the network $\mathbf{W}$. Instead, we further decompose $\eta_i$ into two parts,

$$\eta_i = \beta_\xi \xi_i + \varepsilon_i,$$

where $\beta_\xi$ captures how the unobserved features $\xi_i$ affect the outcome, and the remainder term $\varepsilon_i$ is assumed to be the idiosyncratic part and is independently identically distributed across all VCs $i \in \{1, \ldots, n\}$, with distribution $\varepsilon_i \sim N(0, \sigma^2)$. We summarize this formally in the following assumption.

**Assumption 1.2.** *Assume the conditional distribution of $\varepsilon$ conditional on the observed network structure $\boldsymbol{W}$, observed features $\boldsymbol{X}$ of VCs, and the unobserved features $\xi$ are given by*

$$\varepsilon \mid \boldsymbol{W}, \boldsymbol{X}, \xi = (\varepsilon_1, \ldots, \varepsilon_n)^{\mathrm{T}} \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I}_n\right).$$

Therefore, in the case with endogeneity, we can rewrite the outcome model as

$$\mathbf{Y} = \beta_0 \mathbf{1}_n + \beta_x \mathbf{X} + \beta_{\bar{x}} \mathbf{G}\mathbf{X} + \beta_{\bar{y}} \mathbf{G}\mathbf{Y} + \beta_\xi \xi + \varepsilon. \tag{1.5}$$

Rearranging terms in Equation (1.5), we have

$$\mathbf{Y} = \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \beta_0 \mathbf{1}_n + \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \left(\beta_x + \beta_{\bar{x}} \mathbf{G}\right) \mathbf{X} + \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \beta_\xi \xi + \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \varepsilon.$$

Therefore, the conditional likelihood function of the outcome variable $\mathbf{Y}$, conditional on the observed network structure $\mathbf{W}$, features $\mathbf{X}$, unobserved features $\xi$, and the parameters, is a multivariate normal distribution with mean $\mu_Y$ and variance $\Sigma_Y$, where

$$\begin{aligned}
\mu_Y &= \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \beta_0 + \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \left(\beta_x + \beta_{\bar{x}} \mathbf{G}\right) \mathbf{X} + \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \beta_\xi \xi, \\
\Sigma_Y &= \sigma^2 \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-1} \left(\mathbf{I} - \beta_{\bar{y}} \mathbf{G}\right)^{-\mathrm{T}},
\end{aligned}$$

and

$$\mathbb{P}\left\{\mathbf{Y} \mid \mathbf{X}, \mathbf{W}, \xi; \beta, \beta_\xi, \beta_{\bar{y}}, \sigma^2\right\} = \frac{1}{\left|2\pi \Sigma_Y\right|^{1/2}} \exp\left\{-\frac{\left(Y - \mu_Y\right) \Sigma_Y^{-1} \left(Y - \mu_Y\right)^{\mathrm{T}}}{2}\right\} \tag{1.6}$$

where $\beta = (\beta_0, \beta_x, \beta_{\bar{x}})^{\mathrm{T}}$. Therefore, we can use the conditional likelihood functions of outcome variable $\mathbf{Y}$ and network structure $\mathbf{W}$ to update the prior and get the posterior distribution of our parameters. The detailed estimation procedures and results are discussed in the next subsection.

The outcome model provides underlying economic motivation as well. We apply the network model of peer effects constructed in Calvó-Armengol et al. (2009) to our VC network case and show that the outcome model in Equation (1.5) is indeed the unique Nash Equilibrium under mild conditions. Consider our network of $n$ VCs as $n$ agents who simultaneously decide how much effort (denoted as $y_i$) to put into selection and the due diligence process when they make investment decisions, considering how much effort other VCs would exert. Let $\mathbf{y} = (y_1, \ldots, y_n)$ denote the vector that contains the effort levels of all VCs. Following the literature in games on networks (Ballester et al., 2006; Calvó-Armengol et al., 2009), we assume the utility function of a VC $i$ is

$$u_i(\mathbf{y}, \mathbf{W}) = \underbrace{-\frac{1}{2}y_i^2}_{\text{cost}} + \underbrace{\theta_i y_i}_{\text{benefits from own effort}} + \underbrace{\beta_{\bar{y}} \sum_{j=1}^{n} G_{ij} y_j y_i}_{\text{benefits from connected VCs' efforts}}. \tag{1.7}$$

The utility of a VC $i$ consists of three parts: the cost of effort $-\frac{1}{2}y_i^2$, the benefit VC $i$ gets from its own effort $\theta_i y_i$, and benefits from the connected VCs' efforts $\beta_{\bar{y}} \sum_{j=1}^{n} G_{ij} y_j y_i$.

Calvó-Armengol et al. (2009) restrict the coefficient $\theta_i$, which captures the exogenous heterogeneity, to depend only on observables. We relax this restriction and further decompose the heterogeneous $\theta_i$:

$$\theta_i(\mathbf{X}, \xi) = \beta_0 + \beta_x X_i + \beta_{\bar{x}} \sum_{j=1}^{n} G_{ij} X_j + \beta_\xi \xi_i + \varepsilon_i,$$

where $\mathbf{X}$ and $\xi$ are the observed and unobserved features that help with accounting for differences in VCs. This decomposition shows that the exogenous heterogeneity mainly contains three parts: ex-ante idiosyncratic heterogeneity that can be explained using the observed characteristics of both oneself and the connected VCs, $\beta_0 + \beta_x X_i + \beta_{\bar{x}} \sum_{j=1}^{n} G_{ij} X_j$; some unobserved idiosyncratic features $\beta_\xi \xi_i$; and some IID noise $\varepsilon_i$.

The endogenous peer effects from the network, $\beta_{\bar{y}} \sum_{j=1}^{n} G_{ij} y_j y_i$, depends on the role a VC plays in the network, e.g., its centrality and location in the network. This part captures the benefit from sharing of information and resources. In our definition of connections between VCs, two VCs are connected if they invest in the same startup. Each VC spends its own time and effort on research and due diligence. Sharing information on the current status and progress of the startup obviously increases the utility of both VCs, which is summarized in the endogenous peer effects in our model. We follow the standard literature to assume the positive cross derivatives. Therefore, the utility $u_i(\mathbf{y}, \mathbf{W})$ in Equation (1.7) is concave in one's own effort $y_i$ and satisfies the law of diminishing marginal utility in one's own choice level. To solve for the Nash Equilibrium in the game where all VCs simultaneously choose their effort level $y_i$ to maximize their own utility, we calculate the first order condition. The optimal $y_i$ for VC $i$ in network $\mathbf{W}$ is therefore

$$\begin{aligned} y_i^* &= \theta_i + \beta_{\bar{y}} \sum_{j=1}^{n} G_{ij} y_j^* \\ &= \beta_0 + \beta_x X_i + \beta_{\bar{x}} \sum_{j=1}^{n} G_{ij} X_j + \beta_\xi \xi_i + \beta_{\bar{y}} \sum_{j=1}^{n} G_{ij} y_j^* + \varepsilon_i. \end{aligned}$$

By Theorem 1 of Ballester et al. (2006), under the regularity condition that $\mu_1(\beta_{\bar{y}}\mathbf{G}) < 1$, where $\mu_1(\cdot)$ is the spectral radius of a matrix, the joint decision $\mathbf{y}^* = (y_1^*, \ldots, y_n^*)$ is the unique and pure strategy Nash Equilibrium. This matches the matrix form of our outcome model as in Equation (1.5).

## Estimation procedure

In this subsection, we describe the detailed estimation procedures using Bayesian approaches. We first discuss Manski's linear-in-means model where the network is treated as exogenously given as a starting point. In this model, the parameters to be estimated are $\Theta = \left(\beta_0, \beta_x, \beta_{\bar{x}}, \beta_{\bar{y}}, \sigma_\eta^2\right)$. We specify the following independent prior distributions:

$$\begin{aligned}(\beta_0, \beta_x, \beta_{\bar{x}}) &\sim N\left(\mathbf{0}, \mathbf{I}\right), \\ \beta_{\bar{y}} &\sim \text{Uniform}[-1, 1], \\ \sigma_\eta^2 &\sim \text{Inverse } \chi^2(10).\end{aligned}$$

We set the prior distribution of $\beta_{\bar{y}}$ to be Uniform $[-1, 1]$ following suggestions in Smith and LeSage (2004) and Kelejian and Prucha (2010). We restrict the absolute value of $\beta_{\bar{y}}$ to be no more than $1/\mu_1(\mathbf{G})$, the spectral radius of $\mu_1(\mathbf{G})$ (which is equal to the largest absolute value of its eigenvalues), to make sure $(\mathbf{I} - \beta_{\bar{y}}\mathbf{G})$ is invertible and the joint Nash equilibrium exists.

Rearranging Equation (1.2), we have

$$\mathbf{Y} = \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-1} \beta_0 \mathbf{1}_n + \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-1} \left(\beta_x + \beta_{\bar{x}}\mathbf{G}\right)\mathbf{X} + \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-1} \eta.$$

It follows that the conditional likelihood of $\mathbf{Y}$ given all parameters is given by

$$\mathbb{P}\left\{\mathbf{Y} \mid \mathbf{X}, \mathbf{W}; \Theta\right\} = \frac{1}{\left|2\pi\Sigma_Y^{(0)}\right|^{1/2}} \exp\left\{-\frac{\left(Y - \mu_Y^{(0)}\right)\left(\Sigma_Y^{(0)}\right)^{-1}\left(Y - \mu_Y^{(0)}\right)^{\mathrm{T}}}{2}\right\},$$

where $\mu_Y^{(0)} = \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-1} \beta_0 \mathbf{1}_n + \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-1} \left(\beta_x + \beta_{\bar{x}}\mathbf{G}\right)\mathbf{X}$ and $\Sigma_Y^{(0)} = \sigma_\eta^2 \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-1} \left(\mathbf{I} - \beta_{\bar{y}}\mathbf{G}\right)^{-\mathrm{T}}$.

We then update the posterior distribution of the parameters using the prior distribution and the conditional likelihood function,

$$\lambda(\Theta = \theta \mid \mathbf{Y}, \mathbf{X}, \mathbf{W}) = \frac{\pi(\theta)\mathbb{P}\left\{\mathbf{Y} \mid \mathbf{X}, \mathbf{W}; \theta\right\}}{\int \pi(s)\mathbb{P}\left\{\mathbf{Y} \mid \mathbf{X}, \mathbf{W}; s\right\} d\mu(s)},$$

where $\pi(\Theta)$ denotes the prior distribution and $\lambda(\Theta \mid \mathbf{Y}, \mathbf{X}, \mathbf{W})$ denotes the joint posterior distribution. The denominator of the posterior distribution which acts as a normalizing constant is challenging to calculate in practice. Therefore, instead of deriving the closed form of the posterior distribution and directly calculating the posterior summary statistics, we use the Markov Chain Monte Carlo (MCMC, Gelfand and Smith 1990) to generate samples from the posterior distribution and use the empirical summary statistics to conduct estimation and inference. More specifically, we can use the Metropolis-Hastings algorithm and Gibbs sampling to update the parameters sequentially by drawing from their conditional posterior distributions. We omit the details of implementation here.

#### Network effects under endogenous networks

We also use the Bayesian estimation approach in our empirical study when endogeneity of network formation is present. The problem becomes more complicated due to two facts: (1) We need to take care of the unobserved confounding variable $\xi$, specify its prior distribution and update its posterior along with other parameters. (2) Unlike the previous case, we need to focus on both the network formation model and the outcome model simultaneously, as the unobserved $\xi$ contributes to both processes. To solve these problems, we update and estimate a larger model which incorporates all parameters in the network formation and outcome models, as well as the unobserved feature $\xi$. We specify that their prior distributions are independent and are as follows:

$$\begin{aligned}
\xi &\sim N\left(\mathbf{0}, 10^2\mathbf{I}\right), \\
(\gamma_0, \gamma_1, \gamma_2, \gamma_3, \delta) &\sim N\left(\mathbf{0}, 10^2\mathbf{I}\right), \\
(\beta_0, \beta_x, \beta_{\bar{x}}, \beta_\xi) &\sim N\left(\mathbf{0}, \mathbf{I}\right), \\
\beta_{\bar{y}} &\sim \text{Uniform}[-1, 1], \\
\sigma^2 &\sim \text{Inverse } \chi^2(10).
\end{aligned}$$

Again, the prior distribution of $\beta_{\bar{y}}$ is set to Uniform $[-1, 1]$. The prior variance of parameters in the network formation model is set to be larger than those in the outcome model, taking the scale of our outcome variable (the success rate, which is of a very small scale) and the logistic transformation of the network variable $\mathbf{W}$ into consideration.

Similarly, we update the prior distribution using the observed data and the likelihood functions in Equations (1.4) and (1.6). Faced with the same computational issue, we also use the samples generated from the posterior distribution using the MCMC method to calculate the estimated values and credible intervals. We also omit the conditional posterior distribution formulas of the parameters and unobserved feature $\xi$ due to the similarity with previous procedures.

### Empirical results

Table 1.9 reports the results from the posterior distribution of parameters ignoring the endogeneity of network formation as a benchmark. In this naive model where the endogenous formation procedure of the network is not accounted for, the endogenous network effect $\beta_{\bar{y}}$ is significantly positive with a magnitude of 0.136, which means that if we are able to directly permute the average outcome of all peers connected with a VC to be 1 percentage point higher, the network effect would lead to an increase of 0.136% in the rate of success of this VC. The exogenous network effect of the feature AUM is also significantly positive, which measures the network effect of connected VCs' AUM on the performance of this VC. The effect of the AUM on one's own performance is only marginally significant. We also include the total number of investment professionals and preferred industry sector in $X_i$ and $\bar{X}_{(i)}$, and we omit the detailed results on their posterior distributions for the sake of conciseness. Figure 1.2 plots the detailed density functions for the posterior distributions.

Table 1.9: Summary statistics for posterior distribution: outcome model with exogenous networks

| | | Posterior | | | |
|---|---|---|---|---|---|
| | | mean | std | hdi_2.5% | hdi_97.5% |
| $\beta_{\bar{y}}$ | $\overline{\text{outcome}}_{(i)}$ | 0.136*** | (0.032) | 0.074 | 0.199 |
| $\beta_{\bar{x}}$ | $\overline{\text{AUM}}_{(i)}$ | 0.028*** | (0.007) | 0.013 | 0.042 |
| $\beta_x$ | $\text{AUM}_i$ | 0.009* | (0.005) | -0.002 | 0.020 |

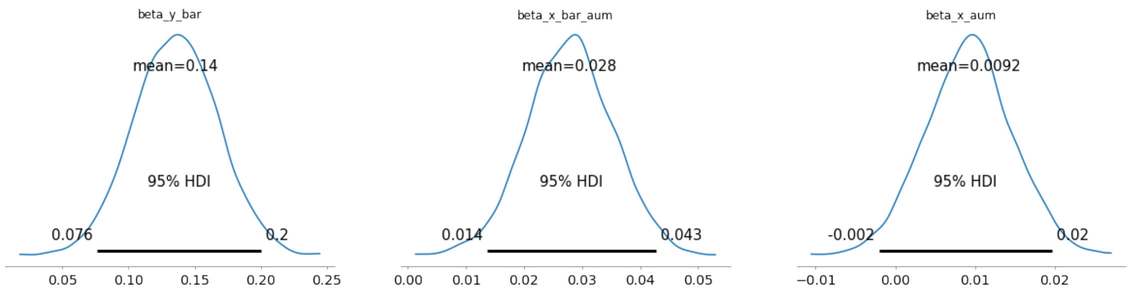
*Notes:* Summary statistics for the posterior distribution of selected parameters. All statistics are numerical from the samples of posterior distribution drawn using Markov Chain Monte Carlo (MCMC). The total number of samples is 4,000 in our results, after discarding the first 4,000 tuning samples. The standard deviations are in parentheses. The last two columns are the lower and upper bound of the 95% highest density intervals of the posterior distribution. ***, **, * indicate that zero is not contained in a 99%, 95%, 90% highest density intervals, respectively. The AUM variable is heavy-tail and we take its logarithm in our analysis. Besides the AUM variable, we also include the total number of investment professionals and preferred industry sector in both $X_i$ and $\bar{X}_{(i)}$. We omit the detailed results of their coefficients.

### Network effects under endogenous network

In order to take care of the underlying endogeneity problem in the network formation, we use the Bayesian approach to estimate the network formation model and outcome model with endogeneity as described in the previous subsections. Note that the unobserved feature $\xi$ is modeled in both models as a confounding variable, so we need to update the posterior distribution of all coefficients in both models and the unobserved feature $\xi$ simultaneously. We report them separately only to make the illustration clear.

Table 1.10 reports the results on network formation model with endogeneity. First, since our graph is essentially modeled as undirected, the effects of AUM are symmetric. VCs with

Figure 1.2: Density function for posterior distribution: outcome model with exogenous networks




*Notes:* This figure plots the density functions of the three parameters of interest in the outcome model with exogenous network: $\beta_{\bar{y}}$, $\beta_{\bar{x}}$, and $\beta_x$. The curves are the posterior density functions, and the solid lines at the bottom of each subplot are the 95% high-density intervals, with the lower bounds and upper bounds of the intervals explicitly presented beside the lines.

more assets under management are more likely to form connections with others. For the dyad-specific features, it is less likely for VCs to form connections with others that are in different locations and industries from themselves. A larger difference in ages also decreases the likelihood of connection. Interestingly, the coefficient of the distance in unobserved features is significantly negative, indicating that similarity in the unobserved feature also contributes to the establishment of connections.

Taking the unobserved endogeneity into consideration, the results of the outcome model are summarized in Table 1.11. The endogenous network effect remains significantly positive. This shows that the positive network effects exist even after controlling for unobserved confounding issues. Again, Figures 1.3 and 1.4 provide detailed density curves for the posterior distributions.

## Heterogeneous network effects

In this subsection, we further explore the heterogeneity of peer effects in different networks following methods in Del Bello et al. (2015). We decompose the connections into those with the same and different preferred industries. Consider two network structures $\mathbf{W}^S$ and $\mathbf{W}^D$. $W_{ij}^S = 1$ if and only if VC $i$ and $j$ ever invested in the same startup and they have the same preferred industry. On the contrary, $W_{ij}^D = 1$ if and only if they invested in the same startup but have different preferred industries. We have the decomposition $\mathbf{W} = \mathbf{W}^S + \mathbf{W}^D$.

We would like to estimate the network effects from both $\mathbf{W}^S$ and $\mathbf{W}^D$. While the network effects from connected VCs in the same industry can be explained by both the flow of information and sharing of resources, the network effects from connected VCs in different

Table 1.10: Summary statistics for posterior distribution: network model with endogenous networks

| | | Posterior | | | |
|---|---|---|---|---|---|
| | | mean | std | hdi_2.5% | hdi_97.5% |
| $\gamma_1$ | $\text{AUM}_i$ | 0.167*** | (0.013) | 0.142 | 0.193 |
| $\gamma_2$ | $\text{AUM}_j$ | 0.167*** | (0.012) | 0.143 | 0.192 |
| | $\mathbf{1}(\text{Loc}_i \neq \text{Loc}_j)$ | -0.874*** | (0.051) | -0.976 | -0.777 |
| $\gamma_3$ | $\mathbf{1}(\text{Ind}_i \neq \text{Ind}_j)$ | -0.723*** | (0.051) | -0.824 | -0.622 |
| | $\lvert \text{age}_i - \text{age}_j \rvert$ | -0.016*** | (0.003) | -0.022 | -0.011 |
| $\delta$ | $\lvert \xi_i - \xi_j \rvert$ | -0.257*** | (0.013) | -0.282 | -0.232 |

*Notes:* Summary statistics for the posterior distribution of selected parameters. All statistics are numerical from the samples of posterior distribution drawn using Markov Chain Monte Carlo (MCMC). The total number of samples is 4,000 in our results, after discarding the first 4,000 tuning samples. The standard deviations are in parentheses. The last two columns are the lower and upper bound of the 95% highest density intervals of the posterior distribution. ***, **, * indicate that zero is not contained in a 99%, 95%, 90% highest density intervals, respectively. The AUM variable is heavy-tail and we take its logarithm in our analysis.

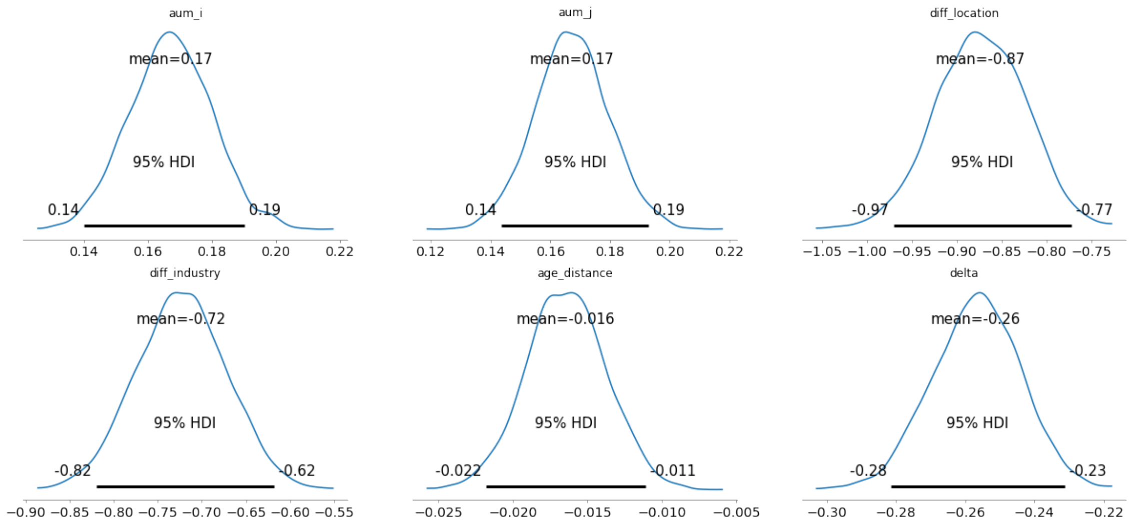
industries are mainly induced by the sharing of resources and expertise. Therefore, by taking a closer look at the heterogeneous effects in the two different networks, we are able to better decompose and distinguish the two channels. We model the network formation using the same features as in the previous subsections, incorporating the endogenous unobserved feature into the network formation model as well. We model networks within and across industries separately. The outcome model under two networks becomes

$$\mathbf{Y} = \beta_0 \mathbf{1}_n + \beta_x \mathbf{X} + \beta_{\bar{x}}^S \mathbf{G}^S \mathbf{X} + \beta_{\bar{x}}^D \mathbf{G}^D \mathbf{X} + \beta_{\bar{y}}^S \mathbf{G}^S \mathbf{Y} + \beta_{\bar{y}}^D \mathbf{G}^D \mathbf{Y} + \beta_\xi \xi + \varepsilon,$$

where $\beta_{\bar{y}}^k$ and $\beta_{\bar{x}}^k$ measures the endogenous and exogenous network effects from network $k$, respectively, for $k \in \{S, D\}$, and the same set of observed and unobserved features are used.

Assuming the connection is made if and only if both VCs are willing to build the link with each other in both networks, under Assumption 1.1, we have the same model of the joint probability of the two networks $\mathbf{W}^S$ and $\mathbf{W}^D$. We assume independence between connections in the two networks. We rewrite the outcome model to get the explicit form of the posterior

Figure 1.3: Density function for posterior distribution: network formation model with endogenous network




*Notes:* This figure plots the density functions of the six parameters of interest in the network formation model with endogenous networks. The curves are the posterior density functions, and the solid lines at the bottom of each subplot are the 95% high-density intervals, with the lower bounds and upper bounds of the intervals explicitly presented beside the lines.

distribution. The posterior distribution of $\mathbf{Y}$ given observed features, the two observed networks, and the unobserved features could be easily written into a normal distribution with a closed-form mean vector and variance-covariance matrix under Assumption 1.2. Together with the two network formation models, we follow the procedure in Del Bello et al. (2015) and use Bayesian estimation methods to get the estimated posterior distribution of our parameters of interest. The prior distributions of parameters and the unobserved features are similarly specified as in previous subsections. We omit the details to avoid repetition.

Table 1.12 reports the posterior distributions of parameters of interest in the outcome model. Decomposing the VC network into two networks: within and across industries, we observe significantly positive endogenous network effects in both networks, while the magnitude within the same industry is about 10% higher than across industries. The exogenous network effects become insignificant in the network across different industries, indicating that when considering connections across industries, one's performance is not significantly affected by its peers' observed features in other industries. The significantly positive network effect $\beta_{\bar{y}}^{D}$ indicates that not only does the sharing of information within the same industry matter, but the sharing of resources and expertise also adds value to the connections across different industries in the network.

Table 1.11: Summary statistics for posterior distribution: outcome model with endogenous networks

| | | Posterior | | | |
|---|---|---|---|---|---|
| | | mean | std | hdi_2.5% | hdi_97.5% |
| $\beta_{\bar{y}}$ | $\overline{\text{outcome}}_{(i)}$ | 0.127*** | (0.028) | 0.072 | 0.182 |
| $\beta_{\bar{x}}$ | $\overline{\text{AUM}}_{(i)}$ | 0.028*** | (0.007) | 0.015 | 0.041 |
| $\beta_x$ | $\text{AUM}_i$ | 0.004 | (0.006) | -0.007 | 0.015 |
| $\beta_\xi$ | $\xi_i$ | 0.001 | (0.002) | -0.002 | 0.004 |

*Notes:* Summary statistics for the posterior distribution of selected parameters. All statistics are numerical from the samples of posterior distribution drawn using Markov Chain Monte Carlo (MCMC). The total number of samples is 4,000 in our results, after discarding the first 4,000 tuning samples. The standard deviations are in parentheses. The last two columns are the lower and upper bound of the 95% highest density intervals of the posterior distribution. ***, **, * indicate that zero is not contained in a 99%, 95%, 90% highest density intervals, respectively. The AUM variable is heavy-tail and we take its logarithm in our analysis. Besides the AUM variable, we also include the total number of investment professionals and preferred industry sector in both $X_i$ and $\bar{X}_{(i)}$. We omit the detailed results of their coefficients.

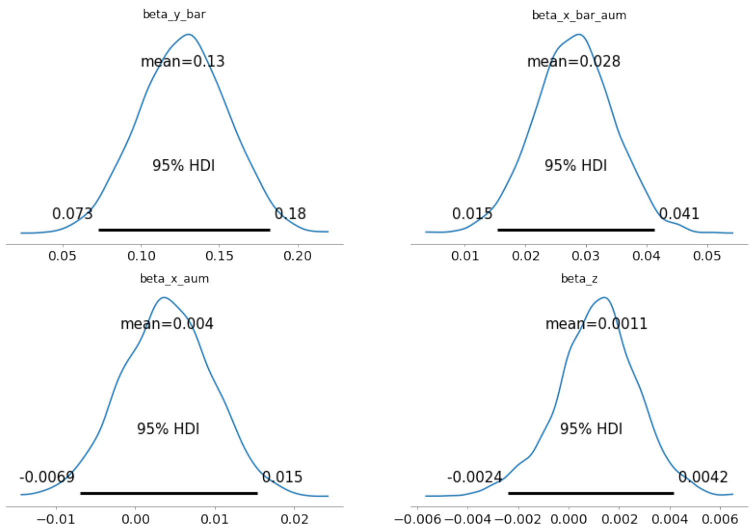
## 1.5 Empirical study II: generalized difference-in-differences method

The previous section estimates the network effects among VCs by simultaneously modeling the network formation and outcome models. In this empirical study, we estimate the peer effects from a more intuitive perspective. We aim to answer the question that if a VC encounters a very successful event, will its connected peers' performance be affected in the following years. More precisely, we want to learn if an extremely successful event such as an IPO in the portfolio of a VC $i$ will positively impact the performance of this VC (main effects) and its linked VCs (peer effects) in the following years.

By the nature of this empirical question, we need panel data for all VCs, including time-varying measures of performance and features. Our outcome variables of interest, measuring the VC's performance, are the total number of startups in VC $i$'s portfolio that

- received funding within in one, two, or three years, i.e., within the interval of month $[t, t+12)$, $[t, t+24)$, or $[t, t+36)$, respectively ($t$ is at the month level);

Figure 1.4: Density function for posterior distribution: outcome model with endogenous networks




*Notes:* This figure plots the density functions of the four parameters of interest in the outcome model with endogenous network: $\beta_{\bar{y}}$, $\beta_{\bar{x}}$, $\beta_x$, and $\beta_\xi$. The curves are the posterior density functions, and the solid lines at the bottom of each subplot are the 95% high-density intervals, with the lower bounds and upper bounds of the intervals explicitly presented beside the lines.

- went IPO or was acquired within one, two, or three years;
- went bankrupt within one, two, or three years.

We also consider the dynamic VC network in this study. Connections among VCs are not static and can change over time, especially during our sample period when entries and exits of VCs are common. We construct the dynamic VC network using the information in all deals in the previous five years[1]. As for the time-varying features, we calculate their age, the degree centrality based on the dynamic network, and the total number of startups in the portfolio, etc. We specify the detailed control variables included in the regressions in the following subsection.

We conduct a generalized difference-in-differences study (Goodman-Bacon, 2021) to measure both effects—the main effects and the peer effects. Difference-in-differences (diff-in-diff) is a quasi-experimental research design used to measure the effect of events such as policy changes, climate or environmental changes, medical treatments or surgeries, etc. It compares the differences between two sets of changes: the difference in the level of the outcome before

[1]We use the 5-year window following the literature (Hochberg et al., 2007).

Table 1.12: Heterogeneous network effects

| | | Posterior | | | |
|---|---|---|---|---|---|
| | | mean | std | hdi_2.5% | hdi_97.5% |
| $\beta_{\bar{y}}^{S}$ | $\overline{\text{outcome}}_{(i)}^{S}$ | 0.144*** | (0.038) | 0.067 | 0.219 |
| $\beta_{\bar{y}}^{D}$ | $\overline{\text{outcome}}_{(i)}^{D}$ | 0.128*** | (0.041) | 0.046 | 0.207 |
| $\beta_{\bar{x}}^{S}$ | $\overline{\log(\text{AUM})}_{(i)}^{S}$ | 0.019*** | (0.009) | 0.001 | 0.038 |
| $\beta_{\bar{x}}^{D}$ | $\overline{\log(\text{AUM})}_{(i)}^{D}$ | 0.005 | (0.008) | -0.011 | 0.021 |
| $\beta_{x}$ | $\log(\text{AUM})_i$ | 0.012*** | (0.0075) | 0.001 | 0.022 |
| $\beta_{\xi}$ | $\xi_i$ | 0.001 | (0.001) | -0.002 | 0.003 |

*Notes:* Summary statistics for the posterior distribution of selected parameters. All statistics are numerical from the samples of posterior distribution drawn using Markov Chain Monte Carlo (MCMC). The total number of samples is 4,000 in our results, after discarding the first 4,000 tuning samples. The standard deviations are in parentheses. The last two columns are the lower and upper bound of the 95% highest density intervals of the posterior distribution. ***, **, * indicate that zero is not contained in a 99%, 95%, 90% highest density intervals, respectively. The AUM variable is heavy-tail and we take its logarithm in our analysis. Besides the AUM variable, we also include the total number of investment professionals and preferred industry sector in both $X_i$ and $\bar{X}_{(i)}$. We omit the detailed results of their coefficients.

and after the change/treatment in the treated group, and that same difference in the control group. The most crucial underlying identification assumption is the parallel trend condition between two groups. More precisely, in the absence of the change/treatment, the trends of the outcome variables in two groups should be parallel to each other, i.e., the difference between the two groups should be a constant that does not vary across time.

In our empirical study, we use an IPO as an extremely successful event. Since IPOs occur throughout our data, the treatments can occur at different time periods. We follow the literature (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Goodman-Bacon, 2021) and estimate the treatment effect using a generalized diff-in-diff estimator, which is defined as the coefficient of the treatment dummy variable in the regression of outcome variable on treatment dummy, control variables, and several fixed effects.

## Main effects

We first estimate the main effects. Define the event time $E_{it}$ as the first month VC $i$ has at least one startup IPO in month $t$, and define the treatment variable $D_{it}$ as the time-varying indicator variable $D_{it} = 1\{t \geq E_{it}\}$. For each of the outcome variables, we run the following generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where

- $\tau$ is our parameter of interest that measures the main effects,
- $\alpha_i$ are the VC fixed effects,
- $\lambda_{t,ind}$ are the interaction between time and industry fixed effects, and
- $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$.

To construct a valid sample for the generalized diff-in-diff model, we construct a panel of VCs that have their first equity deal before 2010 and last equity deal after 2021, so that all have the same possible treatment set within this interval. We then run the regressions on all observations of these valid VCs between 2010 and 2021. Robustness checks show that the results are not sensitive to the start and end times. We report the regression results for the outcomes constructed using the information in the following one year in Table 1.13, and relegate regression results for the outcomes using the information in the following two and three years in Tables A.1 and A.2 in the appendix.

The results show significantly positive main effects of an IPO on all measures of the future performance of the VCs. Within one year after the successful event, 0.455 more startups in the portfolio of this VC received new funding, 0.075 more startups went public and/or were acquired, and 0.107 fewer startups went bankrupt, after controlling for several time-varying features of the VCs, the interaction between time and industry fixed effects, and VC fixed effects. The estimated main effects are all significantly positive for the same set of outcome variables within a longer time horizon.

## Network effects

Next, we estimate the peer effects of a highly positive event (an IPO). Similarly, we define the event time $E_{peer,it}$ as the first month $t$ in which an IPO occurs to at least one of VC $i$'s connected peer's portfolio. Define the treatment variable $D_{peer,it}$ as the time-varying

Table 1.13: Main effects of IPO on future performance

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{it}$ | 0.455*** | 0.0752*** | -0.107*** |
| | (0.0394) | (0.0186) | (0.0148) |
| # of invested companies | 0.199*** | 0.0340*** | 0.0393*** |
| | (0.00172) | (0.000810) | (0.000646) |
| squared age | -0.00640*** | 9.00e-05 | 0.00193*** |
| | (0.00105) | (0.000493) | (0.000393) |
| degree centrality | 132.6*** | 10.35*** | -0.312 |
| | (2.094) | (0.987) | (0.787) |
| Constant | 0.122** | 0.113*** | -0.198*** |
| | (0.0537) | (0.0253) | (0.0202) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 33,535 | 33,535 | 33,535 |
| R-squared | 0.838 | 0.484 | 0.462 |

*Notes:* This table reports the main effects of IPO on the future 1-year performance of the VCs. The treatment variable $D_{it} = 1\{t \geq E_{it}\}$ is the time-varying indicator where $E_{it}$ is the event time, the first month VC $i$ has at least a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 1 year, that have successful exits within 1 year, and that file bankruptcy within 1 year, respectively. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

indicator variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$. The outcome variables of interest are defined and constructed the same as in the previous analysis. To tease out the main effects in this generalized diff-in-diff model, we run the regression on the subsample that does not have its own IPO event. We use the generalized diff-in-diff regression model

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where the control variables and fixed effects are the same as in previous analyses. The results for outcome variables using a one-year window are summarized in Table 1.14, and the results for those using longer time horizons are relegated to the appendix. The estimated peer effects are also significantly positive across all time horizons in the sense that once an IPO occurs in a VC's peer's portfolio, this VC's portfolio will also receive more funding opportunities, have more successful exits, and suffer from fewer bankruptcies.

## Decomposition of network effects

To more thoroughly understand the positive peer effects, we further decompose the peer effects into two parts. Figure 1.5 uses two VCs as an example to illustrate our decomposition of the VC portfolios. Suppose there are two connected VCs with their investment portfolios shown as the two circles in the graph. Without loss of generality, we define the VC $i$ as the *ego*—the VC on whom we estimate the peer effect—and VC $j$ as the *peer*—the VC whose effect on the ego we estimate. Their portfolios have overlapped since they co-invest in some common startups. We use $C$ to denote the overlapping set of startups which includes all commonly invested companies. We use $A$ and $B$ to denote the non-overlapping part of the portfolios of VC $i$ and VC $j$, respectively. That is, $A \cup C$ includes the full set of companies invested by VC $i$, and $B \cup C$ includes those by VC $j$. Our previous analyses on peer effects consider the case when VC $i$ does not have any IPO event in its portfolio $A \cup C$, while VC $j$ has (possible multiple) IPO event(s) in its portfolio $B$ (the part that does not overlap with VC $i$). Results in Table 1.14 show the peer effects of VC $j$'s successful event on VC $i$'s performance, i.e., IPOs in set $B$ on the future performance of $A \cup C$ as a whole. We now further decompose the portfolio of the ego VC $i$ into two parts: set $C$ that is commonly owned by both VCs and set $A$ that does not overlap with VC $j$'s portfolio.

The first set of results reported in Table 1.15 shows the effect of an IPO in peer VC $j$'s non-overlapping portfolio (set $B$ in Figure 1.5) on the future performance of the common portfolio $C$. There are significantly positive effects on the total number of startups that received funding and succeeded within one year as shown in Table 1.15. The negative effect on decreasing the total number of bankruptcies is very small and not significant. We report results using outcomes on a longer horizon in the appendix.

The positive effect on set $C$ indicates that one channel of the positive network effects is to share the benefits of the good-performing VCs by co-investing with them and building up the connections. As part of VC $j$'s portfolio, the set $C$ enjoys a positive (main) effect on

Table 1.14: Peer effects of IPO on future performance

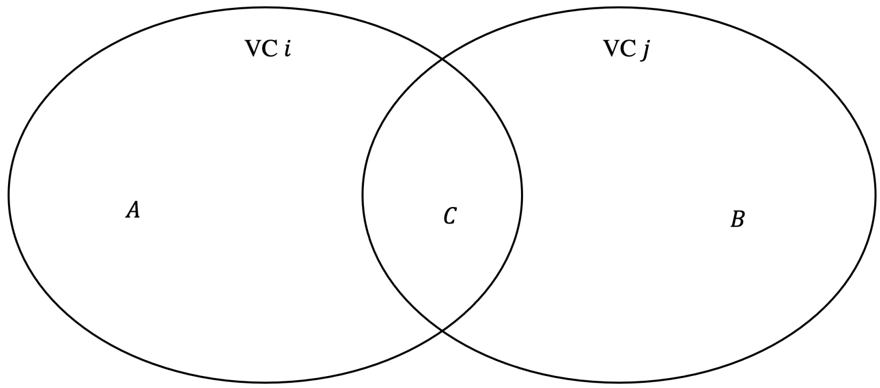
| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.169*** | 0.0723*** | -0.0505*** |
| | (0.0333) | (0.0141) | (0.0136) |
| # of invested companies | 0.154*** | 0.0260*** | 0.0411*** |
| | (0.00195) | (0.000826) | (0.000795) |
| squared age | -0.00158 | 0.000141 | 0.000910* |
| | (0.00116) | (0.000490) | (0.000471) |
| degree centrality | 147.3*** | 12.63*** | 2.777*** |
| | (2.580) | (1.093) | (1.052) |
| Constant | 0.0899 | 0.0348 | -0.139*** |
| | (0.0581) | (0.0246) | (0.0237) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 22,055 | 22,055 | 22,055 |
| R-squared | 0.815 | 0.415 | 0.501 |

*Notes:* This table reports the peer effects of IPO on the future 1-year performance. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 1 year, that have successful exits within 1 year, and that file bankruptcy within 1 year, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1.5: Decomposition of VC portfolios




*Notes:* This figure illustrates the decomposition of VC portfolios. For two connected VCs in the network, we decompose the startups in their portfolio into three sets: set $A$ and $B$ include the non-overlap portfolio of VC $i$ and VC $j$ respectively, and set $C$ includes all the startups that are commonly owned by both VCs.

future performance after VC $j$'s successful event. It is obvious from our results that the ego VC $i$ shares this positive effect since it is connected to VC $j$.

More interestingly, Table 1.16 shows that not only the common portfolio but also the non-overlapping part $A$ benefits from the positive network effects. The treatment effects on the total number of startups that received funding and exited successfully are significantly positive, and the treatment event has a negative effect on the number of bankruptcies. Similar results hold for outcomes calculated using a longer horizon and are reported in the appendix. This decomposition indicates that the VCs not only benefit from the relationship by enjoying their connected successful peer's positive main effects but also improve the performance of their own non-overlapping portfolio when they are connected to peers with excellent performance.

## 1.6 Model

In this section, we build a theoretical model to illustrate the network effects and test the model implications using our empirical data.

### Setup and timeline

Suppose that there are two types of entrepreneurs in the population, a high ability type and a low ability type, denoted by $A_e \in \{h, l\}$. Suppose that the VCs in the population are also heterogeneous in abilities, denoted by $A_{vc}$. We assume that the abilities of VCs are drawn

Table 1.15: Peer effects of IPO on future performance: common portfolio

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.0643*** | 0.0270*** | -3.04e-05 |
| | (0.0167) | (0.00885) | (0.00451) |
| # of invested companies | 0.341*** | 0.117*** | 0.0556*** |
| | (0.0186) | (0.00986) | (0.00502) |
| squared age | -0.00318*** | 0.000593 | -0.000543*** |
| | (0.000717) | (0.000380) | (0.000193) |
| degree centrality | 10.18*** | -2.153*** | -0.178 |
| | (1.268) | (0.672) | (0.342) |
| Constant | -0.0170 | -0.0845*** | -0.0121 |
| | (0.0497) | (0.0263) | (0.0134) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 8,872 | 8,872 | 8,872 |
| R-squared | 0.318 | 0.367 | 0.446 |

*Notes:* This table reports the peer effects of IPO on the future 1-year performance of the common portfolio subsample, i.e., the overlapping set $C$ in Figure 1.5. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the common portfolio that receive new funding within 1 year, that have successful exits within 1 year, and that file bankruptcy within 1 year, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.16: Peer effects of IPO on future performance: non-overlapping portfolio

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.0991*** | 0.0585*** | -0.0518*** |
| | (0.0333) | (0.0139) | (0.0138) |
| # of invested companies | 0.155*** | 0.0268*** | 0.0418*** |
| | (0.00196) | (0.000817) | (0.000812) |
| squared age | 0.00169 | -0.000301 | 0.00101** |
| | (0.00117) | (0.000485) | (0.000482) |
| degree centrality | 138.8*** | 10.32*** | 2.762*** |
| | (2.547) | (1.060) | (1.053) |
| Constant | -0.0454 | 0.0450* | -0.142*** |
| | (0.0586) | (0.0244) | (0.0242) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 21,505 | 21,505 | 21,505 |
| R-squared | 0.812 | 0.417 | 0.504 |

*Notes:* This table reports the peer effects of IPO on the future 1-year performance of the non-overlap portfolio subsample, i.e., the non-overlapping set $A$ in Figure 1.5. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's non-overlap portfolio that receive new funding within 1 year, that have successful exits within 1 year, and that file bankruptcy within 1 year, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

from a uniform distribution on the interval $[0, 1]$. The higher $A_{vc}$ is, the higher their ability is.

When entrepreneurs seek funding opportunities and sell equity in their startups to VCs, we see this process as signal-sending. Assume that the signals entrepreneurs send have two types: a high signal and a low signal, denoted by $S \in \{h, l\}$. The observed signals sent from high-ability entrepreneurs are assumed to be the high type with probability 1, i.e., $\mathbb{P}\{S = h \mid A_e = h\} = 1$[2]. As for the low-ability entrepreneurs, assume they are able to pretend to be capable and manage to send high signals with some (positive) probability. This probability depends on the ability of the VC who does the screening. The higher the VC's ability, the more easily it can distinguish the true ability of the startups. Assume $\mathbb{P}\{S = h \mid A_e = l, A_{vc}\} = p\exp(-A_{vc})$, i.e., the probability a low type entrepreneur succeeds in pretending to be a high type is inversely correlated with the ability of the VC who receives the signal. In other words, the accuracy of the signals received by high-ability VCs is higher.

At time 0, the entrepreneur comes to a VC of ability $A_{vc}$ for funding opportunities and sends a signal $S_0 \in \{h, l\}$. The VC has three options to choose from after observing the initial signal $S_0$:

1. reject the proposal with no investment;
2. invest in the startup alone;
3. propose a syndicated deal to other VCs.

If the VC chooses the first option, no investment will be made and the deal will be closed. With the second option, the VC will complete a standalone deal and invest in the startup by itself. With the third option, if the VC decides to propose a syndicated deal, then at time 1, either the VC or the entrepreneur will have to reach out to other potential investors and attract their interest. The entrepreneur will send a second signal $S_1$ to the peer VCs in order for them to agree with the investment. Again, the signal $S_1 \in \{h, l\}$, where the low ability entrepreneurs send high signals with probability $p\exp(-A_{vc,peer})$ depending on the ability of the peer VC who receives the signal, and the high ability entrepreneurs send high signals with a probability of one. We further assume that conditioning on the true ability of the entrepreneurs and VCs, the peer signal $S_1$ is independent of the initial signal $S_0$. Observing and collecting both signals $S_0$ and $S_1$, the VCs discuss and decide together whether they would like to co-invest in the startup or not. If they decide to co-invest, a syndicated deal is achieved and they negotiate how much share the initial/lead VC takes. The more capable the initial VC is, the more attractive it is to the peer VCs since the probability of successfully

[2]This assumption can be easily generalized to high type sending high signals with probability $p_h$. For a more intuitive discussion, we make this simple assumption for now.

selecting a high-ability type entrepreneur,

$$\begin{aligned}\mathbb{P}\{A_e = h \mid S_0 = h, A_{vc}\} &= \frac{\alpha \mathbb{P}\left\{S_0 = h \mid A_e = h, A_{vc}\right\}}{\alpha \mathbb{P}\left\{S_0 = h \mid A_e = h, A_{vc}\right\} + (1-\alpha)\mathbb{P}\left\{S_0 = h \mid A_e = l, A_{vc}\right\}} \\ &= \frac{\alpha}{\alpha + (1-\alpha) p \exp\{-A_{vc}\}},\end{aligned}$$

is increasing with $A_{vc}$, where $\alpha$ denotes the proportion of high-ability entrepreneurs in the population, possibly a very small number. Thus, higher ability VCs have more bargaining power in the negotiation. We assume that the share $s$ of the initial VC is an increasing function of its ability, $s = \exp\{-(1 - A_{vc})\}$.

At time 2, the return of the investment is realized. Suppose the return of the outside investment opportunity is $r$. This is the cost of the investment VCs make if they decide to invest in the startup. Assume the high-ability entrepreneurs have an expected return of $R_h$ in a standalone deal. When a syndicated deal is achieved, different VCs will cooperate together in the managerial process of the startup and help it using their different expertise, leading to more added value to the startup. Therefore, for syndicated deals, the high-ability entrepreneurs have an expected return of $R_h + v$, where $v > 0$ is the benefit of syndication. The low ability types are assumed to fail for sure with $R_l = 0$ no matter whether they have standalone or syndicated deals.

## VC's decision-making process

Next, we discuss the decision-making processes of the VCs by calculating and comparing their expected returns under different options with different signals.

First, if a low signal is observed at time 0, i.e., $S_0 = l$, the VC will immediately reject the investment no matter what its ability $A_{vc}$ is, since $\mathbb{P}(A_e = l \mid S_0 = l, A_{vc}) = 1$. The expected return of the VC would be 0 if they decide to invest alone or syndicated, while the expected return of rejection would be $r > 0$. The rejection option dominates the other two when $S_0 = l$.

Next, we calculate the expected return of a VC when a high initial signal is observed at time 0. Without loss of generality, we assume the total amount of investment is 1. Let $R$ denote the return realized at time $t = 2$, and use the abbreviations {sta, syn} to denote {standalone, syndication}, respectively. If they choose a standalone investment, the expected return at time 0 is

$$\begin{aligned}R_{sta} &:= \mathbb{E}\left\{R \mid S_0 = h, A_{vc}, \text{sta}\right\} \\ &= \mathbb{E}\left\{R \mid A_e = h, S_0 = h, A_{vc}, \text{sta}\right\} \mathbb{P}\left\{A_e = h \mid S_0 = h, A_{vc}, \text{sta}\right\} \\ &+ \mathbb{E}\left\{R \mid A_e = l, S_0 = h, A_{vc}, \text{sta}\right\} \mathbb{P}\left\{A_e = l \mid S_0 = h, A_{vc}, \text{sta}\right\} \\ &= R_h \mathbb{P}\left\{A_e = h \mid S_0 = h, A_{vc}\right\} + 0\end{aligned}$$

$$= R_h \frac{\alpha \mathbb{P}\{S_0 = h \mid A_e = h, A_{vc}\}}{\alpha \mathbb{P}\{S_0 = h \mid A_e = h, A_{vc}\} + (1-\alpha)\mathbb{P}\{S_0 = h \mid A_e = l, A_{vc}\}}$$
$$= \frac{R_h \alpha}{\alpha + (1-\alpha)\, p \exp(-A_{vc})}, \tag{1.8}$$

where $\alpha$ is the proportion of high-ability entrepreneurs in the population. We assume that it is generally difficult for a low-ability entrepreneur to act as a high-ability type, which means $p$ is a small number. It is reasonable to assume that $R_h \gg r$ so that the expected rate of return of the standalone investment is always larger than $r$ when the initial signal is $h$.

If the VC proposes syndication, there are two possible signals received by their peers at time 1. If the peer's observed signal is low, both the syndicated deal and the standalone deal will be rejected, and the expected return will be $r$ since $\mathbb{P}(A_e = l \mid S_0 = h, S_1 = l, A_{vc}, A_{vc,peer},) = 1$ no matter what the levels of VCs' abilities are. If, however, $S_1 = h$, then the syndication will be achieved and the VC will invest in the startup with a share of $s$. The expected return conditional on the VC's ability $A_{vc}$ and the peer's ability $A_{vc,peer}$ is

$$\begin{aligned}
R_{syn} \mid A_{vc,peer} &:= \mathbb{E}\{R \mid S_0 = S_1 = h, A_{vc}, A_{vc,peer}, \text{syn}\} \\
&= \mathbb{E}\{R \mid A_e = h, S_0 = S_1 = h, A_{vc}, A_{vc,peer}, \text{syn}\} \\
&\times \mathbb{P}\{A_e = h \mid S_0 = S_1 = h, A_{vc}, A_{vc,peer}, \text{syn}\} \\
&= \exp\{-(1-A_{vc})\}(R_h + v)\,\mathbb{P}\{A_e = h \mid S_0 = S_1 = h, A_{vc}, A_{vc,peer}\} \\
&= \exp\{-(1-A_{vc})\}(R_h + v) \\
&\times \frac{\alpha}{\alpha + (1-\alpha)\mathbb{P}\{S_0 = S_1 = h \mid A_e = l, A_{vc}, A_{vc,peer}\}} \\
&= \exp\{-(1-A_{vc})\}(R_h + v) \\
&\times \frac{\alpha}{\alpha + (1-\alpha)\mathbb{P}\{S_0 = h \mid A_e = l, A_{vc}\}\,\mathbb{P}\{S_0 = h \mid A_e = l, A_{vc}\}} \\
&= \frac{\exp\{-(1-A_{vc})\}(R_h + v)\,\alpha}{\alpha + (1-\alpha)\, p^2 \exp(-A_{vc}) \exp(-A_{vc,peer})},
\end{aligned}$$

by the fact that $\mathbb{P}\{S_0 = S_1 = h \mid A_e = h, A_{vc}, A_{vc,peer}\} = 1$. Comparing this conditional expected return of syndication with that of standalone deals as in Equation (1.8), we see the trade-offs of the VCs when they choose between standalone investment and syndication. Compared with standalone deals, proposing a syndicated deal has the following benefits and costs:

**[Benefits]** - the benefits of syndication come from two parts:

1. the higher added-value $(R_h + v)$ compared with the standalone investment case, $R_h$;

2. the better selection process thanks to the information in the additional signal observed by the peer VC, which leads to the lower value of the denominator and thus a higher value of the total expected return.

**[Costs]** VCs do not enjoy these benefits for free. The cost of syndication is due to dilution. Instead of owning the whole startup like in Equation (1.8) in the standalone investment case, after syndication the initial VC only gets a proportion of $\exp(A_{vc} - 1)$ which is no larger than 1 since the ability $A_{vc} \in [0, 1]$. This factor decreases the expected return and thus acts as a cost for VCs.

Integrating over the distribution of the ability of other peer VCs, $A_{vc,peer} \sim \text{Uniform}[0, 1]$, we get the expected return conditioning on the VC's own ability $A_{vc}$,

$$\begin{aligned} R_{syn} &= \mathbb{E}\left\{R \mid S_0 = S_1 = h, A_{vc}, \text{syn}\right\} \\ &= \int_0^1 \mathbb{E}\left\{R \mid S_0 = S_1 = h, A_{vc}, A_{vc,peer}, \text{syn}\right\} d\mathbb{P}\left(A_{vc,peer}\right) \\ &= \int_0^1 \frac{\exp\left\{-(1 - A_{vc})\right\}(R_h + v)\,\alpha}{\alpha + (1-\alpha)\,p^2 \exp\left(-A_{vc}\right)\exp\left(-x\right)} dx \\ &= \exp\left\{-(1 - A_{vc})\right\}(R_h + v)\left\{\log\left(1 + \frac{\alpha\exp\left(A_{vc} + 1\right)}{(1-\alpha)\,p^2}\right) - \log\left(1 + \frac{\alpha\exp(A_{vc})}{(1-\alpha)\,p^2}\right)\right\}. \end{aligned} \tag{1.9}$$

VCs will compare the expected returns of standalone and syndicated deals, $R_{sta}$ and $R_{syn}$, to make their final decision after observing $S_0 = h$. Let $A^*$ denote the solution to the equation $R_{sta} = R_{syn}$, where the existence and uniqueness of this solution under mild assumptions are shown in the proof. Combining all previous results, we summarize the decision-making process of the initial VC into the following proposition.

**Proposition 1.1.** *After observing the initial signal $S_0$ at time 0, the initial VC will*

1. *reject the deal if and only if $S_0 = l$;*
2. *invest in the startup alone if and only if $S_0 = h$ and $A_{vc} \leq A^*$;*
3. *propose syndication if and only if $S_0 = h$ and $A_{vc} > A^*$, and achieve the syndication if and only if the peers' signal $S_1 = h$ at time 1.*

*Proof outline of Proposition 1.1.* We only briefly explain the outline of the proof here and relegate the rigorous proof to the appendix for the sake of conciseness. From the discussion above, the VC chooses between standalone and syndicated deals after observing the high initial signal $S_0 = h$. Their decision is based on the expected returns of these two options, whose closed-form formulas are shown in Equations (1.8) and (1.9). For instance, when $R_{sta} \geq R_{syn}$, they expect a higher rate of return from the standalone deal because the benefits they get from syndication are not enough to compensate for the downside from dilution. Vice versa when $R_{sta} < R_{syn}$. We show that which option is better depends on their ability $A_{vc}$. As the true ability of a VC is known to itself, they make their best choice conditioning on $A_{vc}$. Therefore, to get the optimal decisions for these VCs, it is crucial to solve for the inequality $R_{sta} \geq R_{syn}$ as a function of $A_{vc}$. Both returns $R_{sta}$ and $R_{syn}$ are increasing functions of $A_{vc}$. We prove the existence and uniqueness of the solution by showing the ratio $R_{syn}/R_{sta}$ is also increasing in $A_{vc}$ under mild assumptions. Therefore, for low ability (small $A_{vc}$) VCs, the optimal option is to invest in the entrepreneur alone. As $A_{vc}$ gets large enough for the ratio to exceed one, the best choice switches from standalone investment to proposing syndication. We leave the rigorous proof to the appendix. □

When the lead VC(s) receive a low-ability initial signal, they will always reject the deal since the return is guaranteed to be 0. When a high initial signal is observed, they choose between investing alone and syndication. Which choice generates a higher rate of return depends on the ability of the VC, $A_{vc}$, given all other parameters such as the added-value by sharing expertise $v$, the proportion of high-ability entrepreneurs in the population $\alpha$, the expected realized value of a high ability startup with no syndication $R_h$, and the general difficulty for low-ability entrepreneurs to pretend to be high types $p$. As previously discussed, there is a trade-off between (1) the benefits of higher added value from cooperation and a better selection process due to the additional observed signal; and (2) the cost of dilution by introducing more potential investors. We discuss this as well as other empirical implications in more detail in the following subsection, before which we make the following technical remark on an implicit assumption used in the calculations.

**Remark.** *Throughout this section, we assume that venture capital investors are not financially constrained. If we instead assume that they have binding financial budgets, then the conditional expected return of syndication will be*

$$R_{syn} \mid A_{vc,peer} = \frac{\exp\left\{-(1-A_{vc})\right\}(R_h+v)\,\alpha}{\alpha+(1-\alpha)\,p^2\exp\left(-A_{vc}\right)\exp\left(-A_{vc,peer}\right)} + (1-\exp\left\{A_{vc}-1\right\})r,$$

*where the additional return is from the outside investment opportunity. Our main results of the optimal decisions of VCs and the model implications do not change in this case, as long as the return of the outside investment opportunity, $r$, is small enough compared with the return of high-type startups, $R_h$ (and thus $R_h + v$). However, if the VCs have great outside options, the quality of which may also depend on the abilities $A_{vc}$, the syndication return will be a much more complicated function of $A_{vc}$, so is the optimal choice between standalone investment and syndication. This is beyond the scope of this paper and we leave it for future research.*

## Empirical implications

### Success rate and syndication

First, we calculate the success rates of syndicated deals and compare them with those of standalone deals. The model suggests that the syndicated deals will have a higher probability of successful exits due to the observation of the second signal from peers.

**Corollary 1.1.** *The success rate of standalone deals is lower than that of syndication deals, i.e.,*

$$\mathbb{P}\left\{A_e = h \mid S_0 = h, \mathrm{sta}\right\} < \mathbb{P}\left\{A_e = h \mid S_0 = S_1 = h, \mathrm{syn}\right\}.$$

For the sake of conciseness, we again relegate the rigorous proof to the appendix. Corollary 1.1 suggests that the observed success rate of the syndicated deals should be higher than the standalone deals. We test this empirical implication by regressing the success and failure indicators on the syndication indicator, controlling for the location fixed effects and the interaction between year and industry fixed effects.[3] Tables 1.17 and 1.18 report the results of the empirical tests. Both results indicate that syndicated deals perform significantly better than standalone deals regarding success and failure rates. Table 1.17 suggests that syndicated deals have a 9.65% higher success rate, and Table 1.18 shows that the syndicated deals are 15.3% less likely to end up in bankruptcy.

### Rate of return and syndication

Next, we compare the expected return of the standalone and syndicated deals, conditional on the deals being made. The model implies that the syndicated deals have higher overall returns compared with standalone deals.

**Corollary 1.2.** *Under the previous setting of the VC decision-making process, the overall expected return for syndicated deals is higher than standalone deals given that the investment is made.*

To test this empirical implication, we use the deal size and the startup's last round post-valuation to compute the (monthly) internal rate of return (IRR) for each equity round deal. The IRR is calculated from the following equation

$$\text{deal size} \times (1 + \mathrm{IRR})^{\#\text{ of months}} = \text{last post-valuation}. \tag{1.10}$$

[3] Note that the test on failure is not redundant given the test on success because there are many ongoing startups whose outcomes have not been realized yet. We define an entrepreneur as successful if and only if it goes public or gets acquired, and define failure as if and only if it files bankruptcy. It is not necessary for ongoing startups to be either successful or failed.

Table 1.17: Success exit indicator and syndication indicator

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | succeed | succeed | succeed |
| syndication | 0.0998*** | 0.0962*** | 0.0965*** |
| | (0.00290) | (0.00302) | (0.00302) |
| Constant | 0.144*** | 0.148*** | 0.148*** |
| | (0.00191) | (0.00226) | (0.00226) |
| location FE | | ✓ | ✓ |
| year FE | | ✓ | ✓ |
| industry FE | | ✓ | ✓ |
| year×industry FE | | | ✓ |
| Observations | 72,842 | 66,659 | 66,658 |
| R-squared | 0.016 | 0.115 | 0.118 |

*Notes:* This table summarizes the correlation between whether a deal is syndicated and whether it successfully exits. It reports the results of regressing the successful exit indicator on the syndication indicator. Column (1) is the OLS regression with no fixed effects controlled. Column (2) controls for location fixed effects, deal year fixed effects, and industry fixed effects. Column (3) further controls for the interaction between deal year and industry fixed effects. All regressions are at the deal level. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.18: Failure indicator and syndication indicator

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | bankruptcy | bankruptcy | bankruptcy |
| syndication | -0.198*** | -0.153*** | -0.153*** |
| | (0.00299) | (0.00288) | (0.00287) |
| Constant | 0.312*** | 0.270*** | 0.271*** |
| | (0.00252) | (0.00215) | (0.00215) |
| location FE | | ✓ | ✓ |
| year FE | | ✓ | ✓ |
| industry FE | | ✓ | ✓ |
| year×industry FE | | | ✓ |
| Observations | 72,842 | 66,659 | 66,658 |
| R-squared | 0.060 | 0.134 | 0.138 |

*Notes:* This table summarizes the correlation between whether a deal is syndicated and whether it fails. It reports the results of regressing the bankruptcy indicator on the syndication indicator. Column (1) is the OLS regression with no fixed effects controlled. Column (2) controls for location fixed effects, deal year fixed effects, and industry fixed effects. Column (3) further controls for the interaction between deal year and industry fixed effects. All regressions are at the deal level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We regress the imputed IRR on the syndication indicator, controlling for the location fixed effects and the interaction between year and industry fixed effects. Table 1.19 shows that syndicated deals have a significantly higher monthly IRR than standalone deals, which aligns with our model prediction.

Table 1.19: Rate of return and syndication indicator

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | IRR | IRR | IRR |
| syndication | 0.335*** | 0.121** | 0.123** |
| | (0.0465) | (0.0519) | (0.0520) |
| Constant | -0.514*** | -0.343*** | -0.344*** |
| | (0.0460) | (0.0379) | (0.0379) |
| location FE | | ✓ | ✓ |
| year FE | | ✓ | ✓ |
| industry FE | | ✓ | ✓ |
| year×industry FE | | | ✓ |
| Observations | 30,232 | 27,440 | 27,439 |
| R-squared | 0.002 | 0.016 | 0.019 |

*Notes:* This table summarizes the correlation between whether a deal is syndicated and the internal rate of return (IRR) of the deal. For each equity round deal, the IRR is calculated using the deal size and the startup's last round post-valuation from Equation (1.10). This table reports the results of regressing the imputed IRR on the syndication indicator. Column (1) is the OLS regression with no fixed effects controlled. Column (2) controls for location fixed effects, deal year fixed effects, and industry fixed effects. Column (3) further controls for the interaction between deal year and industry fixed effects. All regressions are at the deal level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

### Extension: syndication, ability and centrality

The two expected returns in Equations (1.8) and (1.9) explicitly show the trade-off between more added value plus better selection, and possible dilution from syndication. In our model, the ability of the VCs is exogenously given. As time passes, VCs who syndicate more can accumulate a reputation for higher success rates and overall rates of return. They also learn more from the cooperation process with each other, compared to those who lack such co-investment opportunities. It is likely that syndication will amplify the advantages of the high-ability VCs, in that it helps them with a higher propensity to syndicate. The more they syndicate, the higher the success rate and rate of return they realize, and the more central they become in the VC network. This VC may attract more followers and gain greater bargaining power in the future.

This indicates that more central VCs may have larger benefits from syndication in the future, leading to higher propensities of syndicated investments. This implies that future syndication decisions may be positively correlated with past centrality. At the same time, the inequality in centrality also increases with time since those who are central have a higher probability of building up relationships with other VCs in the future compared with the isolated or less central ones. We explore both implications empirically using our data.

Table 1.20 reports the results of regressing the syndication indicator on the 1-year lag of degree centrality of each VC on the subsample of lead/sole VC investors, controlling for the location fixed effects and the interaction between year and industry fixed effects. It shows a strong positive correlation between the past centrality of VCs and the probability of syndication. One standard deviation higher in past centrality is correlated with a 2.06% higher probability of leading a syndicated deal.

Moreover, we also explore the trend of inequality in centrality using two different measures of inequality: the Gini coefficient of the centrality and the proportion of total centrality accounted for by the top 1 percentile VC. We calculate these two inequality measures for each year during 2009 - 2021 and plot the trend. Figures 1.6 and 1.7 show the trends. The overall trends are increasing in centrality inequality, although it seems that the increasing curve has flattened in recent years. This may suggest that some learning processes can happen, where VCs become increasingly willing to compromise some of the current returns to achieve higher centrality and ability so that they can benefit more in the future. These implications are beyond the scope of our model and are left for future research.

## 1.7 Conclusion

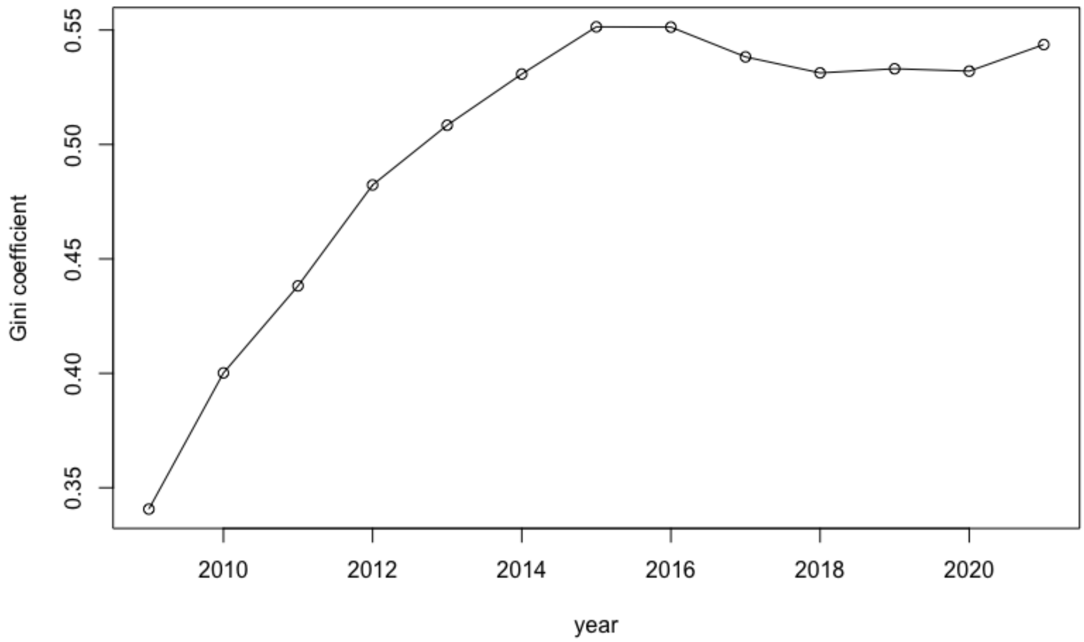
This paper studies the network effects in venture capital financing through empirical tests and a theoretical model. First, as the first paper in the VC literature that estimates the network effects by simultaneously modeling the network formation and the outcome, we

Table 1.20: Syndication and past centrality

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | syndication | syndication | syndication |
| l1.degree centrality | 3.780** | 2.548** | 2.576** |
| | (1.471) | (1.268) | (1.242) |
| Constant | 0.438*** | 0.447*** | 0.444*** |
| | (0.0175) | (0.0168) | (0.0160) |
| location FE | | ✓ | ✓ |
| year FE | | ✓ | ✓ |
| industry FE | | ✓ | ✓ |
| year×industry FE | | | ✓ |
| Observations | 1,112 | 1,107 | 1,101 |
| R-squared | 0.008 | 0.134 | 0.261 |

*Notes:* This table summarizes the correlation between VC's past degree centrality and its propensity to lead a syndicated deal. The outcome variable is an indicator of whether the deal is syndicated, and the l1.degree centrality is the past degree centrality (one-year lag) of each VC. We regress the outcome variable of interest on the past degree centrality, restricting to the subsample consisting of only lead/sole investors. Column (1) is the OLS regression with no fixed effects controlled. Column (2) controls for location fixed effects, deal year fixed effects, and industry fixed effects. Column (3) further controls for the interaction between deal year and industry fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

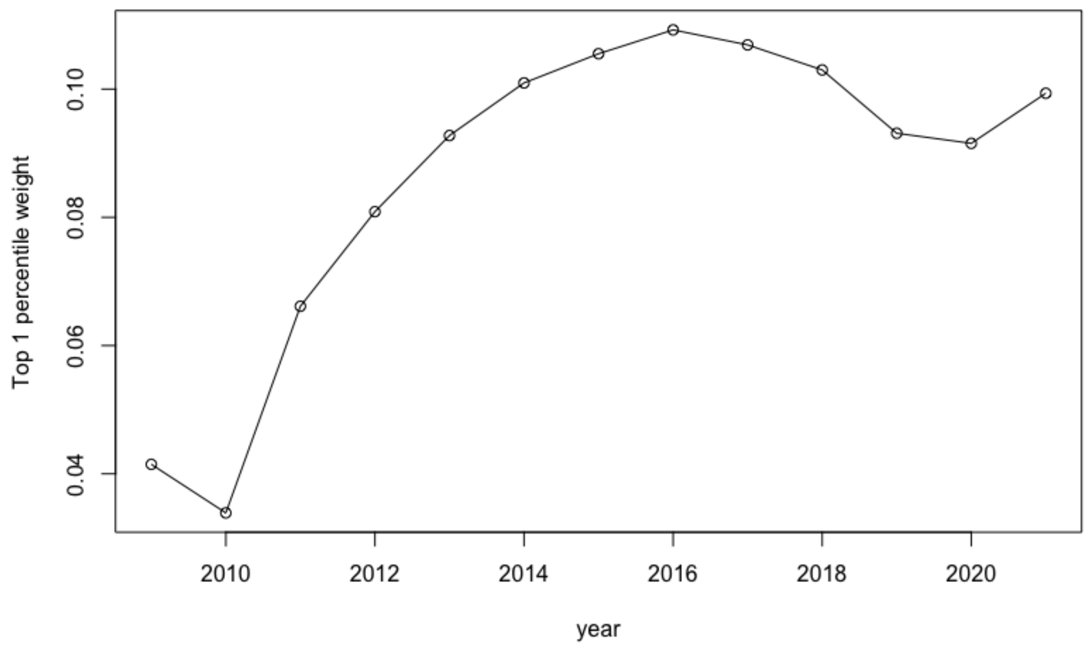
Figure 1.6: Trend of Gini coefficient of centrality




*Notes:* This figure plots the trend of the Gini coefficient of degree centrality in our dynamic VC network.

utilize the econometric theory proposed in Goldsmith-Pinkham and Imbens (2013) to directly tackle such endogeneity problems as the contextual bias and the homophily bias in the VC network. We explicitly model the network formation process and such outcomes as a VC's performance as functions of features of VCs in the network. This takes not only the observed features of a VC itself into account, but also its unobserved features, the peers' observed and unobserved features, and the dyad-specific features of each pair of the VCs. All the unknown parameters and the unobserved features are incorporated and estimated using the Bayesian method. After using the observed data to update the posterior distribution, we get the significantly positive endogenous network effects of 0.127, which means that if we are able to directly permute the average success rate of the connected peers of a VC to be 1% higher, the network effect would lead to an increase of 0.127% in the success rate of this VC. The magnitude of the network effects is large compared to the scale of the success rate in our observed data. Through a more comprehensive study on heterogeneous network effects, we find that positive network effects exist in both the networks within the same industry and across different industries. Therefore, the sharing of expertise and resources between connected VCs in different industries matters greatly, as well as the information and signal

Figure 1.7: Trend of top 1% proportion of centrality




*Notes:* This figure plots the trend of the total proportion accounted for by the top 1% degree centrality VCs in our dynamic VC network.

sharing between connections within the same industry.

Second, we use the quasi-experimental design, a generalized difference-in-differences method, to provide a more intuitive way of understanding where these positive effects come from. Results in this empirical study show that when an extremely successful event such as IPO happens, not only does the VC that owns the IPO startup benefit from this positive event, but its peer VCs' performance is also significantly improved in terms of the ability to attract new funding resources, a higher success rate, and lower probability of bankruptcy.

Last but not least, we build a theoretical model to illustrate the channels through which the network effects contribute. Our theory explicitly models the VCs' trade-off problem when they choose between syndication and standalone investment. In the model, we show that the syndication attracts VCs for two main reasons: the pool of signals to get a more accurate evaluation of the entrepreneur's ability, and the high added value achieved by the sharing of expertise and cooperation between VCs. These benefits come at a cost of dilution since once a syndicated deal is achieved, multiple VCs co-invest in the startup together and the lead VC no longer takes all the pie. We prove that the optimal decision depends on a VC's ability. Moreover, the syndicated deals have a larger propensity for success and higher returns given that the deal is achieved. These empirical implications of the model are tested using observed Pitchbook data. All empirical results are in line with the theoretical model. We also provide extensions on how the network centrality evolves and explanations of why the more central VCs are likely to perform better in the VC financing market.

# Chapter 2

# Common Ownership in Venture Capital Financing

## 2.1 Introduction

Common ownership is widespread in many industries in the US economy and has been increasing in recent years (Azar et al., 2018). Many companies with competing products also share the same set of investors. However, the economic literature has not come to a consensus on the effect of common ownership. Previous literature mainly focuses on the common ownership by large passive investors—large institutions such as BlackRock and Vanguard that employ diversified investment strategies (Azar et al., 2018; Davis, 2013; Fichtner et al., 2016). This study focuses on common ownership in venture capital financing and its effect on the performance of young entrepreneurial startups for two main reasons. First, startups contribute significantly to the business dynamism in the US economy (Lerner and Nanda, 2020). The venture capital investor has a lead role as a supporter of entrepreneurship and startup companies (Bürer and Wüstenhagen, 2008) as they select and screen promising entrepreneurs and provide them with capital, training, and resources to help them realize their potential (Baum and Silverman, 2004). Many large successful companies are backed by venture capital investors before they go public (Lerner and Nanda, 2020). The venture capital-backed startups contribute a large share to innovation in the US economy (Gompers and Lerner, 2004; Kaplan and Lerner, 2010).

Moreover, venture capital investors have relatively strong control rights during the operation process of the startups, participating greatly in their management processes as well (Kaplan and Stromberg, 2001). Compared with the passive investors (Azar et al., 2018), venture capital investors are sophisticated active investors that are more likely to be attentive to spillovers across startups in their portfolio, since they strategically build their portfolio of startups (Li et al., 2021). Information and ideas are vital for startups. Common investors of startups in the same industry act as connections between startups to accelerate the sharing of information among them. Additionally, common ownership in venture capital financing is widespread—65.07% of the startups in our data sample are commonly owned at the industry code level, which is the finest industrial classification level. Several past papers have documented the positive effect of common ownership on innovation, startup performance, and growth (Li et al., 2021; Eldar et al., 2020; González-Uribe, 2020; Lindsey, 2008).

We test the effect of common ownership on both the long-term and short-term performance of startups. We follow the potential outcome framework and aim to estimate the average treatment effect of common ownership on the population of startups. Due to the endogeneity problem of common ownership status, we cannot draw causal conclusions from our reduced form results that commonly owned startups have better long-term performance measured by exiting status. We utilize the matched-pair design as our identification strategy and estimate the spillover-type effects of common ownership on startup performance. We first construct the time-varying common ownership pool using information from the financing history of all startups in our sample. We then answer the question: when a successful funding event occurs to a startup in a common ownership pool, what is the effect of it on the commonly owned startups compared with startups that are not in the common ownership pool? Empirical results show that when successful funding occurs to a startup's peer in the same common ownership pool, this startup will get a new round of financing with a 1.31% higher probability in 180 days, 2.40% higher probability in 365 days, and 3.37% higher probability in 730 days, controlling for several fixed effects. The magnitude of the effect increases with measures of how successful the event to its peer is. Moreover, we find the effects are heterogeneous across different industries. The healthcare industry enjoys the most considerable spillover effect of common ownership, which is a 4.13% higher probability of getting a new round of funding within 180 days. The effects on consumer products and services, business products and services, information technology, energy, and financial services are also significantly positive. We argue with caveat that this can be driven by differences across industries, since the sharing of new information and creative ideas may matter more to industries like healthcare, compared with more traditional industries such as materials and resources that are barely affected by common ownership status.

There are several possible channels through which common ownership positively impacts startup performance. Startups sharing a common venture capital investor are more likely to form alliances (Lindsey, 2008) and share innovation resources (González-Uribe, 2020). Common ownership is also shown to affect innovation, specifically in the pharmaceutical industry (Li et al., 2021). In addition to the effects of successful funding events, we also conduct the same set of analyses on the spillover effect of common ownership when a negative event such as the closure occurs to a startup in the pool. The results show that negative events do not have significant spillover effects on startups within the same common ownership pool, indicating that the positive spillover effects are not likely to be driven merely by reputation.

Our contribution to the literature is threefold. First, when documenting the stylized facts, we employ the proportional hazard model with competing risks in survival analysis (Fine and Gray, 1999) to learn the difference in performance between commonly owned startups and not commonly owned ones. Survival analysis naturally takes the life span of startups into consideration, deals with the time dimension, and considers the three types of exit status (IPO, acquisition, and closure) simultaneously by treating them as competing risks. Second, the matched-pair design used in this paper addresses the endogeneity problem without imposing strong identification assumptions such as the exclusion restriction assumption used in the instrumental variable (IV) setting.[1] Third, our estimated effects of common ownership on both long-term and short-term performance provide empirical evidence of positive spillover effects that when successful events occur to a startup, other startups in the same common ownership pool also benefit from them. However, when a negative event occurs, we observe no evidence of the spillover effects. These results shed light on the possible mechanisms through which common ownership matters and indicate that the effects are not merely driven by reputation.

The remainder of this paper is organized as follows: Section 2.2 introduces data used in this paper and the construction of a time-varying common ownership pool. Section 2.3 documents reduced-form stylized facts on the difference in performance between startups that are commonly owned and those that are not. Section 2.4 describes the matched-pair design model we use as our identification strategy and reports our main empirical results. Section 2.5 discusses potential mechanisms and concludes.

## 2.2 Data and the time-varying common ownership pool

### Pitchbook Data

In this project, we use the information from the Pitchbook database on the startups, their deals, and deal investors. Pitchbook is a Software as a Service company that delivers data, research, and technology covering the private capital markets, including venture capital, private equity, and mergers and acquisitions transactions. It uses machine learning and natural language processing models to review publicly available sources and collects information on companies, deals, and deal investors in the private market, with the entrepreneur capital market as our focus.

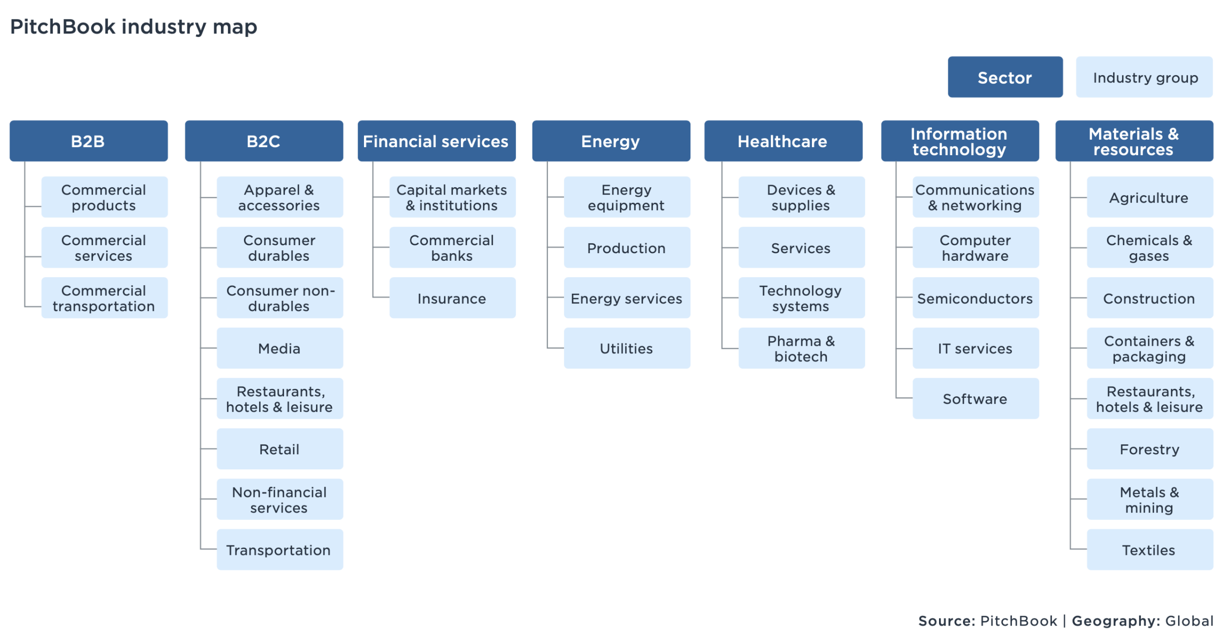
Pitchbook provides detailed features on startups, such as their industry, foundation year, detailed location and address, the total number of employees, various measures of profit and revenue, their financing history, etc. As for the financing history, detailed information on

[1] A commonly used IV of common ownership status in the literature is the geographical distance between startups and investors. While the exogeneity of this IV seems reasonable, it is hard to believe geographical distance and locations affect the outcomes of startups only through the common ownership channel. This violates the exclusion restriction, a crucial assumption for the validity of the IV method.

each financing round is also available in the database, such as the date of a deal, deal type, the stage of a deal if it is an equity round, the debt type of a deal if it is a debt round, the deal size, pre-valuation and post-valuation of a deal, and deal investors. We also observe features of deal investors such as their primary investor type, preferred industries, whether they are the lead/sole investors in each deal, the name and background of the lead partner, the total amount of assets under management, and investment size.

The industry is a crucial feature for constructing this study's common ownership pool. Pitchbook has its own industrial classification system different from SIC or NAICS. Their system has three levels: industry sector, industry group, and industry code. Figure 2.1 shows the first two levels of the Pitchbook industry map. The Pitchbook industrial classification system divides the startups into seven industry sectors: Business products and services (B2B), Consumer products and services (B2C), Financial Services, Energy, Healthcare, Information Technology, and Materials & Resources. Within each industry sector, there are several industry groups, as shown in the figure. Each industry group is further divided into more detailed industry codes, which is the finest level of classification in Pitchbook data.

Figure 2.1: Pitchbook industry map




*Notes:* This figure shows a map of the industrial classification system provided by Pitchbook.

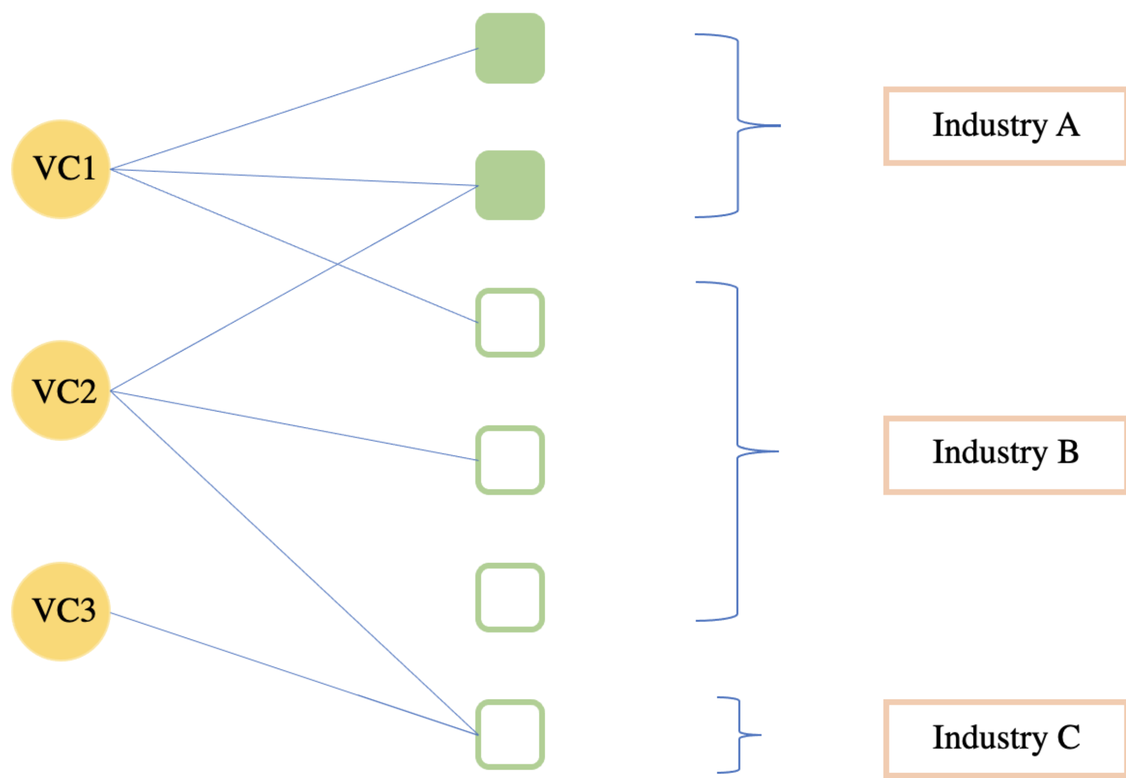
## Construction of the time-varying common ownership pool

As described in the introduction, the phenomenon of common ownership is widespread in the US venture capital market. Many investors choose to invest in multiple startups in the same industry simultaneously. To formally define common ownership in our context and to study the effect of common ownership on the performance of these startups, we first construct a time-varying common ownership pool. A venture capital investor is defined as commonly owning multiple startups if and only if it simultaneously invests in $\geq 2$ startups in the same industry as a leading/sole investor. Figure 2.2 shows the sketch of the definition of common ownership. In the figure, the yellow circles stand for venture capitalists (VCs), and the green squares stand for startups, among which the solid green squares are commonly-owned startups while the hollow green squares are not. For instance, the first two startups are in the same industry A and invested by a common VC1; thus, they are defined as commonly owned. The second and the third startups are not considered commonly owned, as they are not in the same industry although simultaneously invested by the same VC1. Similar logic applies to the remaining startups. As we introduce in the previous subsection, the Pitchbook industrial classification system has three levels of classification. In this project, we define two startups as in the same industry if and only if they have the same *industry code*. We can also change the definition by using industry groups, and the results are robust to the levels we use.

Utilizing the information on deals, we define a pool of commonly owned startups as all startups in the same industry that are simultaneously invested by the same venture capitalist as a lead or sole investor in the financing rounds. The data used in this project contains all backed US-based startups founded between 2008 and 2017 in the Pitchbook database. There are 71,299 startups in total, with 38,572 having the ownership and investor information available. When constructing the common ownership pool, we use only the lead/sole investor of each deal for the syndicated deals since the lead usually takes the most crucial role. If there are multiple leads in one deal, we use all of them, and if there is no information on who is the lead in a deal, we use all of the investors for that deal. We also focus only on the early round venture capital investments, i.e., the equity financing rounds no later than stage B venture capital round. (This includes the angel or seed round, the stage A venture capital round, and the stage B venture capital round.) Among the 38,572 startups with ownership information, 87.73% are commonly owned at the industry sector level, 80.25% are commonly owned at the industry group level, and 65.07% are commonly owned at the industry code level. Even at the finest level, common ownership is very widespread among startups.

Given that there are many entrances and exits of the venture capital investors in startup financing, the pool is time-varying for each venture capitalist, and the startups in the same pool share the same industry code. A startup enters the pool of a venture capitalist of its industry code when it is first invested by this venture capitalist (as a lead or sole investor). A startup exits the pool whenever it is out of business. Taking the time-variant nature of

Figure 2.2: A sketch of common ownership definition




*Notes:* This figure illustrates the definition of common ownership. The yellow circles on the left stand for VCs, and the green squares in the middle stand for startups. The solid green squares are commonly-owned startups, while the empty green squares are not. The orange squares on the right demonstrate their industries.

the common ownership status into consideration, we construct the panel of the common ownership pool and conduct empirical analyses based on this. Figure 2.3 shows a snapshot of the time-varying common ownership pool illustrating what it looks like.

## 2.3 Stylized facts

In this section, we first look at the comparison between the performances of commonly owned startups and not commonly owned ones. We use the following types of exit status as measures of their long-term performance:

Figure 2.3: A snapshot of the time-varying common ownership pool

| | industry | investor | start_date | end_date | companies | num_companies |
|---|---|---|---|---|---|---|
| **67** | Energy Exploration | 54789-49 | 2019-08-26 | current | ['163589-41'] | 1 |
| **68** | Application Software | 10217-17 | 2008-01-01 | 2010-01-01 | ['54040-42'] | 1 |
| **69** | Application Software | 10217-17 | 2010-01-01 | 2010-08-25 | ['54040-42', '60827-77'] | 2 |
| **70** | Application Software | 10217-17 | 2010-08-25 | 2010-10-25 | ['60827-77'] | 1 |
| **71** | Application Software | 10217-17 | 2010-10-25 | 2012-01-01 | ['55301-77', '60827-77'] | 2 |
| **72** | Application Software | 10217-17 | 2012-01-01 | 2012-06-26 | ['97666-12', '55301-77', '60827-77'] | 3 |
| **73** | Application Software | 10217-17 | 2012-06-26 | 2012-09-01 | ['97666-12', '55301-77', '60827-77', '54744-67'] | 4 |
| **74** | Application Software | 10217-17 | 2012-09-01 | 2012-12-10 | ['54744-67', '94918-78', '60827-77', '97666-12... | 5 |
| **75** | Application Software | 10217-17 | 2012-12-10 | 2013-01-01 | ['54744-67', '94918-78', '60827-77', '97666-12... | 6 |
| **76** | Application Software | 10217-17 | 2013-01-01 | 2013-01-11 | ['54744-67', '60652-90', '94918-78', '60827-77... | 7 |
| **77** | Application Software | 10217-17 | 2013-01-11 | 2013-02-01 | ['54744-67', '60652-90', '94918-78', '60827-77... | 8 |

*Notes:* This figure shows a snapshot of the time-varying common ownership pool we construct using the Pitchbook data. For each industry and investor, the common ownership pool contains a set of startups owned by the investor within a certain period (from the start to the end date). The common ownership pool is time-varying due to new investments and exits of the venture capitalists.

- `ipo` is an indicator of whether the startup goes public,
- `acq` is an indicator of whether the startup is acquired, and
- `closed` is an indicator of whether the startup closed.

We run the following reduced-form cross-sectional regression to test the correlation between whether a startup is commonly owned and its long-term performance,

$$\texttt{outcome} = \alpha_{\text{state}} + \lambda_{\text{ind,foundyear}} + \tau \cdot \texttt{common\_code} + \gamma^{\mathsf{T}} X + \epsilon,$$

where $\alpha_{\text{state}}$ are the location fixed effects at the state level, $\lambda_{\text{ind,foundyear}}$ are the two-way interacted fixed effects between industry and found-year fixed effects, `common_code` is the right-hand-side variable that indicates whether a startup is commonly owned or not at the industry code level, $\texttt{outcome} \in \{\texttt{ipo, acq, closed}\}$ is the measure of long-term performance, and $X$ contains several features of the startups as control variables. The main features of startups we use in the regressions are the logarithm of the total number of employees and the indicator of whether the startup is invested in by relatively experienced investors. To measure the investor's experience, we compute the total number of investment rounds and the total number of early rounds made by each investor. For each startup, we then average over all its investors' scores of experience as a measure of how experienced its investors are.

We control for whether the average number of rounds of all investors of the startup is above the median in the regressions.

Table 2.1 reports the results of regressing the outcome on the common ownership indicator when the total number of investment rounds is used in the definition of experienced investors. The results show that commonly owned startups have a 0.59% higher probability of going public and a 3.78% higher probability of being acquired, controlling for the covariates, the location fixed effects, and the two-way interacted fixed effects between industry and found year. Results in the first two columns indicate that common ownership is associated with better long-term performance in terms of successful exits. Column (3) of Table 2.1 shows that common ownership has little correlation with the closure of startups. All results are robust to the same set of regressions using the total number of early rounds in the definition of experienced investors, and are reported in the appendix.

In addition to the reduced-form cross-sectional facts, we also utilize the time dimension of the operating status of each startup, and use models in survival analysis to compare the performance of commonly owned startups versus not commonly owned startups. Define the time to the first major event as $T$ and call it exit time. We treat an exit as occurring whenever the startup goes public, gets acquired, or closes, whichever occurs first. Since there are three types of possible exit events for the startups, we use the survival analysis with the competing risks framework proposed in Fine and Gray (1999). These three types of exit events are the so-called competing risks in survival analysis. Define the cumulative incidence function for each possible exit event $j \in \{\texttt{ipo, acq, closed}\}$ as

$$F_j^z(t) = \mathbb{P}\left(T(z) \leq t, J(z) = j\right),$$

where $t$ is the time point of interest, $z \in \{0, 1\}$ is the indicator of whether a startup is commonly owned or not, $T(z)$ and $J(z)$ denote the exit time of a startup and the type of exit event under common ownership status $z$. In words, the cumulative function $F_j^z(t)$ stands for the probability that a startup exits before time $t$ due to the type of event $j$ under common ownership status $z$. We compare two cumulative incidence functions $F_j^1(t)$ and $F_j^0(t)$ to see the correlation between common ownership status and the probability curves of going public, getting acquired, and closing. The proportional hazard model in Fine and Gray (1999) also allows for covariate adjustment in the estimation procedure of cumulative incidence functions. Thus, we incorporate industry fixed effects, found year fixed effects, and location fixed effects into the estimation model. Figure 2.4 plots the cumulative incidence functions of the three types of exit event of interest, namely `ipo`, `acq` and `closed`. The first subplot in Figure 2.4 shows that the commonly owned startups have a higher propensity of going public at $t \geq 3$, indicating the positive correlation between common ownership and successful exit. The second subplot illustrates a larger effect of common ownership on the probability of acquisition. Commonly owned startups always have a higher probability of being acquired. As for the correlation with closure, there is no obvious difference in probabilities of closure between commonly owned startups and not commonly owned ones,

Table 2.1: Performance and common ownership indicator

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | ipo | acq | closed |
| common_code | 0.00587* | 0.0378*** | 0.00310 |
| | (0.00309) | (0.0105) | (0.00930) |
| Constant | -0.0291*** | 0.117*** | 0.365*** |
| | (0.00323) | (0.0109) | (0.00972) |
| control variables | ✓ | ✓ | ✓ |
| location FE | ✓ | ✓ | ✓ |
| industry×found year FE | ✓ | ✓ | ✓ |
| Observations | 8,698 | 8,698 | 8,698 |
| R-squared | 0.067 | 0.045 | 0.093 |

*Notes:* This table reports the relationship between long-term performance and common ownership indicator. Results are from the regressions

$$\texttt{outcome} = \alpha_{\text{state}} + \lambda_{\text{ind,foundyear}} + \tau \cdot \texttt{common\_code} + \gamma^{\mathrm{T}} X + \epsilon,$$

where `outcome` $\in$ {`ipo, acq, closed`} is the measure of long-term performance and `common_code` is the treatment indicator of whether a startup is commonly owned. We control for startup features such as the logarithm of the total number of employees and whether the average number of rounds of its investors is above the median. Location fixed effects and the two-way interacted fixed effects between industry and found-year are also included in the regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

as shown in the third subplot. The two cumulative incidence functions are very close to each other with an almost negligible difference.

Stylized facts in both reduced-form cross-sectional and survival analysis suggest seemingly beneficial effects of common ownership on the long-term performance of the startups. However, these pieces of empirical evidence are not enough to conclude that the positive effect is *caused* by common ownership, as there are possibly unobserved confounding variables of ownership status and performance, even after controlling for the observed features and fixed effects. For instance, the abilities of the venture capitalists can be an unobserved confounding variable. High-ability venture capitalists may have more adequate funding resources and better preferences on which industry is promising in the following several years, so they may invest more in this industry, leading to a higher propensity of simultaneous investment in the startups in this industry. Additionally, they are equipped with higher abilities that could add more value to the startups and lead to better performance of startups in their portfolio. Similar stories may exist for other unobserved confounding variables. Therefore, we conduct more rigorous causal inference and discuss our identification strategy more in the following section.

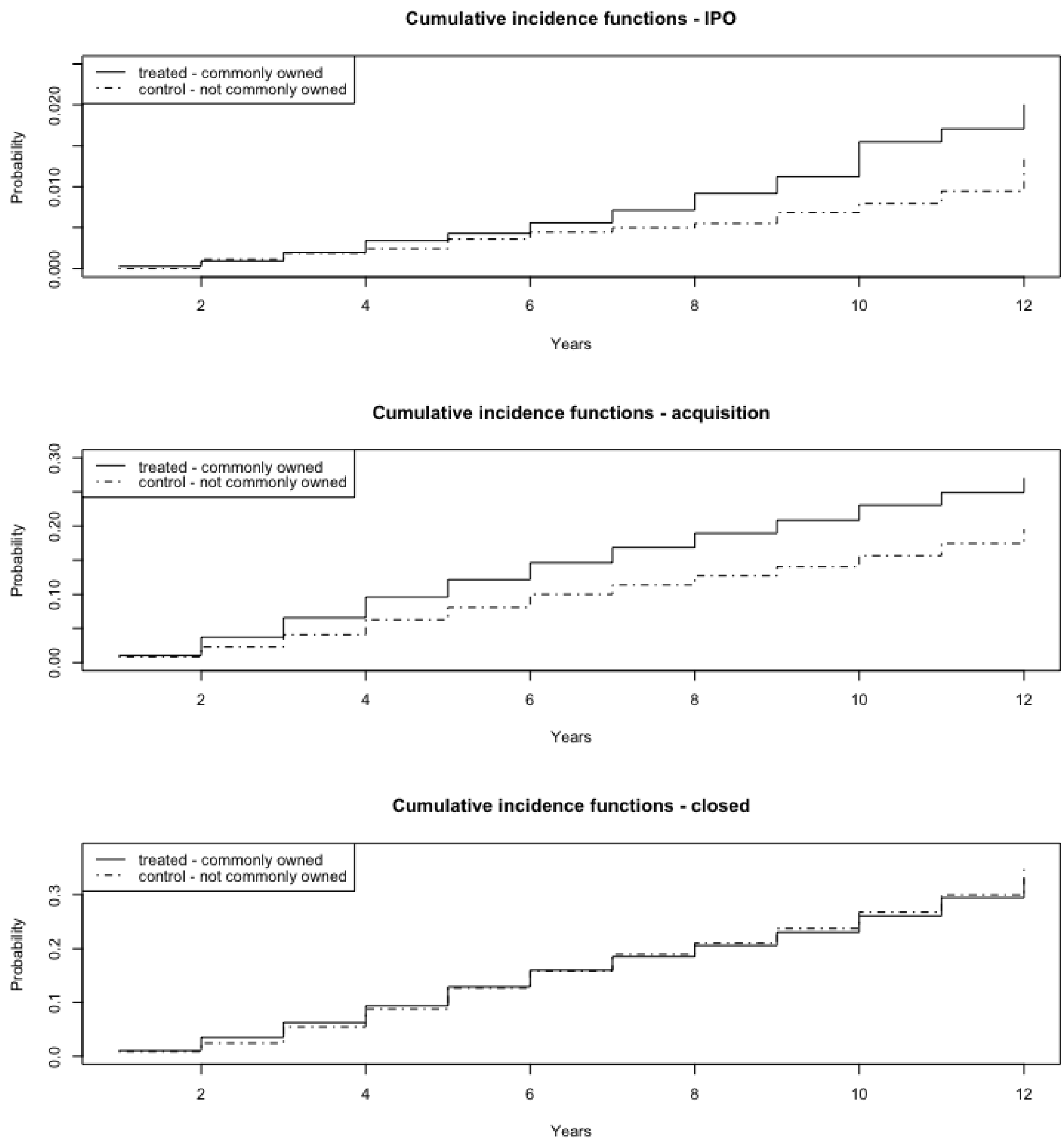
## 2.4 Identification and estimation

We follow the potential outcome framework (Neyman, 1923; Rubin, 1974) under the super population setting. In our setup, there is a binary treatment $Z_i \in \{0, 1\}$ where $Z_i = 1$ if the startup is in the common ownership pool and $Z_i = 0$ if not. Each startup $i$ has two potential outcomes $Y_i(1)$ and $Y_i(0)$, where $Y_i(1)$ stands for the potential outcome under treatment and $Y_i(0)$ under control. The fundamental problem of making causal inferences is that we only observe one of the two potential outcomes. The observed outcome $Y_i$ depends on the realized common ownership status of a startup, i.e., $Y_i = Z_i Y_i(1) + (1 - Z_i)Y_i(0)$. For each startup, the individual treatment effect of common ownership is $\tau_i = Y_i(1) - Y_i(0)$, the comparison in potential outcomes between treatment and control. Our main causal parameter of interest is the average treatment effect of the whole population:

$$\tau = E\{\tau_i\} = E\left\{Y_i(1) - Y_i(0)\right\}.$$

Ideally, we want to learn the average treatment effect from the experiment in which we can manipulate the treatment assignment of each startup. In our setting, the ideal experiment is to randomly assign some startups to the common ownership pool by assigning them the same venture capital investors and the remaining startups into the control group. Such an experiment is obviously unrealistic and we need to answer the question through observational studies where the treatment is not randomly assigned. In our observational study, to achieve identification, we first impose the classic stable unit treatment value assumption (SUTVA, Imbens and Rubin 2015) in the potential outcome framework:

Figure 2.4: Cumulative incidence functions




*Notes:* These three figures plot the cumulative incidence functions for each possible exit event $j \in \{\texttt{ipo, acq, closed}\}$, respectively. In each subplot, the solid line plots the cumulative incidence function under the treatment status `common_code` $= 1$, while the dashed line plots that under the control group.

**Assumption 2.1** (Stable unit treatment value assumption (SUTVA)). *Assume that*

1. *There is only one version of the treatment assignment for the startups, and*
2. *one startup's response depends only on the treatment assignment of itself, independent of the treatments assigned to any other startups in the population.*

Another crucial identification assumption of the average treatment effect in observational studies is the selection on observables:

**Assumption 2.2** (Selection on observables). *The treatment assignment is conditionally independent of potential outcomes conditioning on the observed covariates $X_i$, i.e.,*

$$Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\} \mid X_i$$

*for any $i = 1, 2, \ldots, n$.*

Under Assumptions 2.1 and 2.2, the average treatment effect is nonparametrically identifiable and can be estimated through several different estimators. Specifically, we use the nearest neighbor matching method to construct our estimator (Rubin, 1973; Rosenbaum and Rubin, 1983). The idea of matching is that for each treated unit, we match it with its nearest neighboring control unit(s) and measure the difference in the observed outcomes. The average of all these differences is a valid estimator of the average treatment effect under Assumptions 2.1 and 2.2. Also, it is difficult to directly measure the treatment effect of common ownership given our available data, as it is hard to believe that the common ownership treatment is exogenous, even after controlling for the observed features. Unobserved confounding variables, as well as reverse causality problems, prevent us from directly identifying and estimating the treatment effect of common ownership, i.e., the direct comparison between the two potential outcomes when the treatment variable is commonly owned or not.

In our empirical study, we take one step further to answer the question: for startups in the same common ownership pool, if one of them is successfully funded, what is the effect of this event on the other startups in the same pool? This spillover-type of parameter indirectly measures the treatment effect of common ownership. Specifically, we use the matching method to identify and estimate the effect of common ownership using the constructed time-varying pools. For the startups in the same common ownership pool, whenever company A in this pool has some good deals at time $t$, we match all the other companies in this pool to the nearest (possibly several) companies outside the pool, but in the same industry code. We then measure the difference in financing performance between the companies inside and outside the pool within a specific time window (180 days, 365 days, and 730 days).

For features used in the matching algorithm, we first require an exact match of industry code (the finest industrial classification level in Pitchbook data) and the first financing deal type of the startups. In addition, we calculate the Mahalanobis distance of every startup in the control group with the treatment startup using covariates of startups and choose the five closest control startups as the matched pairs. The covariates included in the Mahalanobis distance functions are the found year, the first financing size, the pre-and post-valuation of the last financing round up until time $t$, the deal type and deal size of the last financing round up until time $t$, and the business status up until time $t$.[2] For the missing value problems of the observed features in the Pitchbook data, we follow suggestions in Zhao and Ding (2021) by including an indicator of whether the value of the variable is missing and the interaction between this indicator and the covariate in the distance function.

For the measure of performance, we use both the long-term and short-term performance measures. For the long-term performance, we use the probability of successful exits and the probability of failure. For the short-term performance, we measure the ability of a startup to attract new funding by calculating whether the startup gets new rounds of funding within a certain time window (180 days, 365 days, and 730 days). We regress each outcome variable of interest on the treatment indicator, controlling for the matched-pair fixed effects and the four-way interaction among industry code, year founded, deal year, and location (state level) fixed effects. Our parameter of interest is the coefficient $\tau$ in the following regression

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it}, \tag{2.1}$$

where

- $\tau$ is the average treatment effect;
- $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$ or not;
- $\alpha_{\text{pair}}$ are the matched-pair fixed effects; and
- $\lambda_{\text{ind,foundyear,location},t}$ are the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$.

We report regression results using the 1-5 matching algorithm, in which we match the five nearest startups in the control group for each treated startup. All our results are robust to other numbers of matched control units, and we report results on 1-1 matching in the appendix. Table 2.2 reports the effects of common ownership on the long-term performance of

[2] The paper Gompers et al. (2020) on how venture capitalists make decision provide foundation and evidence of the reliability of the selection on observables assumption since we include the crucial factors in venture capitalist's decision-making process.

startups. The results show a 0.215% higher probability for the commonly owned startups to get IPO after their peer startups in the same common ownership pool experience a successful financing round. There is also a 0.735% higher probability of getting acquired and a 0.445% lower probability of failure.

Despite the positive effects of common ownership on a startup's long-term performance, the time horizon of outcomes matters as the treatment occurs whenever a startup's peer in the same common ownership pool has good events. Long-term performance stands for the final exit status of the startups, and thus does not incorporate the time horizon or sufficiently reflect the effects as time goes by. Therefore, we further construct the time-varying short-term performance measure: the financing performance of the startups within a 180-day, 365-day, and 730-day window from the time of the "good event". For each startup, the outcome is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window. Results in Table 2.3 show that when successful funding occurs to a startup's peer in the same common ownership pool, this startup will get new rounds of funding with a 1.31% higher probability in 180 days, 2.40% higher probability in 365 days, and 3.37% higher probability in 730 days. These results on short-term financing behavior imply a positive effect of common ownership on the startup's capacity to raise new funding.

Moreover, we further explore the heterogeneity in treatment effects of common ownership across different industries to see whether the benefits from common ownership vary across industries. We run the regression in Equation (2.1) on the subsample of each industry, replacing the four-way interacted fixed effects with the three-way interacted $\alpha_{\text{foundyear,location},t}$ and all other specification the same. Table 2.4 reports the results where the outcome variable of interest is the indicator of getting financed within 180 days. The results show that the healthcare industry has the largest treatment effects[3]—startups in the common ownership pool are 4.13% more likely to get new funding within 180 days from the time of the successful event. Following the healthcare industry are the consumer products and services (B2C), business products and services (B2B), and information technology industries, with treatment effects of 3.01%, 2.75%, and 2.06%, respectively. The positive treatment effects of common ownership are all significant with a p-value of 0.000 in these four industries. For the energy and financial services industries, the magnitudes are also around 2-3% with a slightly higher p-value. For the materials and resources industry which is more traditional, the effect is not significant at 0.1 level, indicating that the flow of information, funding, or resources across startups within the same common ownership pool rarely matters in this industry. This conclusion comes with a caveat. As shown in Table 1.1, only a small proportion of startups in our data are from the traditional industries such as materials and resources and energy, the lack of statistical power of estimated treatment effects in these industries may also be due to the small sample sizes. We leave this for future research.

[3]This is consistent with the results of Li et al. (2021), which show the positive effect of common ownership in the improvement of innovation efficiency.

Table 2.2: Effects of common ownership on long-term performance

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | ipo | acq | closed |
| treatment | 0.00215*** | 0.00735*** | -0.00445*** |
| | (0.000115) | (0.000502) | (0.000427) |
| Constant | 0.00780*** | 0.155*** | 0.121*** |
| | (4.03e-05) | (0.000175) | (0.000149) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 5 | 5 | 5 |
| Observations | 3,582,330 | 3,582,330 | 3,582,330 |
| R-squared | 0.498 | 0.415 | 0.472 |

*Notes:* This table reports the effects of common ownership on the long-term performance of the startups. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where $\texttt{outcome} \in \{\texttt{ipo, acq, closed}\}$ is the measure of long-term performance and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-5 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.3: Effects of common ownership on short-term financing behavior

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | 0.0131*** | 0.0240*** | 0.0337*** |
| | (0.000583) | (0.000666) | (0.000669) |
| Constant | 0.169*** | 0.295*** | 0.420*** |
| | (0.000204) | (0.000233) | (0.000234) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 5 | 5 | 5 |
| Observations | 3,582,330 | 3,582,330 | 3,582,330 |
| R-squared | 0.266 | 0.352 | 0.439 |

*Notes:* This table reports the effects of common ownership on the short-term financing behavior of the startups. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-5 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.4: Heterogeneous effects on financing behavior across industries

| Industry | $\hat{\tau}$ | p-value |
|---|---|---|
| Healthcare | 0.0413*** | 0.000 |
| Consumer Products and Services (B2C) | 0.0301*** | 0.000 |
| Business Products and Services (B2B) | 0.0275*** | 0.000 |
| Information Technology | 0.0206*** | 0.000 |
| Energy | 0.0342** | 0.011 |
| Financial Services | 0.0205** | 0.041 |
| Materials and Resources | 0.00929 | 0.729 |

*Notes:* This table reports the heterogeneous effects of common ownership on the financing behavior of the startups. The second column of each row reports the estimated average treatment effect of common ownership on whether a startup successfully gets new rounds of funding within 180 days of the treatment, $\hat{\tau}$, in that specific industry. Results are from the regressions of the outcome on the treatment controlling for control for the matched-pair fixed effects and the three-way interacted fixed effects among found year, location of the headquarters, and the deal year $t$. We use the 1-5 matching algorithm in these regressions. *** p<0.01, ** p<0.05, * p<0.1. The third column reports the p-values of each regression using subsamples in different industries.

We define the treatment event in the previous regressions as whenever early-round equity financing occurs. We now explore whether the treatment effects would be larger for "better" events by evaluating the size of each financing round and focusing only on the high-value deals. Ideally, we want to have some measure of the success for each financing round, e.g., the ratio between the deal size and its pre-valuation. Unfortunately, the coverage rate of pre-valuation is low in our data, so we are not able to observe the valuation information for most of the deals. Instead, we use the ratio between the deal size and the post-evaluation of the last financing round as a proxy. We compute the 75th percentiles of the ratio's distribution and construct the subsample by selecting events with a ratio larger than or equal to p75. We run the same set of regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it} \tag{2.2}$$

on each subsample, with all definitions of parameters the same as before. Compared with the long-run outcome measures, the short-term ones are more relevant to our setting, so we use the short-term financing behaviors as the outcome. Table 2.5 reports the treatment effects of common ownership on short-term financing behavior within the subgroup whose ratio is larger than or equal to the 75th percentile. Results of the same set of regression using the 50th percentile as the cutoff are reported in the appendix. The magnitude of the treatment effect is larger on the subsample compared with that on the full sample, indicating that better events generate larger spillover effects.

Next, we do similar exercises on the opposite side. We explore the treatment effects when a negative event happens to a startup in the pool. We define a negative event as the closure of a startup. Whenever a startup in the common ownership pool closes at time $t$, we match all other startups in the same pool to the nearest five startups outside the pool, but in the same industry code. The features and functional form used in the matching algorithm are the same as before, and we focus on the more relevant short-term financing behavior measure as the outcome of interest. We do the same set of regressions as in Equation (2.1) and summarize the results in Table 2.6. The treatment effects of bad events are negative but minimal in magnitude and statistically insignificant. These results suggest that the mechanism through which the common ownership influences the performance is not merely driven by reputation.

## 2.5 Conclusion

This paper studies the effect of common ownership on the performance of startups. We document the stylized facts that commonly owned startups have better long-term performance in both reduced-form cross-sectional regressions and survival analysis after controlling for their observed features and several fixed effects. The commonly owned startups are shown to be 0.59% more likely to go public and 3.78% more likely to be acquired. Further addressing the endogeneity problem, we utilize a matched-pair design to identify and estimate the causal effect of common ownership on both the long-term and short-term performance of the startups. The results show that once a startup is successfully funded, others in the same common ownership pool have a 0.215% higher propensity to go IPO, 0.735% higher propensity to be acquired, and 0.445% lower probability of closure. Their short-term financing behaviors are also positively affected, with a 1.31% higher probability of getting new funding rounds in 180 days, 2.40% higher in 365 days, and 3.37% higher in 730 days. These results in short-term financing behavior imply a positive effect of common ownership on the startup's capacity to raise new funding. We also document the heterogeneity in treatment effects of common ownership across different industries. Our results also shed light on the plausible mechanisms of effect. We find that better funding events generate larger spillover effects, indicating that the mechanism of the common ownership effect could be driven by the loosened financial constraints of the investors and sharing of resources among the same common ownership pool, and also that failed startups do not have the spillover effects their peers, suggesting that the reputation channel cannot be the only driving force of common ownership effect.

Table 2.5: Subgroup effects on short-term financing behavior

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | 0.0209*** | 0.0358*** | 0.0486*** |
| | (0.00163) | (0.00194) | (0.00200) |
| Constant | 0.175*** | 0.309*** | 0.444*** |
| | (0.000653) | (0.000777) | (0.000801) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 5 | 5 | 5 |
| Observations | 398,177 | 398,177 | 398,177 |
| R-squared | 0.096 | 0.131 | 0.196 |

*Notes:* This table reports the subgroup effects of common ownership on the short-term financing behavior when we use only the top 25% most valued financing rounds as the successful treatment events. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-5 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Effects on financing performance when a negative event occurs

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | -0.00123 | -0.000203 | -3.80e-05 |
| | (0.000913) | (0.00123) | (0.00153) |
| Constant | 0.0286*** | 0.0546*** | 0.0930*** |
| | (0.000362) | (0.000486) | (0.000606) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 5 | 5 | 5 |
| Observations | 241,615 | 241,615 | 241,615 |
| R-squared | 0.120 | 0.152 | 0.193 |

*Notes:* This table reports the effects of common ownership on the short-term financing behavior when a negative event, closure, occurs. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-5 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

# Chapter 3

# Signaling in Venture Debt and Capital

*Coauthored with Can Huang*

## 3.1 Introduction

Conventional wisdom regards equity as the pivotal financing vehicle for new firms, especially high-tech firms that invest vast money in research and development before generating positive cash flows. Under greatly asymmetric information and a high risk of failure, classic corporate finance theory suggests that debt is not a wise choice in startup financing. Under the context of the agent problem, classic models predict that fast-growing industries and firms with negative cash inflows and high risks should have low leverage (Jensen and Meckling, 1976; Leland and Pyle, 1977; De Meza and Webb, 1987). Winton and Yerramilli (2008) find that under high uncertainty of continuation with risky cash flow distribution, low liquidation, low probability of success, and high returns if successful, equity financing is optimal. Ueda (2004) argues that entrepreneurs seek venture capital financing when they have little collateral and when they require larger investment amounts.

However, in contrast to the prediction classic theory makes, debt financing in the early-round financing market, known as venture debt, is unexpectedly active, and has experienced steady growth in recent years. Roughly \$1 - \$5 billion of venture debt in total are offered to startups annually (Ibrahim, 2010), and are observed in 28-40% of startups with venture financing (Davis et al., 2018). Venture debt is usually offered to venture-backed startups in the technology and health care industry at all stages, mostly after series A or series B. It has a maturity of three to five years, usually starting with a 6-to-12-month interest-only period. Unlike traditional loans, venture debt does not require property or equipment as collateral, which adds a higher risk to venture debt. To compensate for the high risk, venture debt investors ask for warrants of around 8% of the deal size, typically less than 1% of the company's total equity. Unlike convertible bonds with features similar to equity, venture debt is non-convertible. While seemingly risky, venture debt requires interest rates around 2% plus the prime rate, and the rates may be higher when the deal size is larger.

The puzzling existence of venture debt has attracted scholarly attention, with a growing literature documenting and studying the use of venture debt. A pervasive opinion explains the demand side of the venture debt market, arguing that startups are in favor of venture debt since it reduces dilution from equity financing and helps them reach milestones before the next round of funding. Both models and empirical results illustrate that venture debt helps firms that face high dilution and low pre-money valuation to reduce dilution by extending the runway (Ibrahim, 2010; Davis et al., 2018).

However, the supply side of venture debt remains understudied. To the best of our knowledge, this paper is the first to answer the puzzle of why venture debt investors are willing to offer seemingly risky debt at a relatively low interest rate. Are these debts mispriced, or are these debts far less risky than presumed? According to statistics from Silicon Valley Bank, which is a primary investor in venture debt, the default rate of venture debts in 2019 was 0.3%, a number that is fabulously low compared with the failure rate of venture-backed startups, which is around 25% according to industry reports. Based on this fact, this paper states that venture debt, though unsecured by conventional collaterals, is far less risky than presumed. This paper will elucidate a signaling channel that contributes to diminishing the risk of venture debt, supported by empirical results.

The startup financing market features tremendous asymmetric information. When venture debt does not play a role in startup financing, venture capital investors have to exert great effort to eliminate asymmetric information and invest in promising firms. However, the entrance of venture debt provides venture capital investors with more choices. The return of venture debt depends on the likelihood of a firm's continuing operation and getting the next round of funding from venture capital, so venture debt investors screen startups that have outstanding performance and may well succeed in raising funds. Venture debt prefers startups that are invested in by well-known venture capital investors in previous rounds, or with a steady and fast growth rate. Therefore, the firms venture debt investors select will on average have a higher probability of success and better performance than those without venture debt. In the next financing round, venture capital investors will take venture debt as a positive signal, presuming that firms with venture debt are more likely to be good firms. To save on the high cost of due diligence, venture capital investors investigate these firms less and invest. In this way, venture debt increases the probability of a firm getting next-round funding, making the venture debt more likely to be paid off and hence lower the risk of venture debt itself. Anticipating this logic, venture debt investors have the incentive to do a relatively rough screening and issue debt to a small portion of low-quality firms. Though the quality of firms with venture debt is not ideal, the signaling effect still exists as long as the overall performance is sufficiently better than firms without venture debt. Venture capital investors will still invest in firms with venture debt without careful due diligence since they can cover the loss of investing in low-quality firms by saving due diligence costs when they are sufficiently high.

To illustrate this signaling, we establish a three-period simplified model, involving three parties, startups, venture capital investors (VCs), and banks that provide the venture debt. There are two types of startups, high-type, and low-type. While only startups know their type, banks can choose the level of screening cost to filter out some portion of the low types, and VCs can uncover the type after investigation at some cost. At time 0, startups can borrow venture debt from the bank, and the bank decides how much effort to take in screening and whether to lend the venture debt. At time 1, firms have their states realized. After observing the realization of firms at time 1 and banks' investment decisions, the VC updates the conditional probability of a startup firm being high-type, then chooses whether to investigate at some cost to reveal the type of the firm and whether to invest in it. After the VC's decision, the debt starts to get paid. If the VC invests in the firm with debt, then the debt is paid off at the end of time 1. Otherwise, the firm defaults, and banks receive nothing. We solve for equilibria and prove the existence of the signal effect. When the due diligence cost is sufficiently high, if the firms with debt are in a good state at time 1, it is optimal for the venture capital investors to take banks' screening into account and invest in firms with debt without due diligence, to save costs. However, in terms of investment performance, the signal effect results in over-investment in low-type firms.

The model provides four predictions that will be supported by empirical results. First, firms with venture debt have a shorter time gap between funding rounds. As suggested in the model, in equilibrium, venture capital investors do less careful due diligence when assessing a startup firm with venture debt, and this will shorten the period of investigation. As a result, it takes less time for startups with venture debt to get the next round of funding than those without debt. Second, firms with venture debt have better long-term performance. Because banks filter out a portion of low-type firms by their screening, the pool of startups with venture debt will have better performance on average. However, our third prediction shows that conditional on getting the next rounds of venture capital, the performance of firms with venture debt is worse compared with the counterfactual world without venture debt. As we have seen, taking debt financing as a signal, VCs prefer to do less careful due diligence to save costs when confronting firms with venture debt. Since banks do not prudently filter out all low-type firms, next-round venture capital investors will invest in both high-type and low-type entrepreneurs. However, if VCs see a startup without venture debt in the next round of funding, they *will* do sufficient due diligence and invest only in high-type firms, resulting in a much lower probability of investing in low-type firms. Therefore, conditional on getting next-round financing, the average long-term performance of firms with venture debt is expected to be worse than the average of those without venture debt. Finally, our model predicts that more severe asymmetric information reinforces the signaling effect. When faced with more severe asymmetric information, namely higher due diligence costs, VCs get more benefit from taking the signal of venture debt and thus are more willing to get a free ride on the banks' screening results, leading to stronger signaling effects when venture capital investors have less information.

In the empirical section, we test these four predictions. We find strong empirical evidence in line with the model predictions. We use CrunchBase data, which contains the funding round history of startups with information on the investors, investment types, series rounds, funding size, and the announcement dates. First, we find that it takes a startup with venture debt a significantly shorter time to get the next round of funding, which indicates that VCs indeed take venture debt as a good signal when making their investment decisions. Comparing the length of time between two funding rounds, it takes around 93 days less for startups with venture debt to get funds from a next-round investor. Second, in the long time horizon, we find that startups with venture debt have significantly better performance. To measure long-term performance, we investigate several variables of a startup: whether it has closed, gone public, been acquired, or is still operating. Startups with venture debt show better long-term performance evaluated using different measures. They have a marginally (at mean) 3.33% lower probability of suspending operation (at a significance level of 0.01), and are 0.45% more likely to achieve an IPO. However, conditional on getting the next round of funding, our model predictions are reserved. Such startups with venture debt have relatively worse average long-term performance, which is indicated by a marginally (at mean) 1.90% higher probability to close and 1.47% lower likelihood of going public. Again, this results from the VCs taking signals from banks' decisions without careful due diligence and investing in a pool containing both high-type and low-type firms. Lastly, empirical results show that the signaling effect of venture debt is stronger when the asymmetric information problem between the startups and investors is more severe. We test and find that if startups are funded by experienced investors in the next round, the length of time for startups with venture debt to get next round funding is 39 days less compared with if they are invested by inexperienced investors.

Several types of robustness checks are provided to test the sensitivity of our results to different industries and startups founded during different years. The primary empirical evidence remains the same when we do the same types of empirical analysis in various industries and when we use the subsample by choosing companies founded in different time slots. Also, when we test both the unconditional and conditional long-term performance of startups, the startups used in the conditional test are a subsample. We check the robustness of our first empirical prediction on the subsample, which is precisely the same as what we used for the conditional long-term performance test. The pattern that it takes significantly less time for startups with venture debt to raise next-round funding is quite robust.

It is worth noting that when banks make decisions on whether to lend venture debt to a startup, they prefer startups backed by influential venture capital investors. Therefore, one possible concern is that the signaling effect can be caused by influential VCs rather than venture debt. Based on the extraordinary investment performance of these VCs, it is wise for future investors to follow them when making investment decisions. The well-known VCs can also send similar signals. To distinguish the signal of venture capital and venture debt, we do a robustness check to see the impact of being backed by a good venture capital investor. The result shows that while ever invested by a well-performed VC indicates better long-term performance, it also shows better conditional long-term performance, which is substantially distinct from the signaling effect of venture debt.

We enrich the traditional theoretical literature on early-stage startup financing. Agency problems and bankruptcy costs make equity a more favorable financing tool for firms with high risk and low value of the collateral (Jensen and Meckling, 1976; Leland and Pyle, 1977; De Meza and Webb, 1987; Harris and Raviv, 1991). These models do not capture the critical features of venture debt—that it is a relatively short-term contract, and its risk depends on the success of future funding rounds instead of a firm's final success.

We also contribute to the growing literature on venture lending in the early stage of startups by studying the rise of venture debt from the perspective of demand. De Rassenfosse and Fischer (2016) and Davis et al. (2018) show theoretical and empirical evidence that venture debt enables startups to attain more milestones and prevent further dilution. Davis et al. (2018) also find that firms demanding venture debt face higher potential dilution and exhibit lower pre-money valuations. From the perspective of investors, Hochberg et al. (2018) focus on patent-backed venture debt and find that the credibility of VC commitments increases lending. De Rassenfosse and Fischer (2016) argue that being backed by a VC company increases the probability that a startup will obtain venture debt. Cumming and Fleming (2013) study the determinants of returns of venture lending, highlighting the role of time allocation for due diligence and monitoring. This paper builds on this literature in the way that we address the signal effect of having venture debt and explains the seemingly low return of venture debt, which is a puzzle that none of the above papers explains.

We also relate to the literature on the impact of debt financing on startups and innovation in any stage of growth. In addition to the above literature that shows that debt reduces dilution, Geelen et al. (2019) find that while debt hinders innovation due to debt overhang, it encourages entry, fostering growth at the aggregate level. Hombert and Matray (2016) study how relationship lending determines the financing of innovation. Albertus and Denes (2019) document the emergence of debt financing by private equity funds. They find that funds using debt financing tend to reduce the amount of equity invested relative to fund size and delay capital calls. Our research contributes to this literature by providing evidence that as a result of the signaling effect, venture debt induces overinvestment from venture capital investors in later stages.

A branch of related literature studies the signaling effect in venture capital investment, mainly focusing on the patent signaling effect. Howell (2017) show that an early-stage award from the Department of Energy's SBIR grant program significantly increases the probability that a firm receives venture capital. Conti et al. (2013) find that patents serve as a positive signal to attract investors. Our paper also shows the evidence that venture capital investors exploit signal effects when making investment decisions, but focuses on the signal effect of having debt on the balance sheets, which has not been studied yet.

The rest of this paper is organized as follows: Section 3.2 develops our model and the equilibrium. Section 3.3 introduces the database we use in this paper and the variables of interest. It also presents the empirical implications of our model, with the test and results presented in Section 3.4. Section 3.5 provides several further robustness tests, and Section 3.6 concludes.

## 3.2 Model

In this section, we develop a three-period model to illustrate the signaling effect of venture debt. In the model, asymmetric information exists, and both banks (the venture debt providers) and VCs make efforts to screen the firms. Banks move first to do the screening at a cost and provide venture debt to firms that pass the screening. In the second stage, VCs observe the mid-stage realization of firms' valuation and decide whether to do due diligence and invest in the firms. As the banks filter out some bad firms, the VCs may well get a free ride and do less due diligence. Anticipating the VCs' behavior, banks will make fewer efforts in first-stage screening. The following subsections will formally model this intuition.

### Model Setup

There are three periods $t \in \{0, 1, 2\}$ and three parties: startup firms, venture capital investors, and banks. Firms are either high type or low type with a proportion of $\alpha$ being a high type, which is private information to firms. In $t_0$, firms ask for venture debt from banks, and banks do costly investigations into these applicants. By choosing the level of screening cost $C_B$, banks can filter out part of the low-type firms thus increasing the proportion of high-type firms to $\alpha(C_B)$. After screening, banks provide debt of a fixed amount of $D$ to firms passing the screening. $\alpha(C_B)$ is a concave monotonically increasing function whose value range is between $\alpha$ and 1, with $\alpha(0) = \alpha$ and $\lim_{c \to \infty} \alpha(c) = 1$. At time $t_1$, firms' mid-stage value is realized. The value of firms is $V = a$ with probability $P_i$, where $i \in \{H, L\}$ is an indicator of firm type, and $V = b < a$ otherwise. After observing the realization and the bank's investment decisions, the venture capital investor takes action. They decide whether to do due diligence, and whether to invest a fixed amount of $I$ in the firms. Since VCs have more information sources and better knowledge in due diligence, unlike banks, they can reveal the type of the firms at a fixed cost of $C$. Afterward, startup firms that receive venture capital investment pay off the venture debt; otherwise, firms default and banks get 0. In $t_2$, firms' final valuations are realized, whose expectation in $t_1$ is $\mu_i$, where $i \in \{H, L\}$, and VC's return is realized.

Banks are rational and maximize their payoff $E(R) - C_B - D$ by

$$\max_{C_B} \Pi_B = -C_B - D + P_I(C_B)R,$$

where $P_I$ is the probability of debt being paid off, namely the probability of VC investing in the firms. Assuming a competitive venture debt market, banks receive payment of $R$ such that

$$\Pi_B = -C_B - D + P_I(C_B)R = 0.$$

The maximization problem is equivalent to solving the following equations.

$$\begin{aligned} &-1 + \frac{dP_I}{dC_B}R = 0, \quad (\text{ F.O.C. }) \\ &-C_B - D + P_I R = 0. \end{aligned}$$

Suppose banks spend a screening cost of $C_B^*$, then the proportion of good types in firms with venture debt becomes $\alpha^* = \alpha(C_B^*)$. Since the firms without venture debt are low-type firms that are filtered out by banks, VC will only consider firms with venture debt. Based on banks' decisions, venture capital investors will update their belief of the probability of a firm being a high type conditional on having venture debt. When observing $V = a$, the updated probability of being a high type becomes

$$\alpha_a^* = \alpha_a(c_B^*) = \frac{\alpha(c_B^*) P_H}{\alpha(C_B^*) P_H + (1 - \alpha(C_B^*)) P_L},$$

and VCs maximize payoff by

$$\max\left\{-cI + \alpha_a^*\left(\frac{aI}{I+a}\mu_H - I\right), \frac{aI}{I+a}\left(\alpha_a^*\mu_H + (1-\alpha_a^*)\mu_L\right) - I, 0\right\},$$

where $c = C/I$. VCs compare the payoff of three options: doing due diligence and investing in high-type firms, investing in all firms without due diligence, or doing nothing. Three arguments in the maximization problem correspond to these three options. Solving the problem, we get the venture capital investor to do the following actions.

1. When $c < c_a(C_B^*) = (1 - \alpha_a(C_B^*)) \frac{I+a-a\mu_L}{I+a}$ and $c < \alpha_a(C_B^*)\left(\frac{a\mu_H}{I+a} - 1\right)$, do due diligence and invest in high type firms.

2. When $c > c_a(C_B^*)$ and $\alpha_a^*\mu_H + (1-\alpha_a^*)\mu_L - 1 > \frac{I}{a}$, invest in all firms with venture debt.

3. Otherwise, do not invest.

Similarly, when observing $V = b$, the updated probability becomes

$$\alpha_b^* = \alpha_b(C_B^*) = \frac{\alpha_b(C_B^*)\,(1 - P_H)}{\alpha_b(C_B^*)\,(1 - P_A) + (1 - \alpha_b(C_B^*))\,(1 - P_L)},$$

and the venture capital investor maximize

$$\max\left\{-CI + \alpha_b^*\left(\frac{bI}{I+b}\mu_H - I\right), \frac{bI}{I+b}\left(\alpha_b^*\mu_H + (1-\alpha_b^*)\,\mu_L\right) - I, 0\right\}.$$

The actions VC takes are

1. When $c < c_b(C_B^*) = (1 - \alpha_b(C_B^*))\,\frac{I+b-b\mu_L}{I+b}$ and $c < \alpha_b(C_B^*)\left(\frac{b\mu_H}{I+b} - 1\right)$, do due diligence and invest in high-type firms.
2. When $c > c_b(C_B^*)$ and $\alpha_b^*\mu_H + (1-\alpha_b^*)\,\mu_L - 1 > \frac{I}{b}$, invest in all firms with venture debt.
3. Otherwise, do not invest.

Meanwhile, in a world without venture debt, which is equivalent to the circumstance that $C_B = 0$, we assume that VC will always do due diligence and invest in high-type firms. In order to restrict to this condition, $\mu_i$ and $c$ are assumed to satisfy the following conditions.

**Assumption 3.1.** *The following conditions are satisfied.*

$$\begin{aligned}
&\mu_L < \frac{I+a}{a},\\
&\mu_H > \frac{I+b}{b},\\
&c < \alpha_a(0),\\
&c < \alpha_b(0)\left(\frac{b\mu_H}{I+b} - 1\right),\\
&\alpha_a(0)\mu_H + (1-\alpha_a(0))\,\mu_L - 1 > \frac{I}{a},
\end{aligned}$$

*where*

$$\begin{aligned}
\alpha_a(x) &= \frac{\alpha(x)P_H}{\alpha(x)P_H + (1-\alpha(x))P_L},\\
\alpha_b(x) &= \frac{\alpha(x)\,(1-P_H)}{\alpha(x)\,(1-P_H) + (1-\alpha(x))\,(1-P_L)}, \quad \textit{and}\\
c_a(x) &= (1-\alpha_a(x))\,\frac{I+a-a\mu_L}{I+a}.
\end{aligned}$$

The first two assumptions ensure that there exists $c$ that satisfies the other two criteria. The third and fourth assumptions allow the investment in high-type firms after due diligence to be the best choice for VC when there is no venture debt in the market. The last assumption assumes that the payoff of investing in all firms is positive, which will simplify the discussion of equilibria in the following subsection.

## Equilibrium

According to the level of $c$, there are three possible equilibria in this model described formally below.

**Theorem 3.1** (Equilibrium)**.** *There are three possible equilibria in this model:*

1. *When $c < c_a(C_B^0)$, banks take efforts of $C_B^0$ and $R^0 = \frac{C_B^0+D}{\alpha(C_B^0)}$. In $t_1$, VCs always invest in high-type firms after due diligence.*
2. *When $c > c_a(C_B^2)$, banks take efforts of $C_B^2$ and $R^2 = \frac{C_B^2+D}{\alpha(C_B^2)+(1-\alpha(C_B^2))P_L}$. VCs invest in all firms with venture debt when $V = a$, and invest in high-type firms after due diligence when $V = b$.*
3. *When $c \in (c_a(C_B^0), c_a(C_B^2))$, banks take efforts of $C_B^1$ and $R^1 = \frac{C_B^1+D}{\alpha(C_B^1)+(1-\alpha(C_B^1))P_{L}p}$. When $V = a$, VCs play a mixed strategy of investing in all firms with venture debt with some probability of $p$ and investing in high-type firms after due diligence with a probability of $1-p$. When $V = b$, VC invests in high-type firms after due diligence.*

To simplify the discussion on $c$, we impose one assumption on $c$.

**Assumption 3.2.** *Assume that $c_a(0) < c_b(C_B^0)$.*

This assumption allows that when $V = b$, the investigation is a better choice for VC regardless of $c$.

Theorem 3.1 demonstrates the equilibria in this model. The actions of VC highly depend on the level of due diligence cost, as deciding between doing due diligence or not is a trade-off between avoiding bad investment and saving due diligence cost. When VCs observe a bad state, the probability of a firm being a low type is high enough that if the due diligence cost is sufficiently small, although banks partially filter out bad firms, the expected cost of investing in low-type firms still exceeds the cost of due diligence. Under this circumstance, VC will not blindly invest in all firms with venture debt but will still carefully investigate firms. Therefore, there is no signaling effect in this equilibrium. However, when $c$ is high enough to surpass the losses of investing in low-type firms, VCs will take a free ride on banks' screening results and invest in all firms with venture debt. In this case, the signaling of venture debt does take effect. When the due diligence is moderate, VC will partially take the signal by having a mixed strategy of doing due diligence or not. The signaling effect is mitigated but still exists under this condition.

In response to the behavior of VC, banks spend different levels of screening cost and set the payoff $R$ accordingly. The first equilibrium can be treated as the benchmark under which condition there is no signaling effect, and banks spend $C_B^0$ investigating firms and ask for $R^0$ for return. In the other equilibria, due to the existence of the signaling effect, banks set the cost to $C_B^1$ and $C_B^2$. The relation between costs in different situations can be described in the following proposition.

**Proposition 3.1.** *Under the above settings and assumptions,*

$$\begin{aligned} &C_B^0 > C_B^1 > C_B^2, \\ &R^0 > R^1 > R^2. \end{aligned}$$

This proposition indicates that the signaling effect causes banks to do less careful screening, as the probability of a venture debt being paid off is higher when VC regards having debt as a good signal. Therefore, as due diligence cost increases hence signaling effects increase, banks spend less effort on screening. Though lower efforts result in higher default risk, reduced screening cost covers the loss in less careful investment and overall lowers $R$. This proposition gives a possible explanation for the puzzle raised in the beginning that why seemingly highly risky venture debt asks for moderate interest rates.

Besides the nature of venture debt, the model also predicts the performance of firms. More related to the empirical test in the next section, another proposition is straightforward to show that the existence of venture debt providers increases the quality of the pool of firms via screening. Since banks filter out some low-type firms, the average performance of firms receiving venture debt will be better.

**Proposition 3.2.** *Under the model setup,*

$$\mathbb{P}(H|\,VD) = \alpha(c) > \alpha = \mathbb{P}(H)$$

However, as VCs take venture debt as a positive signal and skip the due diligence, the signaling effect can result in venture capital investors over-investing in low-type firms. In a counterfactual world without venture debt, VCs will always do careful due diligence and invest only in high-type firms. Nevertheless, with banks screening at the first stage, VC will save the due diligence cost at the expense of over-investing in low-type firms.

**Proposition 3.3** (Over-investment)**.** *Given the model setup, the introduction of venture debt results in less due diligence and over-investment of VCs in low-type firms, i.e.,*

$$\mathbb{P}^0(L|\mathit{Invest}) = 0,$$

$$\mathbb{P}^1(L|Invest) = \frac{pP_L(1-\alpha(C_B^1))}{pP_L(1-\alpha(C_B^2))+\alpha(C_B^1)} > 0,$$
$$\mathbb{P}^2(L|Invest) = \frac{P_L(1-\alpha(C_B^2))}{P_L(1-\alpha(C_B^2))+\alpha(C_B^2)} > \mathbb{P}^1(L|Invest),$$

*where $\mathbb{P}^i(L|Invest)$ is the conditional probability of a startup firm invested by the VCs being a low-type in equilibrium $i$.*

Proposition 3.3 indicates that when the signaling effect increases, the quality of the firms that VCs invest in decreases. As the venture capital investors rely more on the screening results from banks, the banks are more likely to get the debt paid off and thus do less careful screening, which increases the probability of VCs investing in low-type firms.

Based on the above propositions, we derive four predictions for empirical tests.

**Prediction 3.1.** *Firms with venture debt have a shorter time gap between funding rounds.*

As predicted in the model, in good states VCs will take venture debt as a positive signal and do less careful due diligence, accelerating the financing process.

**Prediction 3.2.** *Firms with venture debt have better long-term performance.*

Before banks issue venture debt, they do screening and filter out some low-type firms. Thus the average quality of firms that pass the screening and get debt will be higher than the total population.

**Prediction 3.3.** *Conditional on getting the venture capital, the performance of firms with venture debt is worse compared with the counterfactual world without venture debt.*

When there is no venture debt and signaling effect, VCs do due diligence, thus firms getting venture capital are high-type firms. However, when venture debt kicks in and VCs take the signaling, some low-type firms can also get funded, impairing the average performance of startup firms that get funding.

**Prediction 3.4.** *More asymmetric information reinforces the signaling effect.*

When the problem of asymmetric information is more severe, venture capital investors are more likely to get a free ride on banks' screening results and rely on the signal, resulting in a stronger signaling effect, thus stronger effects of venture debt on firms' performance.

## 3.3 Data and empirical method

### Model predictions

In our theoretical model, there are four main testable predictions.

First, venture debt acts as a good signal for a startup's next-round financing. Therefore, it's easier for startups with venture debt to get future rounds of VC financing. The signaling effect causes VCs to do less careful screening and overinvest in low-type firms. Empirically, overinvestment causes startups with venture debt to get the next round of funding faster.

Second, as shown in the equilibrium, when $c$ is sufficiently small, the probability of a firm with venture debt being a high type is higher than the proportion of high type in the population. There is an empirical prediction based on this theoretical result. Startups with venture debt are more likely to succeed since all high-type entrepreneurs are willing to use venture debt lending, while only some of the low types are willing to. Therefore, when we look at the long-term performance of all startups, it is not surprising to see those with venture debt are more likely to succeed.

Third, the model predicts that conditional on getting the next round of VC, the performance of startups with venture debt is worse compared to the counterfactual world without venture debt. From the model, in a world without venture debt, VCs will always do due diligence and invest in only high types after the screening. As venture debt currently acts as a good signal, VCs prefer not to do careful due diligence if they encounter a startup with venture debt, and will directly invest in it. As a result of overinvestment, more low-type startups can get VC financed because of the signaling effect of venture debt. Therefore, conditional on the startups achieving later rounds of VC investment, the pool of startups with venture debt performs worse in the long term.

Finally, the model suggests that the signaling effect of venture debt is stronger when the asymmetric information problem between the startups and investors is more severe. To test this prediction, we need suitable measures for the extent of both asymmetric information and the signaling effect. It is natural to think that experienced investors who have invested in a large number of startups or been involved in a large number of funding rounds are considered to have relatively more moderate asymmetric information problems, compared with investors who do not. Assuming this, we use whether an investor is experienced as a proxy of the severity with which it suffers from asymmetric information. As for the signaling effect, it is shown in prediction 1 that startups with venture debt are getting the next round of funding in a shorter period compared with those without venture debt. We utilize this shorter length of time as our measure of the strength of the signaling effect. We compare the number of days that startups with venture debt saved when the next round investor suffers from severe asymmetric information when it only has moderate asymmetry. The model predicts the length of time shortened between two rounds should be significantly less if the next round investor is an experienced investor.

We do our empirical analysis to test these four predictions and check the robustness of our results. We find that the empirical evidence is strong and in line with the model predictions. In the following subsections, we describe the dataset used for the empirical test, introduce our definitions of some important variables, and illustrate and talk in detail about our variables of interest. We show the final results in the next section.

## Data

The data we use in this paper is from CrunchBase, a platform that collects comprehensive information about startups, where companies and investors can get market information or fund data on the platform in exchange for reporting their own information. Data in CrunchBase reports each funding round with information on the investors, investment types, series rounds, funding size, and the announcement dates. For startups, CrunchBase has information on their founded year, the total number of funding rounds, the total amount of funding, and their current status (operating, closed, IPO, or acquired). The industry and the number of employees are also available with some missing values. We use these data to test our model predictions in the following subsections.

## Definition of venture debt

To use CrunchBase data to test our predictions, we need to empirically identify which funding round is venture debt in the data. Venture debt appears in a venture's early stage, usually before series B. Thus, we define funding round as *venture debt* if and only if its investment type is "debt financing" and it's an early round, while we define funding round as *early round* if and only if:

1. the announced date of this round is before the angel, seed, series A, or series B, or
2. the announced date of this round is no more than two years later than the angel, seed, and series A, or
3. this round is right after series A, or
4. this startup's total number of funding rounds is less than or equal to two rounds.

## Variables of interests

In this subsection, we define and explain our variables of interest used in the empirical tests. To test our four model predictions, we need to get the following variables:

- an indicator of whether a funding round is a venture debt round (as defined in the previous subsection),
- the speed of getting next round financing,
- whether a firm ever used venture debt as a financing method,
- measures of the long-term performance of the firms,
- measures of the severity of asymmetric information problem, and
- the strongness of the signaling effect.

First, at the funding round level, the variable `vdebt` indicates whether a funding round is a venture debt round or not, which is equal to 1 if a round has the investment type "debt financing" and is an early-round, and 0 otherwise. Next, `date_diff` is defined as the length of the time interval between a funding round and its next round, and `date_diff2` is similarly defined as the length of the time interval between its next round and the previous round. We use both variables to measure how quickly a venture gets its next-round financing. Then, at the firm level, we define `have_vdebt` as an indicator of whether a startup ever used venture debt, which is equal to 1 if it has used, and 0 otherwise.

For the measure of the long-term performance of startups, we are mainly interested in three outcome variables:

1. `closed` is an indicator of whether a startup has already closed. It is set to be 1 if the startup is closed, and 0 otherwise. If a startup is already closed, we treat this as bad long-term performance.
2. `ipo` stands for whether a venture goes public, and is equal to 1 if it ever managed to go public. This variable is naturally an indicator of long-term success.
3. `acq` indicates whether a startup is acquired, and is equal to 1 if it is acquired. We consider being acquired by other companies as an indicator of good long-term performance for now and will have more discussion on this in the later sections.

It is possible that some startups first went public and got acquired after the acquisition. As for those firms, we treat their values of `acq` as 0 and `ipo` as 1, because it shows enough evidence of good performance if it ever succeeds in going public, and the decisions to acquire a startup and a public firm are very different, we would like to get rid of the latter case in our empirical test.

Finally, to test our last prediction, we need measures of the severity of asymmetric information problems and the strongness of the signaling effect. We define `num_round` as the total number of rounds a VC has ever got involved in and calculate this value for all investors in our dataset. Then, we use whether a VC's value of `num_round` is higher than or equal to the 90th percentile of the whole population. If it is higher, we treat the VC as an experienced investor that suffers from a less severe asymmetric information problem. Similarly, `num_company` is defined as the total number of companies a VC ever invested in, also calculated and compared with its 90th percentile to get another measure of an experienced investor. To check the stability of the cutoff (90th percentile), we also use the 95th percentile to measure whether the VC is treated as an experienced investor or not. We will show our empirical results in all the cases, and it does not make much difference which one we use, as they all lead to similar results.

Table 3.1 reports summary statistics of our variables of interest. Part A summarizes the 45,350 observations at the funding round level. Part B summarizes the 21,444 observations at the startup level. Part C reports summary statistics of the same variables as Part B but on a subsample of startups whose total number of funding rounds is greater than or equal to 4. This is the subsample we use for the conditional test in the latter part.

## 3.4 Empirical results

### Test 1: Venture debt is a good signal for next round financing

In this subsection, we test whether venture debt acts as a good signal to get the next round of financing. If this is the case, it's easier for startups with venture debt to get funding in the next round; thus, it takes a shorter period for startups with venture debt to reach the next round of funding. Our empirical strategy is to regress the length of the time interval between the next funding round and this round on a dummy variable indicating whether this round is a venture debt round or not. Considering that all venture debts are in early rounds, to be comparable, our regression sample only contains early rounds. Our regression equation is

$$\texttt{date\_diff} = \alpha_t + \beta \cdot \texttt{vdebt} + \epsilon$$

where $\alpha_t$ is the fixed effects, `vdebt` is an indicator of whether the financing is a venture debt round, and `date_diff` is the length of the time interval between the next round of funding and this round of funding. In addition, we also do the same test using `date_diff2` as the outcome variable of interest, where `date_diff2` is defined as the length of the time interval between the next round of funding and the previous round of funding.

We report the regression results in Table 3.2. The results indicate that the coefficient of `vdebt` is significantly negative in all the settings, and is robust to whether the year-fixed effects are included or not. This verifies the first model prediction. On average, it takes a startup with venture debt about a hundred days shorter to get the next round of funding. VCs take venture debt as a good signal and do not screen it as carefully as those without

Table 3.1: Summary statistics for funding rounds and startups

**A. funding round level variables - early rounds**

| VARIABLE | N | mean | sd | max | min | p25 | p50 | p75 |
|---|---|---|---|---|---|---|---|---|
| vdebt | 45350 | 0.048 | 0.214 | 1 | 0 | 0 | 0 | 0 |
| date_diff | 45350 | 460.547 | 447.887 | 13283 | 0 | 185 | 362 | 593 |
| date_diff2 | 22763 | 790.461 | 537.580 | 6797 | 0 | 427 | 686 | 1018 |

**B. startup level variables - full sample**

| VARIABLE | N | mean | sd | max | min | p25 | p50 | p75 |
|---|---|---|---|---|---|---|---|---|
| have_vdebt | 21444 | 0.097 | 0.296 | 1 | 0 | 0 | 0 | 0 |
| closed | 21444 | 0.105 | 0.307 | 1 | 0 | 0 | 0 | 0 |
| ipo | 21444 | 0.024 | 0.153 | 1 | 0 | 0 | 0 | 0 |
| acq | 21444 | 0.190 | 0.393 | 1 | 0 | 0 | 0 | 0 |
| # of funding rounds | 21444 | 2.360 | 2.074 | 23 | 1 | 1 | 1 | 3 |

**C. startup level variables - subsample conditional on # funding rounds $\geq$ 4**

| VARIABLE | N | mean | sd | max | min | p25 | p50 | p75 |
|---|---|---|---|---|---|---|---|---|
| have_vdebt | 4365 | 0.179 | 0.383 | 1 | 0 | 0 | 0 | 0 |
| closed | 4365 | 0.042 | 0.201 | 1 | 0 | 0 | 0 | 0 |
| ipo | 4365 | 0.063 | 0.242 | 1 | 0 | 0 | 0 | 0 |
| acq | 4365 | 0.263 | 0.441 | 1 | 0 | 0 | 0 | 1 |
| # of funding rounds | 4365 | 5.759 | 2.156 | 23 | 4 | 4 | 5 | 7 |

*Notes:* This table summarizes the characteristics of the financing deals and startups in our data sample. Part A summarizes the three variables of interest at the funding round level. Part B and C summarize the five variables of interest at the startup level, where Part B describes the full sample, and Part C describes a subsample of startups whose total number of funding rounds is greater than or equal to 4.

venture debt. These results are robust to the two different definitions of the length of the time interval between financing rounds, `date_diff` and `date_diff2`.

Table 3.2: Test of signaling effect of venture debt

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | date_diff | date_diff | date_diff2 | date_diff2 |
| vdebt | -121.9*** | -93.05*** | -141.7*** | -118.6*** |
| | (8.207) | (8.033) | (12.98) | (12.88) |
| Constant | 466.4*** | 13,283 | 799.0*** | 607.0 |
| | (2.169) | | (3.699) | |
| year FE | | ✓ | | ✓ |
| Observations | 45,350 | 45,350 | 22,763 | 22,763 |

*Notes:* This table reports the effect of venture debt round on the length of time interval until the next round of funding. The treatment variable `vdebt` is an indicator of whether the funding is a venture debt round. The outcome variable of interest in the first two columns is `date_diff`, defined as the length of the time interval between the next round of funding and this round of funding. The outcome variable of interest in the last two columns is `date_diff2`, defined as the length of the time interval between the next round of funding and the previous round of funding. The regression models in Columns (1) and (3) are OLS regressions, while those in Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Our results are robust to outliers. When we winsorize the dependent variables, the results do not change much, and the coefficient remains significantly negative in all cases. However, we should interpret these results with caveats. Common sense in the literature and industry reports is that venture debt rounds usually come together with or right after venture capital rounds in the early phases of startup financing. Even so, it is possible that some startups seek venture debt when they are about to achieve a milestone and want to use debt financing as a way to avoid equity dilution. In that case, the power of our results as a proof of model prediction is reduced.

## Test 2: Long-term performance of startups with venture debt

We test the second prediction of the model in this subsection. The model predicts that in general, startups with venture debt are likely to have better long-term performance since all high-type entrepreneurs are willing to borrow venture debt while only some low types do so. In our empirical study, we would like to test whether startups with venture debt are more likely to succeed in the long term (e.g., IPO or acquisition), and less likely to get closed. Our empirical strategy is to use the Probit model to regress the measure of long-term performance on the dummy variable indicating whether a startup ever used venture debt or not. We use three measures of a firm's long term performance: `closed`, `ipo`, and `acq`, as defined in the previous section. Specifically, the Probit regression models we use are:

$$\begin{aligned}\texttt{outcome}' &= \alpha_t + \beta \cdot \texttt{have\_vdebt} + \epsilon,\\ \texttt{ipo}' &= \alpha_t + \beta \cdot \texttt{have\_vdebt} + \epsilon,\\ \texttt{acq}' &= \alpha_t + \beta \cdot \texttt{have\_vdebt} + \epsilon,\end{aligned}$$

where

- $\alpha_t$ are the year fixed effects;
- `have_vdebt` is an indicator equal to 1 if a startup ever used venture debt in the financing history, and 0 otherwise;
- `closed` is an indicator of whether the startup is closed;
- `ipo` is an indicator of whether the startup goes public in the end;
- `acq` is an indicator of whether the startup is acquired by an acquirer; and
- $Y' = \Phi^{-1}(Y)$ for $Y \in \{\texttt{closed}, \texttt{ipo}, \texttt{acq}\}$, where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution.

The sample we use here are startups founded between 2001 and 2011. CrunchBase uses a back-filling way to retrieve the data in the past. To avoid the measurement error caused by the fact that funding rounds information in the early years is not accurate enough, we decide not to use data on startups founded before 2001. On the other hand, many companies founded after 2011 are still operating now. Their long-term performances are yet to see and hard to predict, which is the reason we decide to exclude these startups from our test sample as well.

We report the regression results in Table 3.3. Startups with venture debt have a significantly lower probability of getting close, indicating their better long-term performance under the measure of closure, no matter whether the year fixed effect is considered or not. We also show results using IPO as a measure of success. Startups with venture debt also have a higher probability of going public, while the statistical power is limited. While going public seems to show the promise of young startups, recent literature documents that the number of public firms in the US has declined significantly recently, and one big reason some successful startups are shying away from IPOs is that public listings do not offer enough benefit to them Doidge et al. (2018). In our sample, we also see a vast number of startups that have been successfully operating for over a decade but never went public. Instead of not performing well enough to get into the public market, most of them do not seek going public as their ultimate goal and prefer to operate the business sustainably. Considering this fact, we construct a subsample of the firms excluding those with a low propensity to go public. We first predict each startup's intention of going public, then drop those with IPO probability less than the 10th percentile of the population. The prediction of the intention is based on their length of operation and the size of the company. This could be improved if a larger set of data on the properties of these companies is available. After excluding startups not willing to go public, the measure `ipo` is considered to be a better measure of success compared with not excluding them. The results of the same Probit regression on the subsample are reported in the last two columns of Table 3.3. Startups with venture debt have a significantly higher probability of going public than those without venture debt, which indicates the better long-term performance of firms that ever used venture debt.

The interesting results here are the ones using acquisition as dependent variables. Results indicate that startups with venture debt are significantly less likely to be acquired. Acquisitions have various purposes and are complicated in reality. Some promising startups may be purchased by some giant companies in the end, while others may feel not confident about their future and agree to sell the company at a low price. It will be more clear if we are able to distinguish between these two types of acquisition. We can use the premium of acquisition, defined as the ratio of the deal price over the book value of the company (Masulis and Nahata, 2011), as an indicator of whether an acquisition is a success for the startup or not. We then decompose the set of acquisitions into good and bad groups and use only the good acquisitions as a measure of success. This type of exercise is not doable due to the limit of our data and is left for future research.

Table 3.3: Effects of venture debt on long-term performance of startups

| | full sample | | | | | | subsample | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| VARIABLES | closed | closed | ipo | ipo | acq | acq | ipo | ipo |
| have_vdebt | -0.175*** | -0.185*** | 0.0821 | 0.0898 | -0.121*** | -0.116*** | 0.108* | 0.143** |
| | (0.0419) | (0.0420) | (0.0594) | (0.0611) | (0.0345) | (0.0355) | (0.0613) | (0.0642) |
| Constant | -1.238*** | -1.413*** | -1.988*** | -1.615*** | -0.866*** | -0.410*** | -1.947*** | -1.505*** |
| | (0.0120) | (0.0600) | (0.0197) | (0.0674) | (0.0103) | (0.0421) | (0.0203) | (0.0764) |
| year FE | | ✓ | | ✓ | | ✓ | | ✓ |
| Observations | 21,444 | 21,444 | 21,444 | 21,444 | 21,444 | 21,444 | 18,786 | 18,786 |

*Notes:* This table reports the effect of having venture debt on the long-term performance of startups. The treatment variable `have_vdebt` is an indicator of whether the startup ever used venture debt in the financing history. The outcome variables of interest stand for the exit status of startups, where `closed`, `ipo`, and `acq` are indicators of the startup going closed, public, and acquired, respectively. The first six columns report the regression results on the full sample. The last two columns are results of a subsample of likely-IPO startups that have a predicted propensity of going public larger than or equal to the 10th percentile. The odds number columns are Probit regressions with no fixed effects, while the even number columns control for the year fixed effects. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

### Test 3: Conditional long-term performance of startups with venture debt

In this subsection, we test the conditional long-term performance of startups with venture debt, which is the third model prediction. Our model predicts that conditional on getting the next round of venture capital, the performance of startups with venture debt is worse compared with the world without venture debt. Our empirical strategy is similar to that in test 2. We use the Probit regression of different measures of long-term performance on the dummy variable indicating whether a startup ever used venture debt. The sample is also restricted to startups founded between 2001 and 2011 for the same reason stated in the last test. However, to test the *conditional* long-term performance, we only use startups getting enough next round funding (measured as total funding round greater than or equal to 4 in this case).

The regression equation and definitions of all variables are the same as in Test 2. The only difference is imposing the condition that startups in this subsample already get their next round of funding. For those startups with venture debt, we are able to track whether they get the next round of funding. However, we need to construct a comparable subgroup for those startups that have never borrowed venture debt. To deal with this problem, we construct the subsample by filtering the total number of funding rounds of startups and keep only those with the total number of funding rounds greater than or equal to 4. As shown in Table 3.4, conditionally, startups with venture debt have a significantly higher probability of closing, no matter whether we control for year-fixed effects or not, indicating they have worse long-term performance. As for going public, conditionally, startups with venture debt have a lower probability of going public. Similar to what we do in Test 2, we do the test of IPO on the subsample with a large enough predicted propensity to go public. The results of regressions run on the subsample are reported in the last two columns of Table 3.4. The results using IPO as the dependent variable verifies our prediction that startups with venture debt have worse long-term performance. When we use `acq` as the measure of success, the result is inconsistent with the others, which is not surprising since the acquisition indicator has the same problem as discussed in the previous test. A better method to solve this problem is in need to come up with a better way of measure of success.

Table 3.4: Effects of venture debt on conditional long-term performance of startups

| | full sample | | | | | | subsample | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| VARIABLES | closed | closed | ipo | ipo | acq | acq | ipo | ipo |
| have_vdebt | 0.187** | 0.230*** | -0.176** | -0.142 | -0.151*** | -0.0872 | -0.163* | -0.0938 |
| | (0.0818) | (0.0832) | (0.0841) | (0.0876) | (0.0547) | (0.0565) | (0.0870) | (0.0926) |
| Constant | -1.762*** | -1.586*** | -1.508*** | -1.273*** | -0.607*** | -0.151* | -1.466*** | -1.231*** |
| | (0.0383) | (0.131) | (0.0323) | (0.111) | (0.0224) | (0.0813) | (0.0334) | (0.126) |
| year FE | | ✓ | | ✓ | | ✓ | | ✓ |
| Observations | 4,365 | 4,365 | 4,365 | 4,365 | 4,365 | 4,365 | 3,885 | 3,885 |

*Notes:* This table reports the effect of having venture debt on the conditional long-term performance of startups, conditioning on the startups that have already received at least four rounds of funding. The treatment variable `have_vdebt` is an indicator of whether the startup ever used venture debt in the financing history. The outcome variables of interest stand for the exit status of startups, where `closed`, `ipo`, and `acq` are indicators of the startup going closed, public, and acquired, respectively. The first six columns report the regression results on the full sample. The last two columns are results of a subsample of likely-IPO startups that have a predicted propensity of going public larger than or equal to the 10th percentile. The odds number columns are Probit regressions with no fixed effects, while the even number columns control for the year fixed effects. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

### Test 4: More asymmetric information reinforces the signaling effect

The last test provides empirical evidence for our fourth model prediction: the signaling effect of venture debt is stronger when the asymmetric information problem between the startups and investors is more severe. To test this, we construct measures of the extent of the signaling effect and measures of the asymmetric information problem severity as described in Section 3.3.

The regression equation we use for this test is:

$$\texttt{date\_diff} = \alpha_t + \beta_1 \cdot \texttt{vdebt} + \beta_2 \cdot \texttt{experienced} + \gamma \cdot (\texttt{vdebt} \times \texttt{experienced}) + \epsilon,$$

where

- $\alpha_t$ are the year fixed effects;
- `date_diff` is the length of the time interval between the next round of funding and this round of funding;
- `vdebt` indicates whether this is a venture-debt round; and
- `experienced` indicates whether the next round investor is an experienced investor, where an investor is defined as *experienced investor* if the total number of rounds they get involved is greater than the 90th percentile.

Table 3.5 reports the regression results. Consistent with test 1, the length of the time interval between the funding round and its next round is 100 days shorter when it is a venture-debt round. However, focusing on $\gamma$, the coefficient of the interaction term, we see if an experienced investor invests in the next round, the time shortened is significantly less, indicating the signaling effect of venture debt is a lot weaker among these experienced investors. Robustness checks in the next section use the 95th percentile and another variable—the total number of companies invested in—to define experienced investors. We show that our results are not sensitive to the threshold or the definition. All results provide strong empirical evidence of the fact that the signaling effect of venture debt tampers when the asymmetric information problem is more moderate.

## 3.5 Robustness testing

### Effect of good venture capital investors

As we briefly mentioned in the introduction, when the venture debt issuers make decisions on whether to lend money to a startup, in addition to their own screening process, whether the startup is backed by an influential VC also matters. Influential VCs not only invest a

Table 3.5: Intensity of signaling effects on severity of asymmetric information problem

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | date_diff | date_diff | date_diff2 | date_diff2 |
| vdebt × 1{#rounds ≥ p90} | 38.76** | 3.475 | 92.23*** | 56.48* |
| | (19.31) | (18.95) | (29.68) | (29.03) |
| vdebt | -144.0*** | -96.02*** | -191.8*** | -150.7*** |
| | (14.75) | (14.58) | (20.52) | (20.46) |
| 1{#rounds ≥ p90} | -3.315 | 7.204 | -21.67** | -24.31*** |
| | (5.328) | (5.213) | (9.006) | (8.760) |
| Constant | 473.0*** | 13,283 | 814.6*** | 631.3*** |
| | (3.531) | | (5.781) | (8.760) |
| year FE | | ✓ | | ✓ |
| Observations | 31,924 | 31,924 | 16,118 | 16,118 |

*Notes:* This table summarizes the test results on whether more asymmetric information reinforces the signaling effect. Regression equations are following

$$\texttt{date\_diff} = \alpha_t + \gamma \cdot (\texttt{vdebt} \times \texttt{experienced}) + \beta_1 \cdot \texttt{vdebt} + \beta_2 \cdot \texttt{experienced} + \epsilon,$$

where `vdebt` is an indicator of venture debt round and `experienced` is an indicator of experienced investor, defined as whether the total number of investment rounds $\geq$ 90th percentile. The outcome variable of interest in the first two columns is `date_diff`, defined as the length of the time interval between the next round of funding and this round of funding. The outcome variable of interest in the last two columns is `date_diff2`, defined as the length of the time interval between the next round of funding and the previous round of funding. The main parameter of interest is $\gamma$, the coefficient of the interaction term in the second row. The regression models in Columns (1) and (3) are OLS regressions, while those in Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

considerable amount of funds to support the research and development of the startups they invested in but also provide them with extraordinary management and ensure they are in good shape when they grow. Venture debt lenders take this into consideration and prefer startups backed by good VCs, and following the well-known VCs when making investment decisions is a seemingly secure rule for them. Therefore, it is likely that whether a startup is backed by renowned VCs is highly correlated with whether a startup has venture debt, and we need to distinguish the effect of being backed by influential VCs and the signaling effect of venture debt. To distinguish these two effects, we do our first robustness check to test the effect of good VCs. The empirical regressions are similar to what we do in tests 2 and 3. We regress the measure of success on whether a startup is backed by good VCs. For all results reported in this robustness testing section, we omitted the regression with `acq` as the left-hand-side variable due to the unclearness of whether the acquisitions are successful or not in our data. The regression equations are:

$$\begin{aligned} \texttt{closed}' &= \alpha_t + \beta \cdot \texttt{have\_good\_vc} + \epsilon, \\ \texttt{ipo}' &= \alpha_t + \beta \cdot \texttt{have\_good\_vc} + \epsilon, \end{aligned}$$

where

- $\alpha_t$ are the year fixed effects;
- `have_good_vc` is an indicator equal to 1 if a startup is backed by at least one good venture capital investor, and 0 otherwise;
- `closed` is an indicator of whether the startup was closed;
- `ipo` is an indicator of whether the startup went public in the end; and
- $Y' = \Phi^{-1}(Y)$ for $Y \in \{\texttt{closed}, \texttt{ipo}\}$, where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution.

We searched for top venture capital investors in the US and constructed the list referring to some ranking lists by some convincing and professional organizations. A detailed list of good venture capital investors is relegated to the appendix for the sake of conciseness. From results in Table 3.6, we observe that startups invested by good VCs have a lower probability of closing and a higher probability of going public than those not. This indicates that these good VCs significantly affect the long-term performance of the startups they invest in. Nevertheless, results on the conditional long-term performance in Table 3.7 indicate that even conditionally on getting enough funding rounds, startups backed by good VCs still have significantly better long-term performance. This helps us distinguish the effect of good VCs from the signaling effect of venture debt, where conditionally, the long-term performance of startups with venture debt is not as good as the world when there is no venture debt.

Table 3.6: Effects of good venture capital investors on long-term performance

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | closed | closed | ipo | ipo |
| have_good_vc | -0.198*** | -0.192** | 0.426*** | 0.462*** |
| | (0.0751) | (0.0755) | (0.0659) | (0.0681) |
| Constant | -1.470*** | -1.460*** | -1.812*** | -1.440*** |
| | (0.0202) | (0.0905) | (0.0254) | (0.0875) |
| year FE | | ✓ | | ✓ |
| Observations | 9,654 | 9,654 | 9,654 | 9,654 |

*Notes:* This table reports the effect of having good venture capital investors on the long-term performance of startups. The treatment variable `have_good_vc` is an indicator of whether the startup is ever invested in by at least one good VC. The outcome variables of interest stand for the exit status of startups, where `closed` and `ipo` are indicators of the startup going closed and public, respectively. Columns (1) and (3) are Probit regressions with no fixed effects, while Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Industry

Startups in different industries may have significantly different growth models, and preferences for financing and operating. To check the robustness of our predictions, we want to see whether our signaling model predicts all different industries well. We check the robustness by doing similar regressions among various industries and to see whether the results depend on the industry or not. We do analyses on both unconditional and conditional long-term performance regressions for the technology industry, including all tech firms such as software, hardware, health care, etc. The results are reported in Table 3.8 and Table 3.9. Compared with previous results, we see most of the results are robust. Ventures with venture debt are more likely to be closed unconditionally. Also, conditionally, they have worse long-term performance, indicated by a higher probability of closing and a lower probability of going public and getting acquired.

Table 3.7: Effects of good venture capital investors on conditional long-term performance

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | closed | closed | ipo | ipo |
| have_good_vc | -0.150 | -0.133 | 0.338*** | 0.390*** |
| | (0.107) | (0.108) | (0.0762) | (0.0795) |
| Constant | -1.710*** | -1.579*** | -1.586*** | -1.317*** |
| | (0.0364) | (0.134) | (0.0335) | (0.112) |
| year FE | | ✓ | | ✓ |
| Observations | 4,293 | 4,293 | 4,293 | 4,293 |

*Notes:* This table reports the effect of having good venture capital investors on the conditional long-term performance of startups, conditioning on the startups that have already received at least four rounds of funding. The treatment variable `have_good_vc` is an indicator of whether the startup is ever invested in by at least one good VC. The outcome variables of interest stand for the exit status of startups, where `closed` and `ipo` are indicators of the startup going closed and public, respectively. Columns (1) and (3) are Probit regressions with no fixed effects, while Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Time Period

In order to avoid the trouble that lots of companies founded after 2011 are still operating, thus it's hard to tell their long-term performances yet, we choose to use startups founded between 2001 and 2011 to test our model predictions. Here, to test the robustness of the results, we try to use all startups founded from 2001 to 2016 to see whether the results are very sensitive to the time window we choose. The long-term performance of startups is reported in Table 3.10, and their conditional long term performance is reported in Table 3.11. Here, we can see that the results for the rate of close and going public do not change much, while the results of acquisition are kind of ambiguous. Similar to what we discussed before, this may be caused by the different nature of acquisitions. Being acquired is not a perfect indicator of success for ventures. From these results, we can see that the empirical predictions are robust no matter what specific time slot we use for the test.

Table 3.8: Effects of venture debt on long-term performance in the technology industry

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | closed | closed | ipo | ipo |
| have_vdebt | -0.151*** | -0.162*** | 0.0363 | 0.0463 |
| | (0.0475) | (0.0475) | (0.0650) | (0.0665) |
| Constant | -1.285*** | -1.414*** | -1.921*** | -1.619*** |
| | (0.0142) | (0.0657) | (0.0215) | (0.0741) |
| year FE | | ✓ | | ✓ |
| Observations | 16,182 | 16,182 | 16,182 | 16,182 |

*Notes:* This table reports the effect of having venture debt on the long-term performance of startups in the technology industry. The treatment variable `have_vdebt` is an indicator of whether the startup ever used venture debt in the financing history. The outcome variables of interest stand for the exit status of startups, where `closed` and `ipo` are indicators of the startup going closed and public, respectively. Columns (1) and (3) are Probit regressions with no fixed effects, while Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Measures of experienced investors

To test the stability and robustness of our results on test 4, we use different measures of the severity of the asymmetric information problem and run the same set of regressions as in Section 3.4. We first use a different threshold, 95th percentile, when defining the experienced investors using the total number of investment rounds. Table 3.12 shows our results are robust to the threshold. Next, we define use another measure, the total number of companies a VC invests in, in the definition of experienced investors. Tables 3.13 and 3.14 report the results when the threshold is the 90th percentile an the 95th percentile, respectively. The results do not change much in scale and significance level.

Table 3.9: Effects of venture debt on conditional long-term performance in the technology industry

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | closed | closed | ipo | ipo |
| have_vdebt | 0.179** | 0.219** | -0.178** | -0.146 |
| | (0.0885) | (0.0894) | (0.0887) | (0.0918) |
| Constant | -1.772*** | -1.607*** | -1.485*** | -1.342*** |
| | (0.0417) | (0.140) | (0.0345) | (0.121) |
| year FE | | ✓ | | ✓ |
| Observations | 3,751 | 3,751 | 3,751 | 3,751 |

*Notes:* This table reports the effect of having venture debt on the conditional long-term performance of startups in the technology industry, conditioning on the startups that have already received at least four rounds of funding. The treatment variable `have_vdebt` is an indicator of whether the startup ever used venture debt in the financing history. The outcome variables of interest stand for the exit status of startups, where `closed` and `ipo` are indicators of the startup going closed and public, respectively. Columns (1) and (3) are Probit regressions with no fixed effects, while Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## 3.6 Conclusion

In this paper, we study the supply side of venture debt, providing a resolution to the puzzle that there is growing venture debt with relatively low rates of return while bearing high risk. We build a model where startups use venture debt as a good signal for their financing, and the cost of due diligence for the VC is sufficiently high that venture capital investors prefer to utilize this signal instead of investigating by themselves. We test the four predictions of the model in our empirical study. First, startups with venture debt can get next-round funding faster than those without venture debt, as venture debt is a good signal. Second, in general, startups with venture debt tend to perform better in the long term. They have a lower probability of going out of business and a higher probability of going public. Third, conditional on the startups getting their next round of funding, those with venture debt have worse long-run performance compared with the world without venture debt. Finally,

Table 3.10: Effects of venture debt on long-term performance of startups founded during 2001 - 2016

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | closed | closed | ipo | ipo |
| have_vdebt | -0.0602* | -0.156*** | 0.148*** | 0.0748 |
| | (0.0355) | (0.0365) | (0.0531) | (0.0559) |
| Constant | -1.457*** | -1.199*** | -2.193*** | -1.723*** |
| | (0.00936) | (0.0433) | (0.0163) | (0.0585) |
| year FE | | ✓ | | ✓ |
| Observations | 43,526 | 43,526 | 43,526 | 43,526 |

*Notes:* This table reports the effect of having venture debt on the long-term performance of startups founded during 2001 - 2016. The treatment variable `have_vdebt` is an indicator of whether the startup ever used venture debt in the financing history. The outcome variables of interest stand for the exit status of startups, where `closed` and `ipo` are indicators of the startup going closed and public, respectively. Columns (1) and (3) are Probit regressions with no fixed effects, while Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

the signaling effect is stronger when venture capital investors suffer from more severe asymmetric information problems. We show strong empirical evidence which is in line with our predictions, with various robustness tests provided. This paper documents the signaling effect of venture debt from both theoretical and empirical perspectives.

Table 3.11: Effects of venture debt on conditional long-term performance of startups founded during 2001 - 2016

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | closed | closed | ipo | ipo |
| have_vdebt | 0.241*** | 0.284*** | -0.153** | -0.139* |
| | (0.0701) | (0.0726) | (0.0770) | (0.0827) |
| Constant | -1.837*** | -1.451*** | -1.647*** | -1.279*** |
| | (0.0326) | (0.105) | (0.0285) | (0.0970) |
| year FE | | ✓ | | ✓ |
| Observations | 6,608 | 6,608 | 6,608 | 6,608 |

*Notes:* This table reports the effect of having venture debt on the conditional long-term performance of startups founded during 2001 - 2016, conditioning on the startups that have already received at least four rounds of funding. The treatment variable `have_vdebt` is an indicator of whether the startup ever used venture debt in the financing history. The outcome variables of interest stand for the exit status of startups, where `closed` and `ipo` are indicators of the startup going closed and public, respectively. Columns (1) and (3) are Probit regressions with no fixed effects, while Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.12: Intensity of signaling effects on severity of asymmetric information problem (experienced investors defined as total # of investment rounds $\geq$ p95)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | date_diff | date_diff | date_diff2 | date_diff2 |
| vdebt $\times$ 1{#rounds $\geq$ p95} | 40.48** | 9.348 | 94.83*** | 62.05** |
| | (19.18) | (18.81) | (29.74) | (29.07) |
| vdebt | -145.4*** | -100.1*** | -191.8*** | -152.6*** |
| | (14.27) | (14.18) | (20.08) | (20.07) |
| 1{#rounds $\geq$ p95} | 6.532 | 16.29*** | -16.63* | -18.63** |
| | (5.405) | (5.289) | (9.160) | (8.909) |
| Constant | 468.7*** | 13,283*** | 811.8*** | 625.6*** |
| | (3.386) | (0.000734) | (5.586) | (8.909) |
| year FE | | ✓ | | ✓ |
| Observations | 31,924 | 31,924 | 16,118 | 16,118 |

*Notes:* This table summarizes the test results on whether more asymmetric information reinforces the signaling effect. Regression equations are following

$$\texttt{date\_diff} = \alpha_t + \gamma \cdot (\texttt{vdebt} \times \texttt{experienced}) + \beta_1 \cdot \texttt{vdebt} + \beta_2 \cdot \texttt{experienced} + \epsilon,$$

where `vdebt` is an indicator of venture debt round and `experienced` is an indicator of experienced investor, defined as whether the total number of investment rounds $\geq$ 95th percentile. The outcome variable of interest in the first two columns is `date_diff`, defined as the length of the time interval between the next round of funding and this round of funding. The outcome variable of interest in the last two columns is `date_diff2`, defined as the length of the time interval between the next round of funding and the previous round of funding. The main parameter of interest is $\gamma$, the coefficient of the interaction term in the second row. The regression models in Columns (1) and (3) are OLS regressions, while those in Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.13: Intensity of signaling effects on severity of asymmetric information problem (experienced investors defined as total # of companies invested in ≥ p90)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | date_diff | date_diff | date_diff2 | date_diff2 |
| vdebt × 1{#companies ≥ p90} | 41.34** | 6.515 | 93.78*** | 58.83** |
| | (19.33) | (18.98) | (29.67) | (29.03) |
| vdebt | -145.4*** | -97.74*** | -192.7*** | -152.0*** |
| | (14.80) | (14.68) | (20.57) | (20.54) |
| 1{#companies ≥ p90} | -4.620 | 7.056 | -23.50*** | -24.71*** |
| | (5.328) | (5.215) | (8.997) | (8.753) |
| Constant | 473.6*** | 13,283 | 815.4*** | 631.7*** |
| | (3.531) | | (5.799) | (8.753) |
| year FE | | ✓ | | ✓ |
| Observations | 31,924 | 31,924 | 16,118 | 16,118 |

*Notes:* This table summarizes the test results on whether more asymmetric information reinforces the signaling effect. Regression equations are following

$$\texttt{date\_diff} = \alpha_t + \gamma \cdot (\texttt{vdebt} \times \texttt{experienced}) + \beta_1 \cdot \texttt{vdebt} + \beta_2 \cdot \texttt{experienced} + \epsilon,$$

where `vdebt` is an indicator of venture debt round and `experienced` is an indicator of experienced investor, defined as whether the total number of companies invested in ≥ 90th percentile. The outcome variable of interest in the first two columns is `date_diff`, defined as the length of the time interval between the next round of funding and this round of funding. The outcome variable of interest in the last two columns is `date_diff2`, defined as the length of the time interval between the next round of funding and the previous round of funding. The main parameter of interest is $\gamma$, the coefficient of the interaction term in the second row. The regression models in Columns (1) and (3) are OLS regressions, while those in Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.14: Intensity of signaling effects on severity of asymmetric information problem (experienced investors defined as total # of companies invested in $\geq$ p95)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | date_diff | date_diff | date_diff2 | date_diff2 |
| vdebt $\times$ 1{#companies $\geq$ p95} | 39.44** | 6.049 | 95.01*** | 59.81** |
| | (19.22) | (18.86) | (29.75) | (29.07) |
| vdebt | -144.6*** | -98.16*** | -191.6*** | -151.4*** |
| | (14.39) | (14.31) | (20.11) | (20.06) |
| 1{#companies $\geq$ p95} | 2.918 | 14.38*** | -21.09** | -22.01** |
| | (5.402) | (5.288) | (9.164) | (8.905) |
| Constant | 470.2*** | 13,283 | 813.6*** | 629.0*** |
| | (3.390) | | (5.580) | (8.905) |
| year FE | | ✓ | | ✓ |
| Observations | 31,924 | 31,924 | 16,118 | 16,118 |

*Notes:* This table summarizes the test results on whether more asymmetric information reinforces the signaling effect. Regression equations are following

$$\texttt{date\_diff} = \alpha_t + \gamma \cdot (\texttt{vdebt} \times \texttt{experienced}) + \beta_1 \cdot \texttt{vdebt} + \beta_2 \cdot \texttt{experienced} + \epsilon,$$

where `vdebt` is an indicator of venture debt round and `experienced` is an indicator of experienced investor, defined as whether the total number of companies invested in $\geq$ 95th percentile. The outcome variable of interest in the first two columns is `date_diff`, defined as the length of the time interval between the next round of funding and this round of funding. The outcome variable of interest in the last two columns is `date_diff2`, defined as the length of the time interval between the next round of funding and the previous round of funding. The main parameter of interest is $\gamma$, the coefficient of the interaction term in the second row. The regression models in Columns (1) and (3) are OLS regressions, while those in Columns (2) and (4) control for the year fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

# Bibliography

Albertus, J. F. and M. Denes (2019). “Distorting private equity performance: The rise of fund debt”. *Available at SSRN 3410076*.

Azar, J., M. C. Schmalz, and I. Tecu (2018). “Anticompetitive effects of common ownership”. *The Journal of Finance* 73.4, pp. 1513–1565.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). “Who’s who in networks. Wanted: The key player”. *Econometrica* 74.5, pp. 1403–1417.

Baum, J. A. and B. S. Silverman (2004). “Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups”. *Journal of business venturing* 19.3, pp. 411–436.

Brander, J. A., R. Amit, and W. Antweiler (2002). “Venture-capital syndication: Improved venture selection vs. the value-added hypothesis”. *Journal of Economics & Management Strategy* 11.3, pp. 423–452.

Bürer, M. J. and R. Wüstenhagen (2008). “Cleantech venture investors and energy policy risk: an exploratory analysis of regulatory risk management strategies”. *Sustainable innovation and entrepreneurship*, pp. 290–309.

Callaway, B. and P. H. Sant’Anna (2021). “Difference-in-differences with multiple time periods”. *Journal of Econometrics* 225.2, pp. 200–230.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). “Peer effects and social networks in education”. *The review of economic studies* 76.4, pp. 1239–1267.

Castilla, E. J. (2003). “Networks of venture capital firms in Silicon Valley”. *International Journal of Technology Management* 25.1-2, pp. 113–135.

Conti, A., J. Thursby, and M. Thursby (2013). “Patents as signals for startup financing”. *The Journal of Industrial Economics* 61.3, pp. 592–622.

Cumming, D. and G. Fleming (2013). “Debt investments in private firms: Legal institutions and investment performance in 25 countries”. *The Journal of Fixed Income* 23.1, pp. 102–123.

Davis, G. F. (2013). “After the corporation”. *Politics & Society* 41.2, pp. 283–308.

Davis, J., A. Morse, and X. Wang (2018). “The Leveraging of Silicon Valley”.

De Chaisemartin, C. and X. d’Haultfoeuille (2020). “Two-way fixed effects estimators with heterogeneous treatment effects”. *American Economic Review* 110.9, pp. 2964–96.

De Meza, D. and D. C. Webb (1987). "Too much investment: a problem of asymmetric information". *The quarterly journal of economics* 102.2, pp. 281–292.

De Rassenfosse, G. and T. Fischer (2016). "Venture debt financing: Determinants of the lending decision". *Strategic Entrepreneurship Journal* 10.3, pp. 235–256.

Del Bello, C. L., E. Patacchini, and Y. Zenou (2015). "Neighborhood effects in education". *Available at SSRN 2589818*.

Doidge, C. et al. (2018). "Eclipse of the public corporation or eclipse of the public markets?" *Journal of Applied Corporate Finance* 30.1, pp. 8–16.

Eldar, O., J. Grennan, and K. Waldock (2020). "Common ownership and startup growth". *Duke Law School Public Law & Legal Theory Series* 2019-42.

Fichtner, J., E. M. Heemskerk, and J. Garcia-Bernardo (2016). "Hidden power of the Big Three". *Passive index funds, re-concentration of corporate ownership, and new financial risk* 19.

Fine, J. P. and R. J. Gray (1999). "A proportional hazards model for the subdistribution of a competing risk". *Journal of the American statistical association* 94.446, pp. 496–509.

Geelen, T., J. Hajda, and E. Morellec (2019). "Debt, Innovation, and Growth".

Gelfand, A. E. and A. F. Smith (1990). "Sampling-based approaches to calculating marginal densities". *Journal of the American statistical association* 85.410, pp. 398–409.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). "Social networks and the identification of peer effects". *Journal of Business & Economic Statistics* 31.3, pp. 253–264.

Gompers, P. and J. Lerner (2001). "The venture capital revolution". *Journal of economic perspectives* 15.2, pp. 145–168.

Gompers, P., J. Lerner, and D. Scharfstein (2005). "Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999". *The journal of Finance* 60.2, pp. 577–614.

Gompers, P. A. and J. Lerner (1997). "Venture capital and the creation of public companies: do venture capitalists really bring more than money?" *The Journal of Private Equity*, pp. 15–32.

Gompers, P. A. et al. (2020). "How do venture capitalists make decisions?" *Journal of Financial Economics* 135.1, pp. 169–190.

Gompers, P. A. and J. Lerner (2004). *The venture capital cycle*. MIT press.

González-Uribe, J. (2020). "Exchanges of innovation resources inside venture capital portfolios". *Journal of Financial Economics* 135.1, pp. 144–168.

Goodman-Bacon, A. (2021). "Difference-in-differences with variation in treatment timing". *Journal of Econometrics* 225.2, pp. 254–277.

Harris, M. and A. Raviv (1991). "The theory of capital structure". *the Journal of Finance* 46.1, pp. 297–355.

Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2007). "Whom you know matters: Venture capital networks and investment performance". *The Journal of Finance* 62.1, pp. 251–301.

— (2010). "Networking as a barrier to entry and the competitive supply of venture capital". *The Journal of Finance* 65.3, pp. 829–859.

Hochberg, Y. V., C. J. Serrano, and R. H. Ziedonis (2018). “Patent collateral, investor commitment, and the market for venture lending”. *Journal of Financial Economics* 130.1, pp. 74–94.

Hombert, J. and A. Matray (2016). “The real effects of lending relationships on innovative firms and inventor mobility”. *The Review of Financial Studies* 30.7, pp. 2413–2445.

Howell, S. T. (2017). “Financing innovation: Evidence from R&D grants”. *American economic review* 107.4, pp. 1136–64.

Hsieh, C.-S. and L. F. Lee (2016). “A social interactions model with endogenous friendship formation and selectivity”. *Journal of Applied Econometrics* 31.2, pp. 301–319.

Ibrahim, D. M. (2010). “Debt as venture capital”. *U. Ill. L. Rev.*, p. 1169.

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jensen, M. C. and W. H. Meckling (1976). “Theory of the firm: Managerial behavior, agency costs and ownership structure”. *Journal of financial economics* 3.4, pp. 305–360.

Kaplan, S. N. and J. Lerner (2010). “It ain’t broke: The past, present, and future of venture capital”. *Journal of Applied Corporate Finance* 22.2, pp. 36–47.

Kaplan, S. N. and P. Stromberg (2001). “Venture capitals as principals: Contracting, screening, and monitoring”. *American Economic Review* 91.2, pp. 426–430.

Kelejian, H. H. and I. R. Prucha (2010). “Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances”. *Journal of econometrics* 157.1, pp. 53–67.

Leland, H. E. and D. H. Pyle (1977). “Informational asymmetries, financial structure, and financial intermediation”. *The journal of Finance* 32.2, pp. 371–387.

Lerner, J. and R. Nanda (2020). “Venture capital’s role in financing innovation: What we know and how much we still need to learn”. *Journal of Economic Perspectives* 34.3, pp. 237–61.

Lerner, J. (1994). “The syndication of venture capital investments”. *Financial management*, pp. 16–27.

Li, X., T. Liu, and L. A. Taylor (2021). “Common ownership and innovation efficiency”. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Lindsey, L. (2008). “Blurring firm boundaries: The role of venture capital in strategic alliances”. *The Journal of Finance* 63.3, pp. 1137–1168.

Manski, C. F. (1993). “Identification of endogenous social effects: The reflection problem”. *The review of economic studies* 60.3, pp. 531–542.

Masulis, R. W. and R. Nahata (2011). “Venture capital conflicts of interest: Evidence from acquisitions of venture-backed firms”. *Journal of Financial and Quantitative Analysis* 46.2, pp. 395–430.

Neyman, J. (1923). “On the application of probability theory to agricultural experiments. Essay on principles (with discussion). Section 9 (translated). reprinted ed.” *Statistical Science* 5, pp. 465–472.

Ogburn, E. L. (2018). “Challenges to estimating contagion effects from observational data”. In: *Complex Spreading Phenomena in Social Systems*. Springer, pp. 47–64.

Rosenbaum, P. R. and D. B. Rubin (1983). "The central role of the propensity score in observational studies for causal effects". *Biometrika* 70.1, pp. 41–55.

Rubin, D. B. (1974). "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of Educational Psychology* 66, pp. 688–701.

Rubin, D. B. (1973). "Matching to remove bias in observational studies". *Biometrics*, pp. 159–183.

Sah, R. K. and J. E. Stiglitz (1986). "The Architecture of Economic Systems: Hierarchies and Polyarchies". *The American Economic Review* 76.4, pp. 716–727.

Smith, T. E. and J. P. LeSage (2004). "A Bayesian probit model with spatial dependencies". In: *Spatial and spatiotemporal econometrics*. Emerald Group Publishing Limited.

Sun, L. and S. Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". *Journal of Econometrics* 225.2, pp. 175–199.

Ueda, M. (2004). "Banks versus venture capital: Project evaluation, screening, and expropriation". *The Journal of Finance* 59.2, pp. 601–621.

VanderWeele, T. J. and W. An (2013). "Social networks and causal inference". *Handbook of causal analysis for social research*, pp. 353–374.

Winton, A. and V. Yerramilli (2008). "Entrepreneurial finance: Banks versus venture capital". *Journal of Financial Economics* 88.1, pp. 51–79.

Zhao, A. and P. Ding (2021). "To adjust or not to adjust? Estimating the average treatment effect in randomized experiments with missing covariates". *arXiv preprint arXiv:2108.00152*.

# Appendix A

# Appendix to Chapter 1

## A.1 Proofs

### Lemma

**Lemma 1.** *$R_{syn}/R_{sta}$, the expected return of proposing and achieving syndicated deals over that of investing alone for the VC, is an increasing function of the VC's ability $A_{vc}$.*

*Proof of Lemma 1.* By previous calculation, we have

$$R_{sta} = \frac{R_h \alpha}{\alpha + (1-\alpha)\, p \exp(-A_{vc})} \qquad \text{and}$$
$$R_{syn} = \exp\{-(1-A_{vc})\}\,(R_h+v)\left\{\log\left(1+\frac{\alpha\exp(A_{vc}+1)}{(1-\alpha)\,p^2}\right) - \log\left(1+\frac{\alpha\exp(A_{vc})}{(1-\alpha)\,p^2}\right)\right\}.$$

Therefore, the ratio

$$\begin{aligned}
\frac{R_{syn}}{R_{sta}} &= \exp\{-(1-A_{vc})\}\,(R_h+v)\left\{\log\left(1+\frac{\alpha\exp(A_{vc}+1)}{(1-\alpha)\,p^2}\right) - \log\left(1+\frac{\alpha\exp(A_{vc})}{(1-\alpha)\,p^2}\right)\right\} \\
&\quad \times \frac{\alpha + (1-\alpha)\, p \exp(-A_{vc})}{R_h \alpha} \\
&= \frac{R_h+v}{R_h\alpha}\exp\{-(1-A_{vc})\}\log\left\{\frac{(1-\alpha)\,p^2+\alpha\exp(A_{vc}+1)}{(1-\alpha)\,p^2+\alpha\exp(A_{vc})}\right\}\{\alpha+(1-\alpha)\,p\exp(-A_{vc})\} \\
&= \frac{R_h+v}{R_h\alpha}\left(\alpha\exp\{-(1-A_{vc})\}+(1-\alpha)\,pe^{-1}\right)\log\left\{\frac{(1-\alpha)\,p^2+\alpha\exp(A_{vc}+1)}{(1-\alpha)\,p^2+\alpha\exp(A_{vc})}\right\}.
\end{aligned}$$

The first order derivative of this ratio with respect to $A_{vc}$ is thus

$$\frac{\partial}{\partial A_{vc}}\left(\frac{R_{syn}}{R_{sta}}\right) = \frac{R_h+v}{R_h\alpha}\left\{\alpha\exp(A_{vc}-1)+(1-\alpha)\,pe^{-1}\right\}$$

$$
\begin{aligned}
&\times \left\{ \frac{\alpha \exp\left(A_{vc}+1\right)}{\left(1-\alpha\right)p^2+\alpha\exp\left(A_{vc}+1\right)} - \frac{\alpha\exp(A_{vc})}{\left(1-\alpha\right)p^2+\alpha\exp(A_{vc})} \right\} \\
&+ \frac{R_h+v}{R_h\alpha}\log\left\{ \frac{\left(1-\alpha\right)p^2+\alpha\exp\left(A_{vc}+1\right)}{\left(1-\alpha\right)p^2+\alpha\exp(A_{vc})} \right\} \alpha\exp\left(A_{vc}-1\right) \\
&= \frac{R_h+v}{R_h} \underbrace{\left\{ \frac{1}{\left(1-\alpha\right)p^2\exp\left\{-\left(A_{vc}+1\right)\right\}+\alpha} - \frac{1}{\left(1-\alpha\right)p^2\exp(-A_{vc})+\alpha} \right\}}_{\mathcal{T}_1} \\
&\times \left\{\alpha\exp\left(A_{vc}-1\right)+\left(1-\alpha\right)pe^{-1}\right\} \\
&+ \frac{R_h+v}{R_h} \underbrace{\log\left\{ \frac{\left(1-\alpha\right)p^2+\alpha\exp\left(A_{vc}+1\right)}{\left(1-\alpha\right)p^2+\alpha\exp(A_{vc})} \right\}}_{\mathcal{T}_2} \exp\left(A_{vc}-1\right) \\
&> 0,
\end{aligned}
$$

where the first equality is by Leibniz product rule, the second equality is by rearranging terms, and the final inequality is by the facts that

- $\left(R_h+v\right)/R_h > 0$ by the setup,
- $\mathcal{T}_1 > 0$ because $\frac{1}{(1-\alpha)p^2\exp(-x)+\alpha}$ is an increasing function in $x$ and $A_{vc}+1 > A_{vc}$,
- $\mathcal{T}_2 > 0$ because the ratio inside the logarithm is larger than 1, and
- $\alpha\exp\left(A_{vc}-1\right)+\left(1-\alpha\right)pe^{-1}$ and $\exp\left(A_{vc}-1\right)$ are obviously non-negative values.

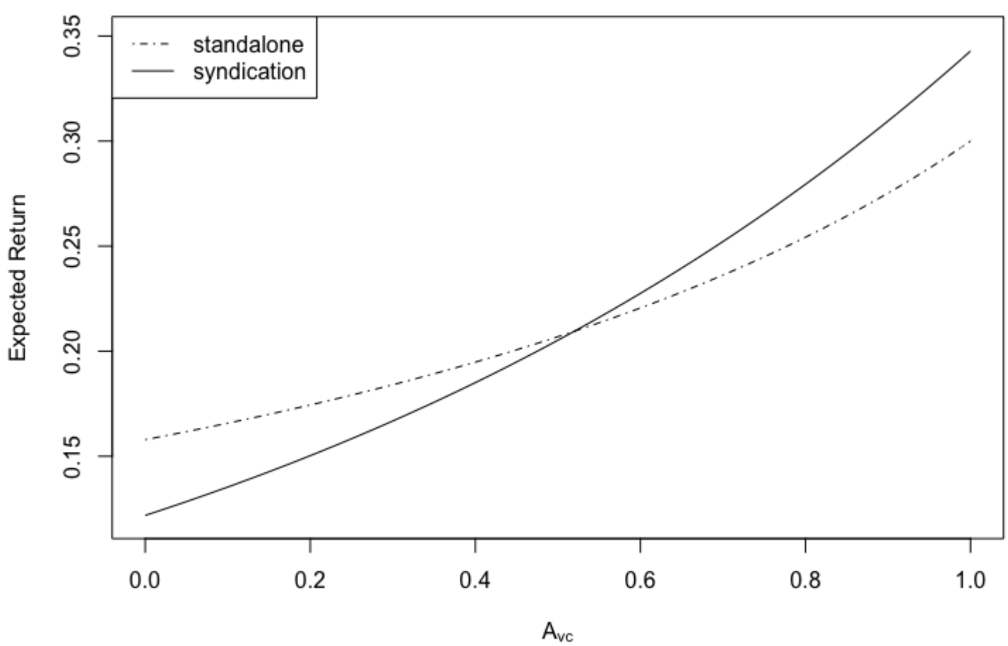
□

## Proof of Proposition 1.1

*Proof.* As calculated in Equations (1.8) and (1.9), the expected return of investing alone is $R_{sta}$, and that of proposing and achieving a syndicated deal is $R_{syn}$. To decide which option is better, the VC should compare these two expected returns depending on their own ability $A_{vc}$. Therefore, to get the optimal decision for the VC, it is crucial to solve for the inequality $R_{sta} \geq R_{syn}$ as a function of $A_{vc}$. It is obvious that both returns $R_{sta}$ and $R_{syn}$ are both increasing functions of $A_{vc}$. By Lemma 1, $R_{syn}/R_{sta}$ is increasing in $A_{vc}$. Use $R_{sta}(a)$ to denote the values of expected return of a standalone deal when the VC's ability is $a$, i.e., when $A_{vc} = a$. Similar notation $R_{syn}(a)$ to denote the expected return of syndicated deal as a function of $A_{vc}$. If $R_{sta}(0) \leq R_{syn}(0)$, then all VCs will choose the syndication option, while if $R_{sta}(1) \geq R_{syn}(1)$, then standalone deal is better for all VCs. Since we observe both standalone and syndicated investments in our dataset, we exclude such cases by assuming $R_{sta}(0) > R_{syn}(0)$ and $R_{sta}(1) < R_{syn}(1)$. Under this mild assumption, since $R_{syn}/R_{sta}$ is increasing in $A_{vc}$, we prove the existence and uniqueness of the solution to the equation $R_{sta} = R_{syn}$, and use $A^* \in (0, 1)$ to denote the unique solution. Figure A.1 illustrates the trends of two expected returns as functions of $A_{vc}$.

Therefore, when $A_{vc} \in [0, A^*)$, $R_{sta} > R_{syn}$ and the VC will choose standalone deals, and when $A_{vc} \in (A^*, 1]$, $R_{sta} < R_{syn}$ and syndication will be preferred. For the VCs with $A_{vc} = A^*$, the two options are indifferent and we assume they choose one of them randomly.

Figure A.1: Expected returns as functions of VC's ability




*Notes:* This figure illustrates the trends of the two expected returns - $R_{sta}$ of the standalone deals and $R_{syn}$ of the syndicated deals - as functions of $A_{vc}$. The solid line plots the expected return of syndicated deals, and the dashed line plots that of standalone deals. Parameters used in this plot are: $R_h = 0.3$, $v = 0.05$, $\alpha = 0.1$, and $p = 0.1$.

□

## Proof of Corollary 1.1

*Proof.* The success probability of a standalone deal given the VC's ability $A_{vc}$ is

$$\mathbb{P}\{A_e = h \mid S_0 = h, A_{vc}, \text{sta}\} = \frac{\alpha}{\alpha + (1-\alpha)\, p \exp(-A_{vc})}.$$

To get the overall success probability of standalone deals, we need to integrate over the distribution of $A_{vc}$ conditional on the standalone option being chosen, i.e.,

$$\begin{aligned}\mathbb{P}\{A_e = h \mid S_0 = h, \text{sta}\} &= \int_0^1 \mathbb{P}\{A_e = h \mid S_0 = h, A_{vc}, \text{sta}\} d\mathbb{P}\{A_{vc} \mid S_0 = h, \text{sta}\} \\ &= \int_0^{A^*} \frac{\alpha}{\alpha + (1-\alpha) p e^{-A_{vc}}} d\mathbb{P}\{A_{vc} \mid 0 \le A_{vc} < A^*\} \\ &< \int_0^{A^*} \frac{\alpha}{\alpha + (1-\alpha) p e^{-A^*}} d\mathbb{P}\{A_{vc} \mid 0 \le A_{vc} < A^*\} \\ &< \frac{\alpha}{\alpha + (1-\alpha) p e^{-A^*}}\end{aligned}$$

since the function $\frac{\alpha}{\alpha+(1-\alpha)pe^{-A_{vc}}}$ is increasing in $A_{vc}$.

On the other hand, the probability of successfully choosing high-ability entrepreneurs after observing two high signals conditional on the ability of the VC $A_{vc}$ is

$$\begin{aligned}&\mathbb{P}\{A_e = h \mid S_0 = S_1 = h, A_{vc}, \text{syn}\} \\ = &\int_0^1 \mathbb{P}\{A_e = h \mid S_0 = S_1 = h, A_{vc}, A_{vc,peer}, \text{syn}\} d\mathbb{P}\{A_{vc,peer} \mid S_0 = S_1 = h, \text{syn}\} \\ = &\int_0^1 \frac{\alpha}{\alpha + (1-\alpha) p^2 e^{-A_{vc}} e^{-x}} dx, \qquad (A.1)\end{aligned}$$

thus the overall success rate for syndicated deals is

$$\begin{aligned}\mathbb{P}\{A_e = h \mid S_0 = S_1 = h, \text{syn}\} &= \int_0^1 \mathbb{P}\{A_e = h \mid S_0 = S_1 = h, A_{vc}, \text{syn}\} d\mathbb{P}\{A_{vc} \mid S_0 = S_1 = h, \text{syn}\} \\ &= \int_{A^*}^1 \mathbb{P}\{A_e = h \mid S_0 = S_1 = h, A_{vc}, \text{syn}\} d\mathbb{P}\{A_{vc} \mid A^* \le A_{vc} \le 1\} \\ &= \int_{A^*}^1 \left(\int_0^1 \frac{\alpha}{\alpha + (1-\alpha) p^2 e^{-A_{vc}} e^{-x}} dx\right) d\mathbb{P}\{A_{vc} \mid A^* \le A_{vc} \le 1\} \\ &\ge \int_{A^*}^1 \left(\int_0^1 \frac{\alpha}{\alpha + (1-\alpha) p^2 e^{-A_{vc}}} dx\right) d\mathbb{P}\{A_{vc} \mid A^* \le A_{vc} \le 1\} \\ &= \int_{A^*}^1 \frac{\alpha}{\alpha + (1-\alpha) p^2 e^{-A_{vc}}} d\mathbb{P}\{A_{vc} \mid A^* \le A_{vc} \le 1\} \\ &> \frac{\alpha}{\alpha + (1-\alpha) p^2 e^{-A^*}},\end{aligned}$$

where the first equation is by the law of iterated expectations, the second is by results in Proposition 1.1, the third equality is by plugging in Equation (A.1), the inequality in the fourth line is by the fact that $e^{-x} \le 1$ on the whole interval $x \in [0, 1]$, the fifth line is by algebraic fact, and the last inequality is by the fact that $\frac{\alpha}{\alpha+(1-\alpha)p^2e^{-A_{vc}}}$ is an increasing function of $A_{vc}$.

Comparing the two success rates, we have

$$\begin{aligned}\mathbb{P}\left\{A_e = h \mid S_0 = h, \text{sta}\right\} &< \frac{\alpha}{\alpha + (1-\alpha)\, pe^{-A^*}} \\ &< \frac{\alpha}{\alpha + (1-\alpha)\, p^2 e^{-A^*}} \\ &< \mathbb{P}\left\{A_e = h \mid S_0 = S_1 = h, \text{syn}\right\}.\end{aligned}$$

□

## Proof of Corollary 1.2

*Proof.* The expected return of standalone deals is

$$\begin{aligned}E\left(R \mid \text{sta}\right) &= E\left\{E\left(R \mid A_{vc}, \text{sta}\right) \mid \text{sta}\right\} \\ &= E\left\{\frac{R_h\alpha}{\alpha + (1-\alpha)\, pe^{-A_{vc}}} \mid \text{syn}\right\} \\ &= \int_0^{A^*} \frac{R_h\alpha}{\alpha + (1-\alpha)\, pe^{-A_{vc}}} d\mathbb{P}\left\{A_{vc} \mid 0 \le A_{vc} < A^*\right\} \\ &< \frac{R_h\alpha}{\alpha + (1-\alpha)\, pe^{-A^*}},\end{aligned}$$

where the first equality is by the law of iterated expectations, the second equality is from the calculation of the expected return of a standalone deal conditional on VC's ability $A_{vc}$, the third equality comes from similar arguments as in the calculation of success rate, and the last inequality is by the fact that $\frac{R_h\alpha}{\alpha+(1-\alpha)pe^{-A_{vc}}}$ is an increasing function in $A_{vc}$. The expected return of syndication deals is

$$\begin{aligned}E\left(R \mid \text{syn}\right) &= E\left\{E\left(R \mid A_{vc}, \text{syn}\right) \mid \text{syn}\right\} \\ &= E\left\{\left(R_h + v\right) \int_0^1 \frac{\alpha}{\alpha + (1-\alpha)\, p^2 e^{-A_{vc}} e^{-x}} dx \mid \text{syn}\right\} \\ &\ge E\left\{\left(R_h + v\right) \frac{\alpha}{\alpha + (1-\alpha)\, p^2 e^{-A_{vc}}} \mid \text{syn}\right\} \\ &= \int_{A^*}^1 \left(R_h + v\right) \frac{\alpha}{\alpha + (1-\alpha)\, p^2 e^{-A_{vc}}} d\mathbb{P}\left\{A_{vc} \mid A^* \le A_{vc} \le 1\right\} \\ &> \left(R_h + v\right) \frac{\alpha}{\alpha + (1-\alpha)\, p^2 e^{-A^*}},\end{aligned}$$

where the first equality is by the law of iterated expectations, the second is by plugging in the expected return of a syndicated deal given the initial VC's ability $A_{vc}$ and integrating over the peer VC's ability $A_{vc,peer}$ over the uniform distribution on the interval $[0,1]$, the third and fourth lines are by similar arguments as in the calculation of success rate, and the last equality is by the fact that $\frac{\alpha}{\alpha+(1-\alpha)p^2e^{-A_{vc}}}$ is an increasing function in $A_{vc}$. Combining these two results, we have

$$
\begin{aligned}
E\left(R \mid \text{sta}\right) &< \frac{R_h \alpha}{\alpha + (1-\alpha)\, p e^{-A^*}} \\
&< \frac{(R_h + v)\, \alpha}{\alpha + (1-\alpha)\, p^2 e^{-A^*}} \\
&< E\left(R \mid \text{syn}\right)
\end{aligned}
$$

given that added-value from syndication $v > 0$ and $p > 0$. □

# A.2 Supplementary regression results

## Outcome variables using financing behavior measured on a longer horizon

Tables A.1 – A.8 provide more regression results to show our previous results in Section 1.5 are robust if we run the same set of regressions using the short-term performance measured on a longer horizon, namely two and three years.

Tables A.1 and A.2 show similar results to the results in the main texts. Compared with the results in the main text, Tables A.3 and A.4 show the peer effects on the ability to raise new rounds of funding and propensity to go IPO remain significantly positive, while that on bankruptcy seems to be vanishing with time going. For the outcome variable that measures the total number of startups going bankrupt, the regression coefficient of $D_{peer,it}$ decreases in absolute value and the significance also decreases.

When we further decompose the portfolio into a common pool and a non-overlapping pool as described in Section 1.5 and run the regressions using outcomes measured on a longer time horizon, we observe stronger significantly positive effects on the number of startups in the common portfolio set $C$ that go public or be acquired. However, as for the total number of startups that received new funding, the results seem negative and not always significant. We show that this is due to the positive effect on successful exits—the positive effect of an IPO on the total number of successful exits is so large that the number of remaining ongoing startups is significantly less, leading to a seemingly negative effect on the funding received. Summing up the total number of startups that either received new funding or successfully exited as in Column (3) of Tables A.5 and A.6, we see the strong positive effects on their performance.

Table A.1: Main effects of IPO on future performance (2 years)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{it}$ | 0.468*** | 0.145*** | -0.163*** |
| | (0.0453) | (0.0257) | (0.0216) |
| # of invested companies | 0.311*** | 0.0645*** | 0.0800*** |
| | (0.00200) | (0.00113) | (0.000950) |
| squared age | -0.00906*** | 0.00191*** | 0.00240*** |
| | (0.00130) | (0.000736) | (0.000619) |
| degree centrality | 165.6*** | 23.44*** | -0.421 |
| | (2.407) | (1.363) | (1.146) |
| Constant | 0.328*** | 0.166*** | -0.303*** |
| | (0.0564) | (0.0319) | (0.0268) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 30,375 | 30,375 | 30,375 |
| R-squared | 0.909 | 0.671 | 0.637 |

*Notes:* This table reports the main effects of IPO on the future 2-year performance of the VCs. The treatment variable $D_{it} = 1\{t \geq E_{it}\}$ is the time-varying indicator where $E_{it}$ is the event time, the first month VC $i$ has at least a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 2 years, that have successful exits within 2 years, and that file bankruptcy within 2 years, respectively. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Main effects of IPO on future performance (3 years)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{it}$ | 0.425*** | 0.204*** | -0.242*** |
| | (0.0489) | (0.0316) | (0.0261) |
| # of invested companies | 0.365*** | 0.0963*** | 0.118*** |
| | (0.00212) | (0.00137) | (0.00113) |
| squared age | -0.0122*** | 0.00515*** | 0.00173** |
| | (0.00150) | (0.000969) | (0.000800) |
| degree centrality | 157.1*** | 33.77*** | 1.300 |
| | (2.512) | (1.624) | (1.340) |
| Constant | 0.499*** | 0.157*** | -0.339*** |
| | (0.0539) | (0.0348) | (0.0288) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 27,228 | 27,228 | 27,228 |
| R-squared | 0.929 | 0.771 | 0.751 |

*Notes:* This table reports the main effects of IPO on the future 3-year performance of the VCs. The treatment variable $D_{it} = 1\{t \geq E_{it}\}$ is the time-varying indicator where $E_{it}$ is the event time, the first month VC $i$ has at least a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 3 years, that have successful exits within 3 years, and that file bankruptcy within 3 years, respectively. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Peer effects of IPO on future performance (2 years)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.339*** | 0.141*** | -0.0500*** |
| | (0.0386) | (0.0195) | (0.0193) |
| # of invested companies | 0.256*** | 0.0469*** | 0.0802*** |
| | (0.00231) | (0.00117) | (0.00116) |
| squared age | -0.00724*** | 0.00272*** | 0.000859 |
| | (0.00148) | (0.000749) | (0.000738) |
| degree centrality | 180.2*** | 34.40*** | 2.596* |
| | (3.035) | (1.538) | (1.516) |
| Constant | 0.339*** | -0.0105 | -0.189*** |
| | (0.0622) | (0.0315) | (0.0311) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 19,975 | 19,975 | 19,975 |
| R-squared | 0.893 | 0.603 | 0.673 |

*Notes:* This table reports the peer effects of IPO on the future 2-year performance. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 2 years, that have successful exits within 2 years, and that file bankruptcy within 2 years, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Peer effects of IPO on future performance (3 years)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.439*** | 0.207*** | 0.000874 |
| | (0.0406) | (0.0231) | (0.0228) |
| # of invested companies | 0.308*** | 0.0669*** | 0.118*** |
| | (0.00248) | (0.00141) | (0.00139) |
| squared age | -0.0102*** | 0.00877*** | -0.000388 |
| | (0.00172) | (0.000979) | (0.000965) |
| degree centrality | 181.6*** | 53.60*** | 2.417 |
| | (3.151) | (1.796) | (1.770) |
| Constant | 0.393*** | -0.142*** | -0.204*** |
| | (0.0596) | (0.0340) | (0.0335) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 17,908 | 17,908 | 17,908 |
| R-squared | 0.916 | 0.714 | 0.776 |

*Notes:* This table reports the peer effects of IPO on the future 3-year performance. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 3 years, that have successful exits within 3 years, and that file bankruptcy within 3 years, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Peer effects of IPO on future performance: common portfolio (2 years)

| VARIABLES | (1)<br># receive_fund | (2)<br># succeed | (3)<br># receive_fund & succeed | (4)<br># bankruptcy |
|---|---|---|---|---|
| $D_{peer,it}$ | -0.0206 | 0.0592*** | 0.0385** | -0.0198*** |
| | (0.0167) | (0.0109) | (0.0184) | (0.00572) |
| # of invested companies | 0.561*** | 0.184*** | 0.745*** | 0.104*** |
| | (0.0186) | (0.0122) | (0.0205) | (0.00635) |
| squared age | -0.00726*** | 0.00101* | -0.00625*** | -0.000680** |
| | (0.000802) | (0.000525) | (0.000884) | (0.000275) |
| degree centrality | 14.97*** | -1.988** | 12.98*** | 0.210 |
| | (1.319) | (0.864) | (1.454) | (0.452) |
| Constant | 0.102** | -0.112*** | -0.0100 | -0.0351** |
| | (0.0484) | (0.0317) | (0.0533) | (0.0166) |
| month×industry FE | ✓ | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ | ✓ |
| Observations | 7,915 | 7,915 | 7,915 | 7,915 |
| R-squared | 0.526 | 0.547 | 0.555 | 0.623 |

*Notes:* This table reports the peer effects of IPO on the future 2-year performance of the common portfolio subsample, i.e., the overlapping set $C$ in Figure 1.5. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the common portfolio that receive new funding within 2 years, that have successful exits within 2 years, and that file bankruptcy within 2 years, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Peer effects of IPO on future performance: common portfolio (3 years)

| VARIABLES | (1)<br># receive_fund | (2)<br># succeed | (3)<br># receive_fund & succeed | (4)<br># bankruptcy |
|---|---|---|---|---|
| $D_{peer,it}$ | -0.0631*** | 0.110*** | 0.0464** | -0.00820 |
| | (0.0164) | (0.0122) | (0.0188) | (0.00626) |
| # of invested companies | 0.599*** | 0.254*** | 0.853*** | 0.117*** |
| | (0.0188) | (0.0141) | (0.0216) | (0.00719) |
| squared age | -0.00908*** | 0.00265*** | -0.00642*** | -0.00135*** |
| | (0.000880) | (0.000658) | (0.00101) | (0.000337) |
| degree centrality | 13.25*** | -0.0876 | 13.16*** | 0.595 |
| | (1.282) | (0.959) | (1.470) | (0.490) |
| Constant | 0.189*** | -0.207*** | -0.0178 | -0.0206 |
| | (0.0457) | (0.0342) | (0.0524) | (0.0175) |
| month×industry FE | ✓ | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ | ✓ |
| Observations | 6,934 | 6,934 | 6,934 | 6,934 |
| R-squared | 0.608 | 0.654 | 0.638 | 0.719 |

*Notes:* This table reports the peer effects of IPO on the future 3-year performance of the common portfolio subsample, i.e., the overlapping set $C$ in Figure 1.5. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the common portfolio that receive new funding within 3 years, that have successful exits within 3 years, and that file bankruptcy within 3 years, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: Peer effects of IPO on future performance: non-overlapping portfolio (2 years)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.303*** | 0.107*** | -0.0428** |
| | (0.0390) | (0.0193) | (0.0196) |
| # of invested companies | 0.258*** | 0.0496*** | 0.0810*** |
| | (0.00235) | (0.00116) | (0.00118) |
| squared age | -0.00219 | 0.00210*** | 0.000955 |
| | (0.00151) | (0.000746) | (0.000759) |
| degree centrality | 167.2*** | 28.55*** | 2.103 |
| | (3.022) | (1.494) | (1.520) |
| Constant | 0.139** | -0.00862 | -0.185*** |
| | (0.0635) | (0.0314) | (0.0319) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 19,444 | 19,444 | 19,444 |
| R-squared | 0.890 | 0.605 | 0.676 |

*Notes:* This table reports the peer effects of IPO on the future 2-year performance of the non-overlapping portfolio subsample, i.e., the non-overlapping set $A$ in Figure 1.5. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's non-overlap portfolio that receive new funding within 2 years, that have successful exits within 2 years, and that file bankruptcy within 2 years, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Peer effects of IPO on future performance: non-overlapping portfolio (3 years)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.434*** | 0.145*** | 0.00140 |
| | (0.0412) | (0.0227) | (0.0233) |
| # of invested companies | 0.309*** | 0.0713*** | 0.119*** |
| | (0.00252) | (0.00139) | (0.00142) |
| squared age | -0.00468*** | 0.00775*** | 8.60e-05 |
| | (0.00176) | (0.000973) | (0.000998) |
| degree centrality | 169.2*** | 45.66*** | 1.600 |
| | (3.145) | (1.735) | (1.779) |
| Constant | 0.198*** | -0.143*** | -0.203*** |
| | (0.0610) | (0.0336) | (0.0345) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 17,400 | 17,400 | 17,400 |
| R-squared | 0.914 | 0.716 | 0.778 |

*Notes:* This table reports the peer effects of IPO on the future 3-year performance of the non-overlapping portfolio subsample, i.e., the non-overlapping set $A$ in Figure 1.5. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's non-overlap portfolio that receive new funding within 3 years, that have successful exits within 3 years, and that file bankruptcy within 3 years, respectively. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Weighted regressions

Tables A.9 – A.11 provide more regression results to show our previous results in Section 1.5 are robust if we run the weighted version of the regressions using the weights of the edges in the VC network, where the weights of an edge between two VCs are defined as the total number of co-investments they make.

Table A.9: Peer effects of IPO on future performance (1 year, weighted regressions)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.0843 | 0.0539** | -0.0978*** |
| | (0.0565) | (0.0239) | (0.0234) |
| # of invested companies | 0.150*** | 0.0251*** | 0.0390*** |
| | (0.00193) | (0.000820) | (0.000800) |
| squared age | -0.000379 | -0.000295 | 0.000685 |
| | (0.00114) | (0.000482) | (0.000471) |
| degree centrality | 150.2*** | 13.22*** | 2.200** |
| | (2.643) | (1.120) | (1.093) |
| Constant | 0.116*** | 0.0848*** | -0.0450*** |
| | (0.0330) | (0.0140) | (0.0137) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 22,055 | 22,055 | 22,055 |
| R-squared | 0.813 | 0.409 | 0.506 |

*Notes:* This table reports the peer effects of IPO on the future 1-year performance. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 1 year, that have successful exits within 1 year, and that file bankruptcy within 1 year, respectively. Observations are weighted using the weights in the VC network, where the weight of an edge between two VCs is defined as the number of times they co-invested in the history. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.10: Peer effects of IPO on future performance (2 years, weighted regressions)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.283*** | 0.0287 | -0.128*** |
| | (0.0606) | (0.0308) | (0.0304) |
| # of invested companies | 0.254*** | 0.0456*** | 0.0774*** |
| | (0.00231) | (0.00117) | (0.00116) |
| squared age | -0.00489*** | 0.00163** | 0.000363 |
| | (0.00146) | (0.000740) | (0.000730) |
| degree centrality | 184.4*** | 35.53*** | 1.509 |
| | (3.109) | (1.579) | (1.558) |
| Constant | 0.273*** | 0.152*** | -0.0331* |
| | (0.0384) | (0.0195) | (0.0192) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 19,975 | 19,975 | 19,975 |
| R-squared | 0.891 | 0.596 | 0.681 |

*Notes:* This table reports the peer effects of IPO on the future 2-year performance. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 2 years, that have successful exits within 2 years, and that file bankruptcy within 2 years, respectively. Observations are weighted using the weights in the VC network, where the weight of an edge between two VCs is defined as the number of times they co-invested in the history. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.11: Peer effects of IPO on future performance (3 years, weighted regressions)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | # receive_fund | # succeed | # bankruptcy |
| $D_{peer,it}$ | 0.297*** | -0.0780** | -0.137*** |
| | (0.0583) | (0.0333) | (0.0329) |
| # of invested companies | 0.306*** | 0.0649*** | 0.116*** |
| | (0.00249) | (0.00142) | (0.00140) |
| squared age | -0.00643*** | 0.00678*** | -0.00126 |
| | (0.00170) | (0.000972) | (0.000959) |
| degree centrality | 186.7*** | 55.55*** | 0.606 |
| | (3.231) | (1.847) | (1.823) |
| Constant | 0.363*** | 0.211*** | 0.0347 |
| | (0.0405) | (0.0232) | (0.0229) |
| month×industry FE | ✓ | ✓ | ✓ |
| investor FE | ✓ | ✓ | ✓ |
| Observations | 17,908 | 17,908 | 17,908 |
| R-squared | 0.914 | 0.706 | 0.780 |

*Notes:* This table reports the peer effects of IPO on the future 3-year performance. The treatment variable $D_{peer,it} = 1\{t \geq E_{peer,it}\}$ is the time-varying indicator where $E_{peer,it}$ is the event time, the first month VC $i$ has at least one peer that has a startup IPO in month $t$. The model used is the generalized diff-in-diff regression

$$Y_{it} = \alpha_i + \lambda_{t,ind} + \tau D_{peer,it} + \gamma^{\mathrm{T}} X_{it} + \epsilon_{it},$$

where $\tau$ is our parameter of interest that measures the main effects, $\alpha_i$ is the VC fixed effects, $\lambda_{t,ind}$ is the interaction between time and industry fixed effects, and $X_{it}$ are other control variables including the total number of startups VC $i$ invested at time $t$, the age and squared-age of VC $i$ at time $t$, and the degree centrality of VC $i$ in the dynamic network at time $t$. The outcome variables of interest in Columns (1) – (3) are the total number of startups in the VC's portfolio that receive new funding within 3 years, that have successful exits within 3 years, and that file bankruptcy within 3 years, respectively. Observations are weighted using the weights in the VC network, where the weight of an edge between two VCs is defined as the number of times they co-invested in the history. The regression is run on the subsample that does not have its own IPO event. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

# Appendix B

# Appendix to Chapter 2

## B.1 Supplementary regression results

This appendix provides several supplementary regression results to those in the main text. Table B.1 reports the correlation between long-term performance and common ownership using the total number of early rounds to define experienced investors, showing similar results to Table 2.1 when we use the total number of investment rounds. Table B.2 reports the subgroup effects of common ownership on the short-term financing behavior using the top 50% most valued financing rounds as the successful treatment events. Tables B.3 – B.7 report the results of some same set of regressions as in the main text using 1-1 matching in the matching algorithm as robustness checks.

Table B.1: Performance and common ownership indicator (use # of early rounds to define experienced investors)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | ipo | acq | closed |
| common_code | 0.00739** | 0.0474*** | 0.00484 |
| | (0.00309) | (0.0105) | (0.00932) |
| Constant | -0.0280*** | 0.120*** | 0.365*** |
| | (0.00326) | (0.0110) | (0.00981) |
| control variables | ✓ | ✓ | ✓ |
| location FE | ✓ | ✓ | ✓ |
| found year×industry FE | ✓ | ✓ | ✓ |
| Observations | 8,698 | 8,698 | 8,698 |
| R-squared | 0.068 | 0.043 | 0.093 |

*Notes:* This table reports the relationship between long-term performance and common ownership indicator. Results are from the regressions

$$\texttt{outcome} = \alpha_{\text{state}} + \lambda_{\text{ind,foundyear}} + \tau \cdot \texttt{common\_code} + \gamma^{\mathrm{T}} X + \epsilon,$$

where $\texttt{outcome} \in \{\texttt{ipo, acq, closed}\}$ is the measure of long-term performance and `common_code` is the treatment indicator of whether a startup is commonly owned. We control for startup features such as the logarithm of the total number of employees and whether the average number of early rounds of its investors is above the median. Location fixed effects and the two-way interacted fixed effects between industry and found-year are also included in the regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.2: Subgroup effects on short-term financing behavior (ratio > p50)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | 0.0209*** | 0.0354*** | 0.0485*** |
| | (0.00115) | (0.00136) | (0.00140) |
| Constant | 0.176*** | 0.313*** | 0.454*** |
| | (0.000459) | (0.000545) | (0.000560) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 5 | 5 | 5 |
| Observations | 810,308 | 810,308 | 810,308 |
| R-squared | 0.096 | 0.134 | 0.202 |

*Notes:* This table reports the subgroup effects of common ownership on the short-term financing behavior when we use only the top 50% most valued financing rounds as the successful treatment events. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-5 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.3: Effects of common ownership on long-term performance (1-1 matching)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | ipo | acq | closed |
| treatment | 0.00114*** | 0.0137*** | -0.00293*** |
| | (0.000167) | (0.000693) | (0.000593) |
| Constant | 0.00917*** | 0.154*** | 0.112*** |
| | (0.000106) | (0.000442) | (0.000378) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 1 | 1 | 1 |
| Observations | 1,114,368 | 1,114,368 | 1,114,368 |
| R-squared | 0.627 | 0.538 | 0.534 |

*Notes:* This table reports the effects of common ownership on the long-term performance of the startups. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where $\texttt{outcome} \in \{\texttt{ipo, acq, closed}\}$ is the measure of long-term performance and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-1 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.4: Effects of common ownership on short-term financing behavior (1-1 matching)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | 0.00945*** | 0.0171*** | 0.0233*** |
| | (0.000874) | (0.000972) | (0.000952) |
| Constant | 0.175*** | 0.306*** | 0.436*** |
| | (0.000557) | (0.000620) | (0.000607) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 1 | 1 | 1 |
| Observations | 1,114,368 | 1,114,368 | 1,114,368 |
| R-squared | 0.327 | 0.430 | 0.523 |

*Notes:* This table reports the effects of common ownership on the short-term financing behavior of the startups. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-1 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: Subgroup effects on short-term financing behavior (ratio > p75, 1-1 matching)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | 0.0196*** | 0.0321*** | 0.0432*** |
| | (0.00220) | (0.00259) | (0.00264) |
| Constant | 0.175*** | 0.311*** | 0.449*** |
| | (0.00155) | (0.00183) | (0.00186) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 1 | 1 | 1 |
| Observations | 124,231 | 124,231 | 124,231 |
| R-squared | 0.165 | 0.204 | 0.268 |

*Notes:* This table reports the subgroup effects of common ownership on the short-term financing behavior when we use only the top 25% most valued financing rounds as the successful treatment events. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-1 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.6: Subgroup effects on short-term financing behavior (ratio > p50, 1-1 matching)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | 0.0199*** | 0.0318*** | 0.0414*** |
| | (0.00155) | (0.00182) | (0.00185) |
| Constant | 0.177*** | 0.316*** | 0.460*** |
| | (0.00110) | (0.00129) | (0.00130) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 1 | 1 | 1 |
| Observations | 251,860 | 251,860 | 251,860 |
| R-squared | 0.166 | 0.210 | 0.279 |

*Notes:* This table reports the subgroup effects of common ownership on the short-term financing behavior when we use only the top 50% most valued financing rounds as the successful treatment events. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window, and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, location of the headquarters, and the deal year $t$. We use the 1-1 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.7: Effects on financing performance when a negative event happens (1-1 matching)

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | financed_180d | financed_365d | financed_730d |
| treatment | -0.00132 | -0.00116 | -0.00121 |
| | (0.00120) | (0.00162) | (0.00201) |
| Constant | 0.0290*** | 0.0562*** | 0.0951*** |
| | (0.000843) | (0.00114) | (0.00141) |
| matched-pair FE | ✓ | ✓ | ✓ |
| four-way interacted FE | ✓ | ✓ | ✓ |
| # matched | 1 | 1 | 1 |
| Observations | 74,116 | 74,116 | 74,116 |
| R-squared | 0.214 | 0.251 | 0.293 |

*Notes:* This table reports the effects of common ownership on the short-term financing behavior when a negative event, closure, occurs. Results are from the regressions

$$\texttt{outcome}_{it} = \alpha_{\text{pair}} + \lambda_{\text{ind,foundyear,location},t} + \tau \cdot D_{it} + \epsilon_{it},$$

where `outcome` is a dummy variable indicating whether it successfully gets new rounds of funding within a certain time window, and the treatment variable $D_{it}$ is the binary indicator of whether a startup is in the common ownership pool at time $t$. We control for the matched-pair fixed effects and the four-way interacted fixed effects among industry, found year, headquarters location, and the deal year $t$. We use the 1-1 matching algorithm in these regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

# Appendix C

# Appendix to Chapter 3

## C.1 Proofs

### Proof of Equilibrium

*Proof.* **No Venture Debt**

We first solve the world without venture debt, where only venture capital takes action. When the venture capital observes a firm with a valuation of $a$, the venture capital updates the belief of the probability of the firm being a high type as

$$\alpha_a = \frac{\alpha P_H}{\alpha P_H + (1-\alpha) P_L}.$$

If the venture capital investigates the firm and invests in high type, the profit of the venture capital is

$$R_I = -cI + \alpha_a \left( a\mu_H \cdot \frac{I}{I+a} - I \right).$$

If the venture capital does not investigate but invests in all firms, the profit of the venture capital is

$$R_{NI} = \frac{aI}{I+a} \left\{ \alpha_a \mu_H + (1-\alpha_a)\, \mu_L \right\} - I.$$

If the venture capital does not invest in any firms,

$$R_N = 0.$$

To optimize the return of venture capital investors, the action of the venture capital is as follows.

- When $c < (1-\alpha_a)\frac{I+a-a\mu_L}{I+a} = C_a^*$ and $c < \alpha_a\left(\frac{a\mu_H}{I+a} - 1\right)$, the venture capital investigates and only invest in high type firms.
- When $c > c_a^*$ and $\alpha_a\mu_H + (1-\alpha_a)\mu_L - 1 > \frac{I}{a}$, the venture capital does not investigate and always invest in firms.
- Otherwise, there is no investment.

If the venture capital observes a firm in the state $b$, the venture capital updates the belief of the probability of the firm being a high type as

$$\alpha_b = \frac{\alpha(1-P_H)}{\alpha(1-P_H) + (1-\alpha)(1-P_L)}.$$

Similarly, to maximize the profits, the venture capital's action is as follows.

- When $c < (1-\alpha_b)\frac{I+b-b\mu_L}{I+b} = C_b^*$ and $c < \alpha_b\left(\frac{b\mu_H}{I+b} - 1\right)$, the venture capital investigates and only invest in high type firms.
- When $c > c_b^*$ and $\alpha_b\mu_H + (1-\alpha_b)\mu_L - 1 > \frac{I}{b}$, the venture capital does not investigate and always invest in firms.
- Otherwise, there is no investment.

**With Venture Debt**

Banks are rational and maximize their payoff $ER - C_B - D$ by

$$\Pi_B = -C_B - D + P_I(C_B)R.$$

Venture capital does not change their actions, except by updating the belief based on the bank's behavior. Therefore, the action of the venture capital is as follows. When the venture capital observes a firm with a valuation of $a$, the venture capital updates the belief of the probability of the firm being a high type as

$$\alpha_a(C_B) = \frac{\alpha(C_B)P_H}{\alpha(C_B)P_H + (1-\alpha(C_B))P_L}.$$

- When $c < (1-\alpha_a(C_B))\frac{I+a-a\mu_L}{I+a} = c_a(C_B)$ and $c < \alpha_a(C_B)\left(\frac{a\mu_H}{I+a} - 1\right)$, the venture capital investigates and only invest in high type firms.

- When $c > c_a(C_B)$ and $\alpha_a(C_B)\mu_H + (1-\alpha_a(C_B))\mu_L - 1 > \frac{I}{a}$, the venture capital does not investigate and always invest in firms.
- Otherwise, there is no investment.

If the venture capital observes a firm in the state $b$, the venture capital updates the belief of the probability of the firm being a high type as

$$\alpha_b(C_B) = \frac{\alpha(C_B)\,(1-P_H)}{\alpha(C_B)\,(1-P_H) + (1-\alpha(C_B))\,(1-P_L)}.$$

Similarly, to maximize the profits, the venture capital's action is as follows.

- When $c < (1-\alpha_b(C_B))\,\frac{I+b-b\mu_L}{I+b} = c_b(C_B)$ and $c < \alpha_b(C_B)\left(\frac{b\mu_H}{I+b} - 1\right)$, the venture capital investigates and only invest in high type firms.
- When $c > c_b(C_B)$ and $\alpha_b(C_B)\mu_H + (1-\alpha_b(C_B))\mu_L - 1 > \frac{I}{b}$, the venture capital does not investigate and always invest in firms.
- Otherwise, there is no investment.

According to the level of $c$, there are three possible equilibria in this model.

1. When $c$ is sufficiently small, based on the reasoning above, venture capital always does due diligence and invests in high-type firms. Therefore, only high-type firms will be invested in by the venture capital. Given venture capital's behavior, $P_I(C_B) = \alpha(C_B)$. The level of screening cost $C_B^0$ banks take solves

$$\begin{cases} -1 + \alpha'(C_B)R = 0, & (\text{ F.O.C }) \\ -C_B - D + \alpha(C_B)R = 0. \end{cases}$$

   If venture capital always does due diligence, $c$ must satisfy

$$c \leq (1-\alpha_a(C_B^0))\frac{I+a-a\mu_L}{I+a} = c_a(C_B^0).$$

   Therefore, when $c < c_a(C_B^0)$, banks spend $C_B^0$ on screening and venture capital always does due diligence and invests in good firms.

2. When $c$ is sufficiently large, venture capital always takes the signal of venture debt and invests in all firms with venture debt when $V = a$, and does due diligence when $V = b$.

In this case, $P_I(C_B) = \alpha(C_B) + (1 - \alpha(C_B))P_L$. The level of screening cost $C_B^2$ banks take solves

$$\begin{cases} -1 + \alpha'(C_B)(1 - P_L)R = 0, \quad (\text{ F.O.C }) \\ -C_B - D + (\alpha(C_B) + (1 - \alpha(C_B))P_L)R = 0. \end{cases}$$

If venture capital always invests in firms with venture debt when $V = a$, $c$ must satisfies

$$c \geq (1 - \alpha_a(C_B^2))\frac{I + a - a\mu_L}{I + a} = c_a(C_B^2).$$

Therefore, when $c > c_a(C_B^2)$, banks spend $C_B^2$ on screening, and venture capital does due diligence and invests in good firms when observing $V = b$, and always invests in firms with venture debt when $V = a$.

3. When $c \in (c_a(C_B^0), c_a(C_B^2))$, there is no pure strategy for venture capital. If banks spend $C_B^0$, since $c > c_a(C_B^{)})$, venture capital will take the signal of venture debt. Anticipating that, banks will change their action and spend $C_B^2$ on screening. Then venture capital has the incentive to ignore the signal as $c < c_a(C_B^2)$. Therefore, venture capital will play a mixed strategy. When $V = a$, venture capital plays a mixed strategy of investing in all firms with venture debt with some probability $p$ and investing in high-type firms after due diligence with probability $1 - p$. When $V = b$, venture capital invests in high-type firms after due diligence. In this case, $P_I(C_B) = \alpha(C_B) + (1 - \alpha(C_B))p \times P_L$. The level of screening cost $C_B^1$ banks take solves

$$\begin{cases} -1 + \alpha'(C_B)(1 - p \times P_L)R = 0, \quad (\text{ F.O.C }) \\ -C_B - D + (\alpha(C_B) + (1 - \alpha(C_B))p \times P_L)R = 0. \end{cases}$$

Since venture capital has no preference over doing due diligence or not when $V = a$, $p$ must satisfies

$$c = (1 - \alpha_a(C_B^1(p)))\frac{I + a - a\mu_L}{I + a} = c_a(C_B^1(p)).$$

□

## Proof of Proposition 2

*Proof.* To generalize the question, let $C_B$ and $R$ denote the solutions to

$$\begin{cases} -1 + \alpha'(C_B)(1 - q)R = 0, \quad (\text{ F.O.C }) \\ -C_B - D + (\alpha(C_B) + (1 - \alpha(C_B))q)R = 0. \end{cases}$$

$C_B^0$, $R^0$ solves the question when $q = 0$, $C_B^1$, $R^1$ solves when $q = p \times P_L$, and $C_B^2$, $R^2$ solves when $q = P_L$. We will prove the proposition by showing that $C_B$ decreases in $q$ and $R$ decreases in $q$.

$$\begin{cases} -1 + \alpha'(C_B)(1-q)R = 0, \quad (\text{ F.O.C }) \\ -C_B - D + (\alpha(C_B) + (1-\alpha(C_B))q)R = 0, \end{cases}$$

$$\begin{aligned} &\Rightarrow \frac{1}{R} = \alpha'(c)(1-q) = \frac{\alpha(c) + (1-\alpha(c))q}{c+D}, \\ &\Rightarrow \alpha'(c)(1-q)(c+D) = \alpha(c) + (1-\alpha(c))q. \end{aligned}$$

Derive both sides with respect to $q$, we get

$$(\alpha''(c)(c+D)(1-q) + \alpha'(c)(1-q))\frac{dc}{dq} - \alpha'(c)(c+D) = \alpha'(c)\frac{dc}{dq} + 1 - \alpha - \alpha'(c)q\frac{dc}{dq},$$
$$\Rightarrow \alpha''(c)(c+D)(1-q)\frac{dc}{dq} = \alpha'(c)(c+D) + 1 - \alpha(c).$$

Since $\alpha''(c) < 0, \alpha'(c) > 0, q < 1, \alpha < 1$, we have $\frac{dc}{dq} < 0$, namely $C_B$ decreases in $q$.

Let $r = \frac{1}{R} = \alpha'(c)(1-q)$.

$$\begin{aligned} \frac{dr}{dq} &= \alpha''(c)(1-q)\frac{dc}{dq} - \alpha'(c) \\ &= \alpha''(c)(1-q)\frac{\alpha'(c)(c+D) + 1 - \alpha(c)}{\alpha''(c)(c+D)(1-q)} - \alpha'(c) \\ &= \frac{\alpha'(c)(c+D) + 1 - \alpha(c)}{c+D} - \alpha'(c) \\ &= \frac{1-\alpha(c)}{c+D} > 0, \\ \Rightarrow &\frac{dR}{dq} < 0. \end{aligned}$$

Thus $R$ decreases in $q$. □

## C.2 List of good venture capital firms

In this appendix, we describe the construction of our list of good venture capital firms in detail. We provide the name of these venture capitals and our reference of the rankings.

The list of good venture capital firms (in alphabetical order):

1. Accel Partners
2. Alexandria Venture
3. Alumni Ventures Group
4. Andreessen Horowitz
5. Bessemer Trust
6. Canaan Partners
7. Founders Fund
8. General Catalyst
9. Goldman Sachs
10. Greycroft
11. GV
12. Higher Ground Labs
13. Insight Venture Partners
14. Khosla Ventures
15. Kleiner Perkins Caufield & Byers
16. Lightspeed Venture Partners
17. New Enterprise Associates
18. Quake Capital Partners
19. Revolution
20. Sequoia Capital
21. Sinai Ventures

22. Social Capital
23. SV Angel
24. True Ventures

This list is constructed referring to three convincing venture capital rankings. We also attach their website addresses for reference:

1. https://www.angelkings.com/top-venture-capital-firms
2. https://www.forbes.com/sites/alejandrocremades/2018/07/18/top-10-venture-capital-investors-that-every-entrepreneur-should-be-pitching-right-now/4017b74f1ced
3. https://pitchbook.com/news/articles/how-20-big-name-us-vc-firms-invest-at-series-a-b