Strategic Network Formation and Performance in the Venture Capital Industry*

Kosuke Uetake[†]

Northwestern University

Job Market Paper

November 2012

Abstract

I investigate the network structure of venture capitalists based on co-investments, and the effects of network structure on investment performance. As venture capitalists select their partners, network structure is endogenously determined in equilibrium. Using comprehensive data on venture capital firms in the U.S., I jointly estimate a model of strategic network formation and a performance equation, taking endogeneity of network structure into account. In the estimation of the strategic network formation model, instead of imposing an equilibrium selection rule, I exploit the partial identification approach. My estimation strategy relies only on the necessary conditions of pairwise stability and is computationally feasible. I find that the network of venture capitalists tends to be homophilous in terms of asset size, but anti-homophilous in terms of investment experience and industry expertise. Moreover, I find that not taking the endogeneity into account results in significant overestimates of the effects of the network structure on investment performance. Lastly, I conduct a counterfactual policy experiment in which the government makes direct investment into the market

*I am especially grateful to Rob Porter for his advice, guidance, and support. I also thank my committee members, Yael Hochberg, Elie Tamer, and Yasutora Watanabe, for helpful discussion and advice. I have also benefitted from comments by Mike Abito, Vineet Bhagwat, Joseph Goodman, Garrett Johnson, Kei Kawai, Konrad Menzel, Aviv Nevo, Ken Onishi, Tiago Pires, Mike Powell, Bill Rogerson and Northwestern Industrial Organization Seminar participants. Remaining errors are my own. Please check my web site for the latest version of the paper: http://gradstudents.wcas.northwestern.edu/~kue667/indexjm.html

[†]Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. Email: uetake@u.northwestern.edu

by establishing a venture capital firm. I examine the effects of the additional venture capital firm on equilibrium network structure and investment performance.

1 Introduction

The venture capital industry is one of the driving forces of U.S. economic growth by promoting the innovative behavior of entrepreneurs. Most of the start-up companies that attract substantial public attention – such as Amazon, Apple, eBay, Google, and Microsoft – have been backed by venture capital funds. Annual venture capital investments have expanded from virtually zero in the mid-1970s to almost $30 billion in 2011, as revenues from venture–backed companies now account for more than 20% of total U.S. GDP.

The significance of venture capitalists has grown in tandem with the industry. Venture capitalists are professional and institutional investors who fund early-stage start-up companies with high-potential and high-risk. They also engage in monitoring and nurturing those companies, which require extensive industry knowledge and investment experience.

A prominent feature of venture capital industry is the networks of venture capitalists, which are observed through co-investment behavior in the start-up companies. Because venture capitalists can exchange information regarding promising investment deals or knowledge about new technologies, impacts of co-investment networks on venture capitalists' operation are substantial. Moreover, the structure of the networks is likely to influence the venture capitalists' investment performance since the network structure may affect the ability to screen high-quality start-up companies (see, e.g., Sah and Stiglitz, 1986) and the ability to monitor and nurture investments (see, e.g., Hellman and Puri, 2002; Kaplan and Stromberg, 2004).

Assessing the causal effects of the network structure is empirically challenging, however, as venture capitalists select their partners, and network structure is thus determined in equilibrium. If unobserved heterogeneity among venture capitalists influences both the formation of the network and performance, then there may be endogeneity bias. For example, a venture capitalist who has greater monitoring ability is more likely to be connected with other venture capitalists, and the venture capitalist's investment performance may be better due to its monitoring ability. Consequently, I need to account for the endogeneity of network structure in estimating its effects on performance,[1] although existing papers studying the

[1] Although we focus on the venture capital network, endogeneity of network structure arises in many different situations such as the networks of airlines through mutual service agreements and the networks

effect of network structure on performance typically take the network structure as given.[2] My first goal in this paper is to provide a useful framework for estimating the effects of the network structure on their economic outcomes.

My second goal is to understand the structure of venture capital networks. Because each venture capitalist has a different investment experience, industry knowledge and access to capital, it is important to identify key characteristics of venture capital networks. In other words, where does the value of connections comes from? The literature (see, e.g., McPherson, Smith-Lovin and Cook, 2001) has documented that social networks such as friendship networks tend to be *homophilous*: People tend to form ties with those who are similar to them in wide range of characteristics. In contrast, less is known about networks in competitive or strategic situations such as the venture capital industry. From a theoretical point of view, the question of whether venture capital networks should be homophilous is ambiguous. One could imagine that, for example, venture capitalists may form ties with those who are similar to them in order to avoid conflicts of interest or to pool similar resources.[3] On the other hand, they may form ties with those who are dissimilar to them because, for example, venture capitalists with more investment experience may prefer being connected with newly established venture capital firms so as to take advantage of their greater bargaining power over younger venture capitalists. Hence, the question as to whether venture capital networks exhibit homophily must be addressed empirically.

In many countries, the government tries to intervene in the venture capital market to support innovation processes, and answering two questions in this paper has potentially large policy implications. If more connections among venture capitalists improve their investment performance, the government is willing to implement policies that can facilitate more connections: e.g., the government may provide subsidies to some venture capitalists, or the government may make direct investment in the market by establishing government funded

of countries based on free trade agreements. In those organizational networks, the networks are generally not randomly formed. In networks of airlines through service agreements, for example, airlines strategically choose their business partners so that they can earn competitive advantages through such connections. Hence, the correlation between the network structure and the outcome may also occur in many other contexts. Our estimation and identification strategies are not specific to the venture capital industry and can be applied to networks in many other industries.

[2]Examples are Conlon and Udry (2010), Hochberg, Ljungqvist and Lu (2007) and Kinnan and Townsend (2011). These papers try to address the potential endogeneity problem by exploiting additional data, imposing timing assumption, or using panel structure of data.

[3]A potential reason that venture capitalists have different interest is that younger venture capitalists may want to spend their capital for accumulating investment experience or industry knowledge, while experienced venture capitalists may want to spend capital only for the best investment deals.

venture capital firms. Understanding the venture capital network structure or the value of connections provides an important guidance for implementing such policies effectively. This paper presents a useful framework for the governments to carry out policies that promote innovations by the start-up companies.

In order to answer these questions, I jointly estimate a structural model of strategic network formation and a performance (or an outcome) equation.[4] A key point of the analysis is that there is unobserved heterogeneity of venture capitalists which influence both the formation of links and performance, and create a selection bias in the performance equation.[5] In my analysis, I first consider the estimation of strategic network formation models and then examine the estimation of the outcome equation together with the strategic network formation model, thereby addressing the endogeneity issue of the network structure.

My model of strategic network formation is based on Jackson and Wolinsky (1996). In this model, each node in a network represents a venture capitalist and there is a link between two of them if and only if they co-invest in the same start-up company (i.e., join the same syndicate). In order to account for the fact that a syndication investment requires mutual agreement among the venture capitalists, I exploit the solution concept of *pairwise stability* proposed by Jackson and Wolinsky (1996), which requires the consent of both players for the formation of a link. The pairwise stability conditions generate inequality conditions, which form the basis of my estimation strategy.

The estimation of strategic-network formation models is not straightforward. There are two main challenges. First, there typically exist multiple equilibria (pairwise stable networks in my case): For a given set of parameter values, the model may admit more than one pairwise stable network. In the absence of equilibrium-selection mechanism, this multiplicity translates into a non-uniqueness of outcome predictions and may result in a lack of point-identification.[6]

[4] We use "outcome equation" and "performance equation" interchangeably.

[5] One way to address the endogeneity problem is, of course, to use instrumental variables. However, as pointed out by Sorensen (2007), it is difficult to find an instrument that is independent of the shock to the outcome but correlated to the network structure in our case. For example, the distance between venture capitalists is a candidate instrument, since it is easier for venture capitalists to form links with those who are located closer. However, this factor may also influence their investment performance because it is easier for them to provide better monitoring service for start-up companies. This lack of suitable instrumental variables is a reason that we use a structural model in this paper.

[6] In order to avoid the lack of point identification, researchers typically impose a certain equilibrium selection mechanism. For example, Mazzeo (2002a) assumes an order of the players' decision node (it is unobservable in the data) to guarantee the uniqueness of Nash equilibrium in an entry game with vertically differentiated players. Bajari, Benkard and Levin (2007) and Bajari et al. (2010) assume that the same equilibrium is played conditional on exogenous characteristics in each market, and then propose a two-step

The second challenge is computational. Even under a unique pairwise stable network for a given equilibrium selection mechanism, the number of possible network configurations increases exponentially as the number of players in the network increases. Hence, even with a relatively small number of players, it can be computationally infeasible to fully solve the model and to evaluate the likelihood or moment conditions.

To deal with the first challenge, I exploit the partial-identification approach and estimate the model by a moment-inequality estimator. In particular, I construct inequality and equality conditions from *only* the necessary conditions of pairwise stability.[7] Since any pairwise-stable networks must satisfy the necessary conditions, my estimation strategy is robust to the multiplicity of pairwise-stable networks, although the parameters of the model are only partially (or set) identified.

The computational cost is reduced in two ways. I construct moment conditions based on deviations from the *observed* network[8] and I exploit the fact that pairwise stability requires a network to be robust to a deviation of *one* link at a time. These facts allow me to estimate the network formation model without finding all pairwise stable networks (fully solving the game). As a result, it is possible to significantly reduce the computational burden to only the number of possible links (i.e., quadratic in the number of players, rather than exponential). The computational cost thus becomes manageable, thereby permitting estimation of strategic network formation models.

After proposing the estimation strategy of the network-formation model, I combine the outcome equation and the strategic network formation model to correct the endogeneity bias of network formation. My model of the investment performance of venture capitalists is similar to the model in Hochberg, Ljungvist and Lu (2007), in which the performance of

estimator of a dynamic game with incomplete information and a static game with incomplete information, respectively.

Another way to achieve point identification is to focus on the property that is invariant across any equilibrium. A seminal paper by Bresnahan and Reiss (1991), for example, estimates an entry game with homogeneous players. They exploit the fact that the number of entering firms does not vary across all Nash equilibria. Using information regarding only the number of entering firms, they can point-identify the player's payoff function.

[7]Our estimation strategy is similar to Bajari, Benkard and Levin (2007), Pakes (2010), and Pakes et al. (2011) in that they construct an inequality-based estimator resulting from necessary conditions of the equilibrium, respectively.

[8]If one could employ information about other pairwise-stable networks as well as the observed network, the identified set could be much sharper, though the computational burden would be extremely severe. For more details on this approach, see Beresteanu, Molchanov and Molinari (2012), Ciliberto and Tamer (2009) or Glichon and Henry (2011).

venture capitalists, measured by the ratio of the number of initial public offerings (IPO) to the number of total invested deals, is related to (endogenous) network characteristics such as degree and betweenness, in addition to exogenous characteristics of venture capitalists and markets. In order to incorporate unobserved heterogeneity, which affects both the formation of the network and the outcome, I allow for the shocks to the preference of the venture capitalists to be correlated with the unobserved shocks that affect their investment performance. I exploit my structural model of the strategic network formation to correct the selection bias in the outcome equation.

The idea of correcting the selection bias is similar to Heckman's (1990) two-step estimation, in which the correction term is computed from the first stage regression and the correction term is used as an input of the second stage regression. I compute the correction term using the estimates of the strategic network formation model, but the existence of multiple stable networks makes it difficult. More precisely, in order to correct the selection bias exactly, the researcher needs to know the equilibrium selection rule, which is not observed. Since I use only the necessary conditions of pairwise stability for estimating the network formation model, it is not possible to compute the correction term exactly. Instead of estimating the equilibrium selection rule, I use the necessary conditions for making the bounds of the correction term, which result in the moment inequality conditions for the outcome equation.[9]

The identification of the strategic network formation model is obtained by an exclusion restriction. The exclusion restriction I require takes advantage of pair-specific characteristics: I consider variables that influence only the payoff of a pair of firms, but do not influence the payoffs of any other combination of firms. Exploiting this exclusion restriction, I can consider a situation where only one pair of firms has the opportunity to form a tie (and any other combination of firms does not). Then, the strategic network formation model is reduced to the double-index model studied by Ichimura and Lee (1991), and the identification follows their results. Regarding the identification of the outcome equation, the other firms' characteristics provide a source of exogenous variation as in Sorensen (2007), since those characteristics are independent of a given venture capitalist's unobserved heterogeneity.[10]

[9] It is possible to regard the model in this paper as a sample selection model, in which a structural model is used for correcting the selection bias. Examples include Mazzeo (2002b) and Ellickson and Misra (2012). Main difference from my paper and these papers is that there exist multiple equilibria in the first stage model. Multiplicity of equilibria leads to indeterminancy of the reduced-form selection equation, and correcting selection bias is not straightforward.

[10] Using the characteristics of other firms as instruments ignores the fact that the observed network structure is determined in equilibrium. Because correctly specifying the selection equation is essential in the

My structural model exploits such exogenous variation to identify the causal effect of the network structure on investment performance.

I find that networks of venture capital firms tend to be homophilous with regard to asset size under management: They prefer being connected with venture capital firms of similar size. In terms of investment experience and industry expertise, venture capital networks are *anti*-homophilous. Hence, they are more likely to be linked with venture capital firms that have different investment experience or different industry expertise. Moreover, the estimation results of the outcome equation indicate that there exists a selection bias resulting from the positive correlation between the unobserved heterogeneity in the preference over networks and the unobserved heterogeneity in the outcome. If the endogeneity of the networks are not taken into account, then the effects of the network structure on investment performance are significantly *overestimated*.

Lastly, I conduct a counterfactual policy experiment to see what if the government makes direct investment by establishing a government-funded venture capital firm. Government interventions into the venture capital industry are prevalent, but what is a good model of government-sponsored venture capital firm is little explored. Specifically, I exogenously add a venture capitalist with given asset size and investment experience to each market, and examine how the equilibrium network structure changes.[11] I then measure the change of investment performance under the new equilibrium network structure. The estimates indicate that the new entrant tends to make the network denser even when it has small asset size and little investment experience. The increase in network density improve their investment performance, which can offset the negative effect of congestion.[12]

The rest of the paper proceeds as follows. First, a literature review is conducted in Subsection 1.1. In Section 2, some additional background information about the U.S. venture capital industry and the data is introduced. A strategic network formation model is presented

sample selection models, a simple 2SLS approach, which assumes a linear relationship between the endogenous network structure characteristics and the exogenous characteristics of the other firms, may not be desirable. That is because the equilibrium aspect of the network structure may lead the selection equation to be highly nonlinear. Moreover, existence of multiple equilibria makes how to specify the selection equation even harder. Our approach uses the structural model of the network formation that accounts for the equilibrium aspect of the network structure in order to deal with the selection effect. Moreover, we are able to conduct counterfactual policy experiments using the estimates obtained from our structural model.

[11] In the counterfactual experiment, I focus on relatively small markets with less than 7 venture capital firms because computational costs are too big for markets with more than 7 firms. In those markets, I find all pairwise stable networks.

[12] Addition of a new venture capital firm creates negative congestion effects because I find that the effect of one additional veture capital firms on investment performance is negative.

in Section 3, and the econometric specification, the estimation strategy, and the identification of the model are described in Section 4. Section 5 reports the results of the estimation and the counterfactual experiment. Finally, Section 6 concludes.

1.1 Related Literature

This paper is related to several strands of literature. First, there is a large body of literature studying the venture capital industry and the role of networks in the industry. Gompers (1995) and Lerner (1995) provide evidence of the monitoring role of venture capitalists. Hochberg (2011) studies how venture-capital backing affects corporate governance. Hochberg, Ljungqvist and Lu (2007) point out the importance of the network structure as one of the determinants of venture capitalists' investment performance. Hochberg, Ljungqvist and Lu (2010) and Hochberg, Mazzeo and McDevitt (2011) examine the relationship between network structure and market competition. Lindsey (2008) and Robinson and Stuart (2007) study how sharing the same venture capitalist as a source of capital makes it easier to form a strategic alliance between companies in which the venture capitalist invests. These papers treat the network as given, while Hochberg, Lindsey, and Westerfield (2011) examine how venture capitalists choose their syndication partner(s) under the assumption that a link is formed if both of firms involved in the link prefer forming the link *regardless of whether other links are formed or not.*[13] They find that venture capitalists form ties in an anti-homophilous fashion in terms of their experience. My paper complements these studies. First, my network formation model concerns a strategic situation, in which I allow the existence of externalities that may accrue venture capitalists beyond each link. Second, my paper addresses the potential issue of the endogeneity of the network structure by jointly estimating a strategic network formation model and an outcome equation.

My paper is also related to the literature on estimating coorperative game models, such as two-sided matching models and coalition formation models.[14] Sorensen (2007) examines a non-transferable utility two-sided matching model between venture capitalists and start-up companies.[15] His paper is close to my paper in that he jointly estimates the matching

[13] In other words, the value of a link is a function of only characteristics of two firms involved, and a link is formed if the value of the link is greater than 0. Under that assumption, the observed network is always efficient, because a given link does not create any externalities for the decisions of other links.

[14] Cooperative games provide useful tools for investigating situations when a Nash equilibrium may not be able to characterize interactions among players very well or when institutional details for writing a specific non-cooperative game may not be observable. Typically, given a cooperative game, there exists a non-cooperative game that induces exactly the same outcomes in equilibrium.

[15] Under the assumption that all agents' preferences are aligned, Sorensen (2007) shows that there is unique

model and an outcome equation, and finds that the sorting effect determined through the equilibrium matching process has significant impact on the entrepreneur's success. In contrast to his paper, I consider the network formation of venture capitalists rather than the matching between venture capitalists and start-up companies. By doing so, I can examine the effects of the externalities among venture capital firms that bind them beyond each investment deal (each matching). Uetake and Watanabe (2012a) propose a new estimation strategy of two-sided matching models using a fixed point characterization of the set of stable matchings. In a separate paper by Uetake and Watanabe (2012b), we consider a non-transferable utility two-sided matching model with externalities by combining a static entry model and a two-sided matching model with contracts. In that paper, we use data on the U.S. commercial banking industry right after the deregulation about de novo branching restrictions, and investigate how entry costs vary for de novo entry and for entry by merger, and where merger synergies or dissynergies come from. My identification strategy of the strategic network formation model is similar to theirs. Fox (2010a and 2010b) study estimation and identification of two-sided matching models with *transferable utility*, in which players are freely able to make side payments and the utilities of the players are comparable in monetary terms. In contrast to a transferable utility model, my model of the strategic network formation concerns a non-transferable utility model. I consider a non-transferable utility model because venture capitalists do not make any side payments among them when they invest, and non-monetary benefits of forming ties such as accumulating experience are important for tie-formation decisions. Finally, Gordon and Knight (2009) and Weese (2011) study a coalition–formation model, respectively. As far as I know, this paper is one of the first papers that studies a strategic network formation model by applying the pairwise stability concept to real data. I contribute to the literature by providing a way to estimating these types of cooperative games.

Finally, there are only a few papers that structurally estimate a model of strategic network formation.[16] Christakis, Fowler, Imbens, and Kalianaraman (2010) examine a dynamic

stable matching. Hence, identification and estimation are conceptually straightforward, though computationally difficult.

[16] Random graph models have been extensively used in sociology and economics, where the probability of forming a link is related to various network characteristics. Among many others, examples include Geyer and Thompson (1992), Snijders (2002), Jackson and Rogers (2007) and Chandrasekhar and Jackson (2012). Geyer and Thompson (1992) and Snijders (2002) propose a Markov Chain Monte Carlo estimation of exponential random graph models. Christakis et al. (2010) and Mele (2011) use a similar algorithm to estimate their strategic network formation model. Jackson and Rogers (2007) consider a dynamic network formation model and calibrate the model to study the degree of randomness in the network formation of co-author relationship networks. As an example of a different type of network formation model can be found in Curraini, Jackson

model of strategic network formation. In their model, a pair of agents is randomly selected in each period to create a new link (or to delete an existing link). When each of agents makes a decision, the authors assume that agents behave myopically. This assumption allows the authors to bypass the issue of multiple equilibria.[17] They then estimate the model using a Bayesian MCMC approach. Goldsmith-Pinkham and Imbens (2011) study a linear-in-means model as in Manski (1993) and point out the possible endogeneity issue associated with the network formation. They address this point by jointly estimating the linear-in-means model and a simple version of Christakis et al.'s (2010) model. Mele (2011) also considers a dynamic process of strategic directed network formation that converges to a unique stationary equilibrium. My paper is different from these papers in that I consider a static model of strategic network formation with pairwise stability as the solution concept. Moreover, I do not impose any equilibrium selection rule, and apply the partial identification approach to estimate the model with multiple equilibria.[18] I also discuss how to identify the model, which is not explicitly examined in the existing papers. Finally, all of the papers above use the Add Health data set and study high school friendship networks, while my study focuses on venture capital investment networks, which exhibit an entirely different structure.

2 Industry Background and Data

2.1 Venture Capital Industry

Venture capitalists are professional and institutional investors who fund early-stage, high-potential, and also high-risk entrepreneurial start-up companies, which usually have a novel technology or business model in high-technology industries, such as biotechnology, IT, software, etc. Venture capital is attractive for new companies with a limited operating history that are too small to raise capital in public markets and have not reached the point where

and Pin (2009), who calibrate a search theoretic model that fits their findings of a homophily structure in high school friendship networks.

[17] In Christakis et al. (2010) and Goldsmith-Pinkham and Imbens (2011), and Mele (2011), there are no multiple equilibria at all because, in every period, randomly selected agents make the best response, which is generically unique, given the current network structure. Hence, the dynamic process uniquely determines the network structure if we are given a matching process. Mele (2011) goes further to show that such best-response dynamics converges to a unique stationary distribution under assumptions on the matching process and players' preferences.

[18] Mele (2011) assumes symmetricity on part of the payoff function in order for the model to have a potential function, which is necessary for his estimation strategy. Our estimation strategy does not require such an assumption and hence a more flexible form of the payoff function can be specified.

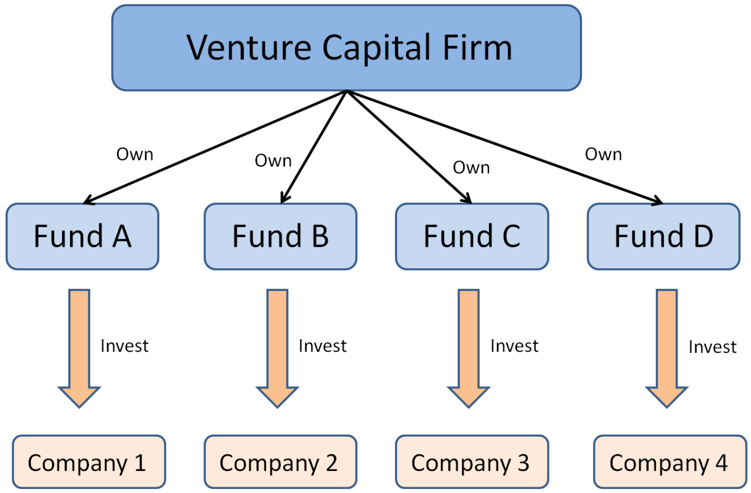

Figure 1: Example of the Relationships Between a Venture Capital Firm and Venture Capital Funds.

they are able to secure a bank loan or complete a debt offering.

Structure of Venture Capital Firms and Funds The typical structure of the venture capital firm is illustrated in Figure 1.

Note first that it is important to distinguish a venture capital *fund* from a venture capital *firm*. A venture capital fund, which is managed by a venture capital firm, invests in start-up companies, and a venture capital firm oftentimes own multiple funds under the management at the same time.[19] My primary focus in this paper is venture capital firms rather than venture capital funds. Finally, I define "company" as a start-up company in which a venture capital fund invests.

The role of venture capitalists is not only to find talented start-up companies and provide capital to those companies, but also to monitor and nurture the companies in which they invest. Because typical venture capital funds own equity in the companies, they have significant control over the companies (see, e.g., Hellman and Puri, 2002). Given the fact that the monitoring and nurturing roles of venture capitalists are essential for start-up companies to be successful, venture capital investments are typically long-term investments. In Figure 2, I draw the typical life cycle of a venture capital fund. Most venture capital funds have a fixed

[19] We abstract away the possibility of a syndication investment in Figure 1. As we will show in Figure 2, a venture capital fund often creates a syndicate with other venture capital funds to invest in an entrepreneur.

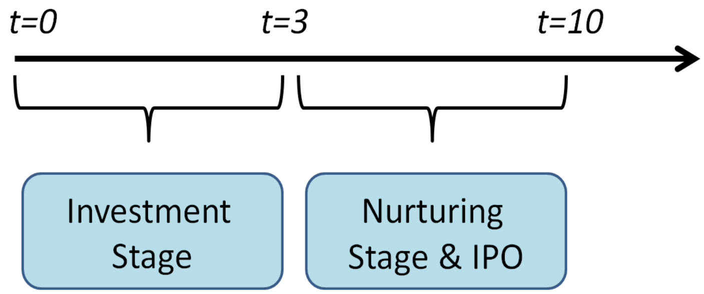

Figure 2: A Typical Life Cycle of a Venture Capital Fund.

life of 10 years.[20] The investing cycle for most funds is generally about three years, after which the focus is on nurturing the companies. The main goal of venture capital investments is going public (i.e., IPO) or being acquired by a big company in the industry (referred to as a "successful exit"). Because of the high-risk nature of venture capital investments, only a small fraction of portfolio companies are able to exit successfully.[21]

Although the life span of a typical venture capital fund is fixed at 10 years, including both the investment stage and the nurturing stage, a venture capital firm continues to be active longer. Kaplan and Shoar (2005) find that the performance of venture capital funds under the same management firm is persistent over the years. Hence, the knowledge of a particular industry or investment experience is likely to be carried over from one venture capital fund to another fund under the same venture capital firm.

Network of Venture Capital Firms Based on a Syndication Investment A remarkable feature of venture capital investments is that most financing involves a syndicate of two or more venture capital funds managed by different venture capital firms. Figure 3 depicts the typical structure of a syndicate that involves four venture capital funds managed by four different venture capital firms. Given a well-known fact in the industry that start-up companies searching for available capital seek it from a venture capital *firm*, rather than a venture capital *fund*, it is natural to regard a syndicate as a collection of different venture capital *firms*.[22]

[20] There is the possibility of a few years of extensions to allow for seeking better options of IPO or M&A.

[21] The average successful exit rate including both IPO and M&A is about 30%.

[22] The actual syndicate formation process is as follows. First, a venture capital fund that finds and contacts to an entrepreneur is called a lead investor. The process of the formation of a syndicate starts with an evaluation of the company by the lead investor. The lead investor then asks another venture capital firm for providing their opinion. If other venture capital firms agree with the project, they participate in the

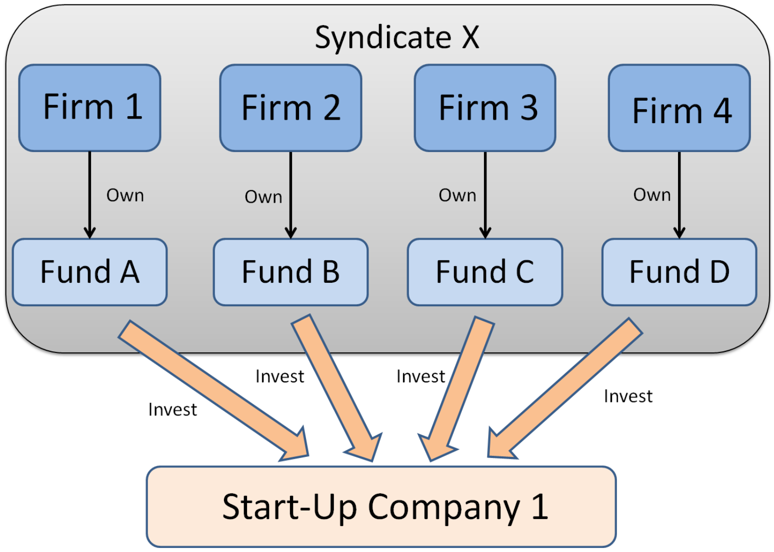

Figure 3: Example of a Syndicate with Four Different Venture Capital Firms.

There are several possible rationales for venture capital firms to make a syndicate. First, the syndication simply provides more capital for cash needs throughout the life span of the portfolio company (Lerner, 1994). Second, syndicating with other venture capital firms works as a double check on decisions. If the return from a project is uncertain, opinions of other venture capital firms who have different industry expertise or investment experience might help with the decision whether to invest (Sah and Stiglitz, 1986; Casamatta and Haritchabalet, 2007). Third, a reason to syndicate is obtaining access to other VCs' deal flow on a reciprocal basis (Bygrave, 1987; Lerner, 1994).[23] Hence, it is natural to investigate, as the starting point of my study, a syndication-based network rather than other types of networks, such as fund managers' networks through board connections.

As I show in Figure 3, a venture capital firm owns multiple venture capital funds at the same time. I consider networks of venture capital firms based on syndicates: in particular, there is a link between two venture capital firms if and only if two venture capital firms join the same syndicate (at least once). In Figure 4, I show an example of a network with four venture capital firms. In this example, firms A and B jointly invest in company 1 and hence there is a link between those two firms. On the other hand, firms A and D are not directly linked, but are linked indirectly through a link between firms A and C and a link between

syndicate.

[23]Other reasons include syndication might work as a device of risk-sharing, ability to draw on the expertise of other venture capitali firms when nurturing start-up companies, etc.

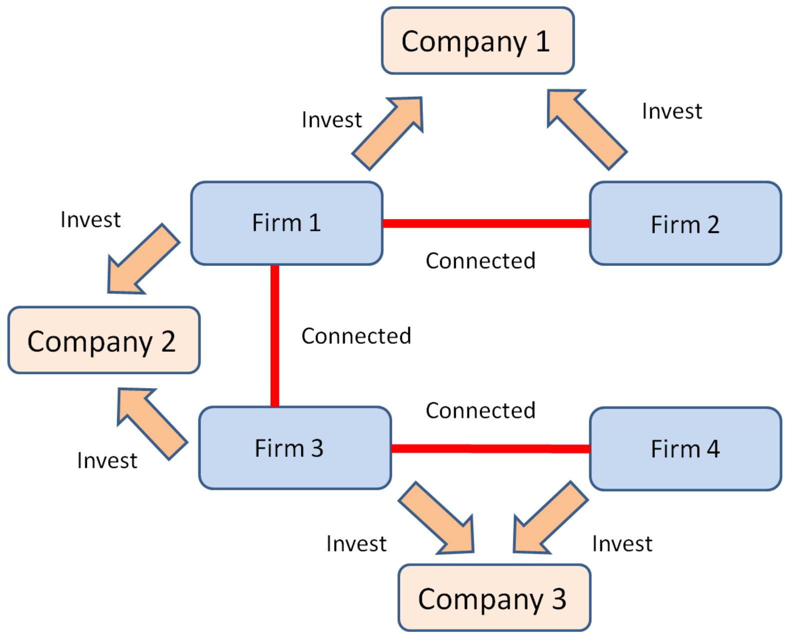

Figure 4: Example of a Network of Venture Capital Firms Based on Syndication Investments.

firms C and D. Note that I use the information of start-up companies only for creating networks of venture capital firms, and the start-up companies are not active players in my model. Thus, I can map the observations of many syndicate investments onto the network of venture capital firms (i.e., red lines in Figure 4).

2.2 Data

The main data I use for the estimation come from Thomson Financial's VentureXpert database. The VentureXpert database covers most of the data on venture capital investments from the early 1960s to today. As Gompers and Lerner (1999) find, the database contains most venture capital investments and the missing investments tend to be the less-significant ones. Hence, I treat the data as the full sample of investments rather than sampled investments.[24] This database contains rich information regarding venture capital firms, funds and investment deals. For more details, see, e.g., Gompers and Lerner (1999). My original sample includes all investments made between 1980 and 2005. I do not use the data before

[24] Chandrasekhar and Lewis (2011) study a potential bias that stems from using sampled networks for studying network effects. We do not think such bias is problematic in our sample.

1980 because there was a significantly smaller number of observations before 1980.[25] I also do not use the data after 2005 as many of the investments made since then have not yet gone public. Note that I use the information on exit events (i.e., IPO) through August 2012, so most of the investments made before 2005 have gone through the exit stage. As in Hochberg, Ljungvist and Lu (2007), I drop all investments made by non-US-based venture capital funds and by angels/buyout funds.[26]

In this paper, I consider the network formation in each geographic area. As demonstrated by Sorenson and Stuart (2000), venture capital firms tend to invest in geographically proximate companies and venture capital markets are known to be local in nature. This is because the nature of the relationships between start-up companies and venture capital firms includes establishing frequent personal contacts, due diligence, research, and the monitoring of portfolio companies. Hence, a venture capital firm/fund in the Chicago area, for example, is significantly more likely to invest in companies in the Chicago area than to ones in the San Francisco area.[27] In particular, following Hochberg, Mazzeo and McDevitt (2011), I define a local geographic market based on the Metropolitan Statistical Area (MSA), where each venture capital firm is operating.

In addition to the geographic area, I also consider a specific time span in which venture capital firms play a strategic network formation game. Given that the investment cycle of venture capital funds lasts about three years, I define this time span as a five-year window, and I do not consider any dynamic interactions among venture capital firms within that window of time and across time periods.[28] In summary, I define a "market" as an MSA in

[25] This is because there were several important policy changes that affected the institutionalization of the venture capital industry before 1980. First, before 1978, the Employee Retirement Income Security Act of 1974 (ERISA) restricted corporate pension funds from holding certain risky investments, including many investments in privately held companies. In 1978, the US Labor Department relaxed certain of the ERISA restrictions, under the "prudent man rule," thus allowing corporate pension funds to invest in private equity, resulting in a major source of capital becoming available for investing in venture capital and other private equity. Second, the Economic Recovery Tax Act of 1981 (ERTA) lowered the top capital gains tax rate, which made high-risk investments even more attractive.

[26] We also drop "markets," which we define below, with more than 25 venture capital firms for the sake of computational time. The number of markets dropped based on this criterion is about 50. Moreover, we do not use any monopoly markets because we cannot study any aspects of the networks from those markets. Those markets are typically very small and the number of such markets is about 180. Our estimation strategy is still feasible even if we drop markets with fewer than 75 venture capital firms. We will estimate the model with those data for a robustness check.

[27] It is possible for venture capital firms to invest in companies in other states, but companies located in the same MSA have priority.

[28] We also consider alternative definitions of a time window as follows: i) the first three years of each five-year window (i.e., 1980-82, 85-87, ...), ii) the first five years of each ten-year window (i.e., 1980-84, 1990-94, 2000-2004), and iii) the first and last ten years in our sample (1980-1989, 1996-2005). In the main text, we

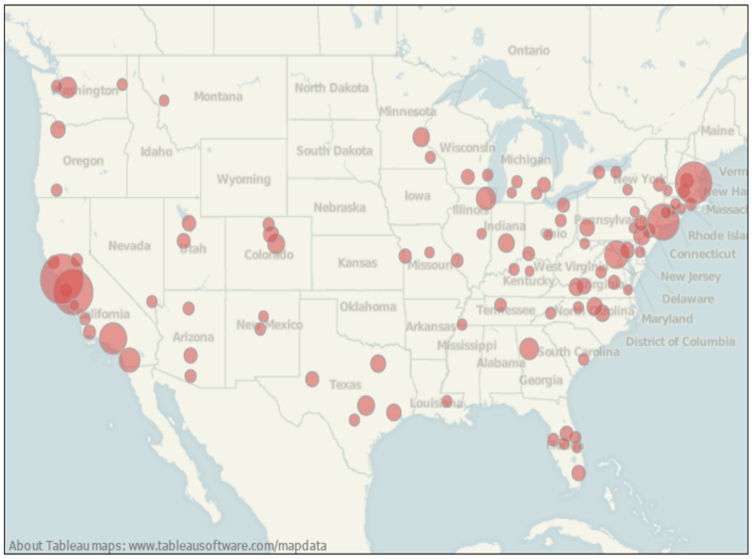

Figure 5: Map of Venture Capital Markets Across the United States. Each circle basically illustrates the size of each market (MSA).

a five-year window, and assume that network formation occurs in each "market."

To see how venture capital markets are distributed across states, in Figure 5, I plot the size of the venture capital markets (MSAs) on the US map.[29] As is clear in the figure, a considerable amount of investments takes place in the Silicon Valley, Route 128 in Boston, and New York areas, while the venture capital markets are spread across the states, and there are many regional markets. My study focuses on these local markets rather than larger markets.

I now explain how I construct a network matrix for each market from the data. In each market m defined above (MSA$\times$5 year), I compute an adjacency matrix g_m, in which each cell, say (i, j)-element, represents the existence of a link between venture capital firms i and j. I assume there exists a link between firms i and j if and only if both firms i and j participate in at least one same syndicate in a given time period.[30]

focus on the original definition, but in the Appendix, we also show the estimates under alternative market definitions, and find that the results are qualitatively similar.

[29] We use software from http://blogs.wsj.com/venturecapital/2011/08/04/interactive-map-the-united-states-of-venture-capital/tab/interactive/. The data for creating this table is from the first half of 2011.

[30] We do not distinguish whether venture capital firms i and j participate in more than one same syndicates from whether they participate in just one syndicate. In other words, we do not consider the strength of links defined by the number of portfolio companies in which they co-invest. Including the strength of links to the model would be a fruitful future research topic.

Some remarks are in order regarding my definition of the networks among venture capital firms. First, I focus in this paper on the networks between venture capital firms that aggregate all syndicates, and do not investigate syndicate formation in deal by deal. As pointed out by Hochberg et al. (2011), this approach allows us to investigate potential externalities that influence venture capital firms beyond the level of each investment-deal-level syndicate formation. With respect to this point, note that my definition of a link is based on venture capital *firms* rather than venture capital *funds*. I could alternatively consider fund-based networks,[31] but that would be much more complicated to analyze because a firm would simultaneously own multiple funds and hence a firm's decision would have to involve all the possible links for all the funds under management.

Second, note that I define a link when firms i and j co-invest in a company in whichever the start-up company is located. In other words, if firms i and j in market m co-invest in a portfolio company in market $m' \neq m$, I assume there is a link between i and j in market m instead of market m'. Given that substantial evidence of venture capital firms' preferences to invest in their own geographic market m (Sorenson and Stuart, 2000; Hochberg and Rauh, 2012), it is not unreasonable to define a link this way.

Third, I do not specify the stage in which firms i and j are co-investors. It is possible that some venture capital firms are more specialized in the nurturing stage than the initial investment stage. Hence, it might be important for venture capital firms to decide when to join a syndicate. Studying such dynamic interactions is left for future research.

Summary Statistics I start my discussion about the variables I will use in the estimation with market-related characteristics. I control for market-level characteristics because market-level demand and cost shifters may affect venture capital firms' tie-forming decisions. In particular, I use the MSA-level population, the MSA-level gross product (hereafter, GDP), and the number of venture capital firms in a market. The MSA-level information on market-level characteristics comes from the Bureau of Economic Analysis.[32]

In Table 1, I show the summary statistics of market-level characteristics. In total, there are 328 markets and there exists substantial variation in each variable across markets. Because I drop the very large MSAs (e.g., San Francisco and Boston), each market-level variable is relatively small compared to the full sample. The average number of venture capital firms

[31] For example, Bhagwat (2011) studies the effect of fund-level networks based on the fund managers' educational relationship on the funding possibility of venture capitalists.

[32] MSA-level R&D expenditure may be an alternative market characteristic, but it is highly correlated with MSA-level GDP, so we use GDP rather than R&D expenditure for the current specification.

	Mean	Std	Min	Max	Obs
Population (Thousand)	2,710	4,383	80	18,400	328
Per Capita Income	24,183	8,856	8,611	68,036	328
MSA GDP ($\$1M$)	266	268	11	1,689	328
# of Firms	6.750	5.543	2	25	328

Table 1: Descriptive Statistics—Market Level

in a market is about seven. This implies that the number of possible links is on average $21 = (7 \times 6 \div 2)$, and the number of possible network configurations is 2^{21} , which is greater than 2 million. Thus, the computational burden of solving the model is still problematic.

Next, I discuss the characteristics of venture capital firms. Table 2 summarizes the firm-level variables. Using recorded information from the VentureXpert database, I compute the size of each venture capital firm and investment experience. As a measure of firm size, I use assets under management, which are total dollars invested by all of currently active funds under management. With respect to the measure of the investment experience of a venture capital firm, I define the overall investment experience by the number of days that have passed since the venture capital firm's very first investment, which is recorded in the VentureXpert database even if the initial investment of the venture capital firm occurred before 1980.

In the first and second rows of the table, I show summary statistics of firm size and investment experience. The average of overall firm size is $\$33M$, but its standard deviation is \$99.5, implying that there is significant variation across firms. The average total experience is about five years, and the standard deviation of it is about eight years. Hence, venture capital firms are very heterogeneous in terms of both asset size and investment experience.

In addition to the venture capital firms' characteristics, I also need a measure of their investment performance. As Hochberg, Ljungvist and Lu (2007) note, it is not straightforward to quantify the performance of venture capital firms. This is because private equity such as venture capital is largely exempt from public disclosure requirements. Moreover, portfolio companies are not yet publicly traded and thus it is difficult to compute the rate of return, which is typically used as the measure of the investment performance of financial institutions. Instead of the rate of return, I define a proxy for the performance of a venture capital firm using the Initial Public Offering (IPO) rate, which is the number of portfolio companies

Category	Mean	Std	Min	Max	Obs
Firm Size ($\$1M$)	33.0	99.5	0	$2,173.6$	$2,214$
Experience (Years)	5.04	7.46	0	40.70	$2,214$
IPO Rate	0.183	0.3	0	1	$2,214$
Distance	15.9	50.7	0	167.2	$2,214$

Table 2: Descriptive Statistics — Venture Capital Firm Level

	Mean	Std	Min	Max	Obs
Degree	0.463	0.642	0	4.2	328
Betweeness	0.572	1.641	0	15.3	328
Diameter	1.420	1.280	0	6	328

Table 3: Descriptive Statistics—Network Statistics

that went public through August 2012 divided by the total number of portfolio companies in which the venture capital firm invested within a particular time period. For example, if venture capital firm i invested in three start-up companies between 1980 and 1985 and only one of them went public before the August of 2012, firm i's IPO rate is 0.3. When the IPO rate is higher, the venture capital firm tends to obtain a greater return from its investment.[33] The average IPO rate is about 18%, implying that venture capital investments are high risk.

In the bottom row of Table 2, I show the summary statistics of the distance between the headquarters of venture capital firms i and j using the zip code of each venture capital firm. I use the distance to investigate how the physical distance affects the likelihood of forming a link.

Finally, I show some measures of network characteristics in Table 3. I report the summary statistics of the average degree, betweenness centrality, and diameter of venture capital networks. The degree of a node is a measure for how well the node is connected, defined by the number of direct links a given node is a part of the links. The betweenness centrality

[33] As Gompers and Lerner (2000) point out, there is another possible variable for the measure of venture capital firms' investment performance: Merger and Acquisition. They compare the IPO rate to Merger and Acquisition and find that both measures provide similar variation. In the future version of the paper, I will also use detailed M&A information to constructt the outcome measure.

captures how important each node is in terms of connecting other nodes, which is defined by the sum of the fractions of all the shortest paths in the network that contain a given node. To arrive at the figures in the table, I first compute each measure for each node within a market, take the average across nodes in each market, and finally take average of each measure across markets. The average degree and betweenness are 0.46 and 0.57, respectively. The diameter of a network is defined as the largest distance (number of links) between any two nodes in the network. The average diameter is 1.4. Overall, these summary statistics of the network characteristics show that the networks in my sample are relatively sparse.

3 A Model of Strategic Network Formation

My model of strategic network formation is based on Jackson and Wolinsky (1996). Because the venture capitalists' decision about syndicate formation is made after the intensive review of funding applications by start-up companies, it is natural to use a *strategic* model rather than a *random* model for network formation. After describing the model, I introduce the solution concept: *pairwise stability.*

3.1 Setup

The strategic network formation model of this paper is a game of complete information based on Jackson and Wolinsky (1996). Let $I = \{1, 2, ..., N\}$ be the set of venture capital firms. A network g is represented by a $N \times N$ adjacent matrix whose (i, j)-element, denoted by g_{ij}, is referred to as the link between firm i and j, which takes the value of 1 if firms i and j are linked, and 0 otherwise. Note that links are undirected, i.e., $g_{ij} = g_{ji}$. It is natural to describe venture capital networks as undirected networks because venture capital firms must agree with the investment deal. In other words, if firm i is a partner of firm j, then it is not possible that firm j is not a partner of firm i. Note also that $g_{ii} = 0$ by construction. I denote $g + ij$ as the matrix replacing $g_{ij} = 0$ by $g_{ij} = 1$ if i and j are not linked in g, and denote $g - ij$ as the matrix replacing $g_{ij} = 1$ by $g_{ij} = 0$ if i and j are linked. I also denote $ij \in g$ if firms i and j are directly connected in network g. Let us also define G as the set of all possible network configurations. Lastly, I denote the set of existing links in network g by $C(g) = \{(i, j) \in I \times I : g_{ij} = 1\}$ and the set of non-existing links in network g by $D(g) = \{(i, j) \in I \times I : g_{ij} = 0, i \neq j\}$.

The payoff of firm i depends not only on the firms with which firm i has a direct link, but also on the firms that are indirectly connected to firm i. I specify the payoff function that the firms obtain from network g as $U_i(g) : G \to \mathbb{R}$, or

$$\begin{aligned} U_i(g) &= u_i(g) + \varepsilon_i(g) \\ &= \sum_{j:d(i,j;g)\leq \overline{d}} \delta^{d(i,j;g)-1} u(x_i, x_j, x_{ij}, z; \theta) + \varepsilon_i(g), \end{aligned}$$

where $u_i(g)$ is the deterministic part of the payoff, which is the discounted sum of pair-specific payoffs, $u(x_i, x_j, x_{ij}, z; \theta)$. I normalize the payoff from a null network is 0, i.e., $U_i(\{\phi\}) = 0$ for all i. The pair-specific payoff of firm i from being directly linked with firm j, $u(x_i, x_j, x_{ij}, z; \theta)$, is the function of firm i's characteristics, x_i, firm j's characteristics, x_j, pair-(i, j) specific variables, x_{ij}, market characteristics, z, and the parameter to be estimated, θ. I denote by $d(i, j; g)$ the length of the minimum path between firms i and j in network g. For example, when firm i and j are directly connected, $d(i, j; g) = 1$, and when firms i and j are only indirectly connected via another firm k, $d(i, j; g) = 2$. I also define $d(i, j; g) = \infty$ if firms i and j are neither directly nor indirectly connected in network g. $\delta \in [0, 1)$ is a discount factor that captures the idea that the payoff of firm i from being indirectly connected with firm j depreciates as the distance between i and j in network g increases. Hence, given the characteristics (x_i, x_j, x_{ij}, z), the payoff obtained from a direct partner is larger than that from indirectly connected partners. $\overline{d} \in \mathbb{N} \cup \{\infty\}$ is the maximum distance at which each firm can receive the payoff. If the distance between firm i and j becomes greater than $\overline{d}$, then the discounted payoff, $\delta^{d(i,j;g)-1} u(x_i, x_j, x_{ij}, z; \theta)$, becomes 0. I use this upper bound for the distance only for computational reasons. Lastly, $\varepsilon_i(g)$ is a stochastic payoff shock to firm i when network g is realized.[34] The dependence of $\varepsilon_i(g)$ on the network structure of g rather than just the direct links that firm i is part of links captures unobserved payoff shock from indirect links.[35] Because I consider a complete-information game, $u(x_i, x_j, x_{ij}, z; \theta)$ and $\varepsilon_i(g)$ are common knowledge to firms, while not observable to an econometrician. For notational simplicity, I denote the vector of the exogenous characteristics of all the firms and the market by $\mathbf{X} = \{(x_i, x_j, x_{ij}, z)\}_{(i,j)\in I\times I}$.

[34] The dependence of ε_i on $g \in G$ is similar to discrete choice models in which g corresponds to a paritcular choice and G corresponds to the entire choice set.

[35] To understand the dependency of ε_i^g on g, let's consider the following simple example. Suppose there are three players, A, B, and C. If only A and B are connected, then each player receives a shock, and if A and B, and B and C are connected, each player receives different shock, respectively. Later, we will assume that these shocks are i.i.d. for simplicity. Thus, other players' decisions only affect whether each player receives a particular shock or not, but ε_i^g itself is indepently drawn from an identical distribution, and not a function of other players' decisions.

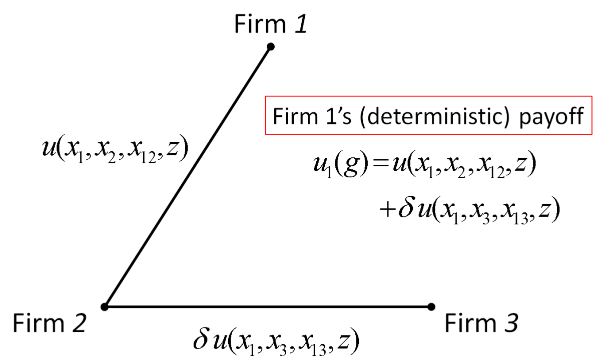

Figure 6: Example of the Deterministic Part of the Payoff, $u_i(g)$.

In Figure 6, I illustrate an example of the structure of the deterministic part of the payoff, $u_i(g)$. In the network in Figure 6, firms 1 and 2 are connected, and firms 2 and 3 are connected, but firms 1 and 3 are not connected. As for firm 1's deterministic-part payoff, firm 1 receives a direct payoff from its connection to firm 2, i.e., $u(x_1, x_2, x_{12}, z)$. In addition to this direct payoff, firm 1 obtains an additional payoff from its indirect connection to firm 3 through the link between firms 2 and 3. Hence, $d(1, 3; g) = 2$, and such indirect payoff is $\delta u(x_1, x_3, x_{13}, z)$. The overall payoff from the deterministic part is then $u(x_1, x_2, x_{12}, z) + \delta u(x_1, x_3, x_{13}, z)$.

Several remarks are in order regarding the specification of the payoff. First, I assume that the deterministic part of the payoff is additively separable by each pair (i, j). The additive-separability assumption is a popular one in the empirical network formation literature, e.g., Mele (2011). Second, note that each firm receives only one payoff shock, $\varepsilon_i(g)$, in my specification. An alternative model might have pair-by-pair additively separable payoff shocks, i.e., $\varepsilon_i(g) = \sum_{j:d(i,j:g)\leq\bar{d}} \delta^{d(i,j:g)-1}\varepsilon_{ij}$. I can interpret such an error term ε_{ij} as pair-level unobserved heterogeneity. Although I am able to have such an error term structure in general, I focus on the original specification for computational simplicity. I will come back to this point in the next section. Third, my network formation model is a non-transferable utility model because venture capital firms do not make any transfers among them when they form a syndicate, and also non-monetary payoffs such as the manager's taste are important motivations for forming a link. Finally, in my model of strategic network formation of venture capital firms, I assume away matching processes between venture capital funds and start-up companies, which occur before the network-formation stage starts. I do so because i) incorporating a model of matching into a strategic network formation model makes the problem even more difficult and the estimation of such a model practically infeasible, and ii) only a few exogenous variables are observable for the portfolio companies, e.g., industry group and investment stage.[36]

3.2 Pairwise Stability

The solution concept I employ is *pairwise stability* proposed by Jackson and Wolinsky (1996), which is extensively used in the literature. The definition of pairwise stability captures the idea that it is necessary for both players to agree to forming a link, while it is possible for each of them to sever the link. This bilateral aspect of the decision to form a link is a key to understanding the determinants of strategic network formation in the venture capital industry. I now define pairwise stability.

Definition 1 (Jackson and Wolinsky (1996)) *A network g is pairwise stable if $\forall ij \in C(g)$, $U_i(g) \geq U_i(g - ij)$ and $U_j(g) \geq U_j(g - ij)$, and $\forall ij \in D(g)$, if $U_i(g + ij) > U_i(g)$, then $U_j(g + ij) < U_j(g)$.*

The first condition states that if one observes a link between two firms, then both firms are better off by forming the link than not forming. The second condition states that if one does not observe a link between two firms, then at least one of the two firms will be worse off by adding the link. In other words, it is not possible for both firms to be better off by adding a non-existing link. Given that investments through syndicates require the involved venture capital firms to agree with the investment deal, it is natural to use pairwise stability to characterize the network formation in the venture capital industry.

In order to understand how the solution concept of pairwise stability characterizes the network structure, it is helpful to compare pairwise stability with other solution concepts, such as Nash equilibrium and strong stability (see, e.g., Jackson and Nouweland, 2005). While a pairwise-stable network is robust to bilateral deviation, a Nash equilibrium is only robust to unilateral deviation and thus may not be able to capture the fact that the joint investment is a bilateral decision.[37] Moreover, when I consider Nash equilibrium as my model's

[36] Sorensen (2007) focuses on this stage and estimates a two-sided matching model.

[37] As in Mele (2011), it is reasonable to consider Nash equilibrium for modeling friendship networks. This is because it is possible in a friendship network for person A to call person B a friend, but for person B not to call person A a friend. In such a network, unilateral deviation is sufficient to characterize the networks.

solution concept, there are generally too many Nash equilibria in the strategic network formation game. Hence, it may not be helpful to predict observed networks in the data using the model.[38]

In addition to Nash equilibrium, an alternative solution concept is *strong stability* (see, e.g., Jackson and Nouweland, 2005), which requires stronger conditions than the pairwise stability. While pairwise stability requires the network to be robust to deleting or adding *one* link at a time, strong stability requires the network to be robust to a change of more than one link at a time. Because the venture capital network is based on the joint investment behavior potentially containing more than one venture capital firms, strong stability may be a more appropriate solution concept. Finding a strong-stable network, however, is much more complicated than finding a pairwise-stable network because it is necessary to check all the possible subsets of all existing links in a given network. Hence, it is not practically possible for us to use strong stability for estimation. Moreover, because pairwise stability is a necessary condition for strong stability, the estimates I obtain under pairwise stability must include the estimates under strong stability. For these reasons, I use pairwise stability as the solution concept of the game.

Multiplicity of a Pairwise-Stable Network One of the difficulties of using a strategic network formation model for prediction is the multiplicity of pairwise-stable networks. In the following example, I show how a market can admit multiple stable networks. Suppose that $N = 3$ (firms A, B, and C) and the payoffs are given as in Figure 7, i.e., if $g = \{\phi\}$, then $U_i(g) = 0$ for all i; if $g \in \{AB, AC, BC\}$, then $U_i(g) = 0$ for all i; if $g \in \{\{AB, AC\}, \{AB, BC\}, \{AC, BC\}\}$, then $U_i(g) = -3$ for all i; and if $g = \{AB\}$, then $U_A(g) = U_B(g) = 1$ and $U_C(g) = 0$. Similarly, if $g = \{AC\}$ (or $\{BC\}$), then $U_A(g) = U_C(g) = 1$ and $U_B(g) = 0$ (and $U_B(g) = U_C(g) = 1$ and $U_A(g) = 0$, respectively). Then, it is easy to see that the full network, $\{AB, AC, BC\}$, is a pairwise-stable network because deleting one of the links makes all firms worse off ($0 \to -3$). Moreover, $\{AB\}$, $\{AC\}$, and $\{BC\}$ are also pairwise-stable networks. This is because, for example, deleting AB from $g = \{AB\}$ makes both A and B worse off ($1 \to 0$), and adding either $\{AC\}$ or $\{BC\}$ makes all the firms worse off ($1 \to -3$). Thus, the set of pairwise-stable networks is not singleton in general. My estimation strategy, however, depends only on the

[38] For example, Myerson (1991) proposes a simultaneous link announcement game, where each player simultaneously declares with whom she wants to form a link, and a link is established if both players prefer having a link. This game always admits the null network as a Nash equilibrium. For other network formation models using Nash equilibrium, see, e.g., Bala and Goyal (2000).

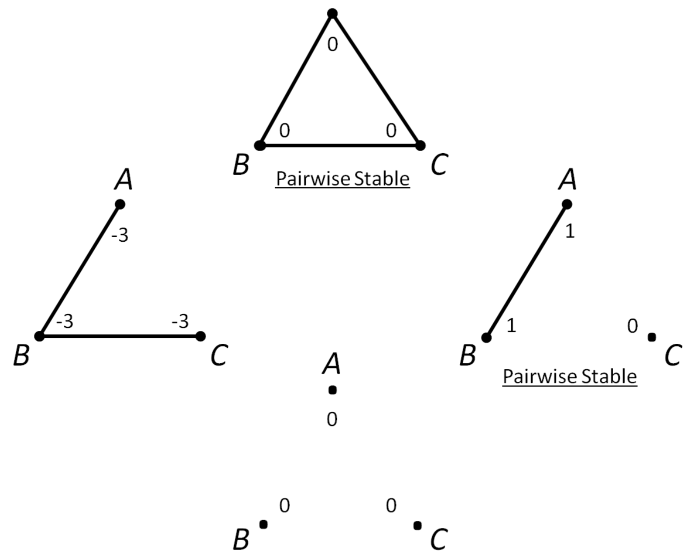

Figure 7: Multiplicity of a Pairwise-Stable Network.

necessary conditions of pairwise stability, thus being robust to multiple equilibria.

Existence of Pairwise-Stable Networks Another issue in studying strategic network formation models is the existence of pairwise-stable networks. Unfortunately, the existence of pairwise stable networks is not well understood in general. Although Jackson and Watts (2001), Chakrabarti and Gilles (2007), and Hellman (2012) provide sufficient conditions for the existence of pairwise-stable networks, the conditions they show may be restrictive from an empirical point of view.[39] Hence, when I discuss my estimation and identification strategies in the next section, I implicitly assume that there exists at least one stable network for any set of parameters, and I will confirm it by simulation once I estimate the model. In fact, as long as I conduct simulations, I found at least one pairwise stable network for any parameters in the confidence set.

[39] Jackson and Watts (2001) and Chakrabarti and Gilles (2007) show that the existence of a network potential function is sufficient for ruling out closed cycles, which in turn implies the existence of pairwise-stable networks. They also argue that if preferences are aligned or if a link between two players is beneficial either to both or to none, then a potential function exists. Hellman (2012) provides another set of sufficient conditions: ordinal convexity and ordinal strategic complements of utility functions. Imposing this symmetricity or complementarity condition to utility functions a priori is in general too strong from an empirical point of view.

4 Estimation and Identification

In this section, I discuss my estimation and identification strategies for the model, which involves both the strategic network formation and the outcome equation. The outcome equation relates the investment performance of venture capital firms with their exogenous characteristics and network characteristics. First, I focus on the estimation of the strategic network formation model. My estimation strategy relies only on the necessary conditions of pairwise stability and hence is robust to the multiplicity of pairwise-stable networks. Moreover, my estimation strategy is computationally feasible due to the fact that I construct moment conditions from observed networks and that pairwise stability is robust to one-link deviation. Second, I discuss joint estimation of the strategic network formation model and the outcome equation. In particular, I allow a shock to the firm's payoff to be correlated with a shock to the outcome equation in order to capture the idea that both performance and network formation may be influenced by firm-level unobserved heterogeneity. Third, I discuss the identification of my strategic network formation model, which is given by an exclusion restriction (I require a variable that affects only a pair of firms and not any other combination of firms) and the identification results of multiple index models as in Ichimura and Lee (1991). Finally, I also briefly discuss the identification of the outcome equation.

4.1 Estimation of the Strategic Network Formation Model

The estimation of strategic network formation models poses challenges. This is because i) there are typically multiple pairwise-stable networks and ii) the number of possible network configurations is too large. Regarding the first challenge, the issue of multiple equilibria often arises in estimating game theoretic models. If the researcher cannot tell which equilibrium corresponds to the data-generating process, it is not straightforward to match the data to the model. More precisely, in the model with multiple equilibria, the model generally becomes incomplete and point identification of the model may be lost as pointed out by Tamer (2003). Researchers therefore typically impose a certain equilibrium selection mechanism to restore the point identification. As for the second challenge, even if one can guarantee that the equilibrium is unique given an equilibrium selection mechanism, the number of possible network configurations is prohibitively large even with a relatively small network. To see this, suppose there are ten players in a market. Even in this small network, there exist $45(= 10 \times (10-1)/2)$ possible links, and hence the number of possible network configurations is $2^{45} \approx 3.5 \times 10^{13}$. Thus, fully solving the game (or finding all pairwise-stable networks) to evaluate the likelihood or the moment conditions is computationally infeasible.

In order to tackle these two challenges, I take advantage of "the partial-identification approach" instead of imposing any equilibrium selection mechanism.[40] In particular, I construct moment conditions derived only from the necessary conditions of pairwise stability, modifying the idea of Bajari, Benkard and Levin (2007), Pakes et al. (2011) and Pakes (2010) who propose a moment inequality estimator resulting from the necessary conditions of equilibrium. This estimation strategy allows us to bypass two major problems in estimating strategic network formation models: It does not depend on any assumption regarding equilibrium selection mechanism, and hence is robust to multiple equilibria. Moreover, it is not necessary to consider all possible-network configurations or to find other pairwise-stable networks, and thus it is computationally feasible.

I start the discussion of my estimation strategy by introducing an assumption on the payoff shock, $\varepsilon_i(g)$.

Assumption 1: The payoff shocks $\varepsilon(g) = \{\varepsilon_i(g)\}_{i \in I}$ are i.i.d. across $i \in I$ and $g \in G$. $\varepsilon_i(g)$ is drawn from a distribution function F_ε that is known up to a finite dimensional parameter. Also, the payoff shocks have mean 0. Moreover, $\varepsilon_i(g)$ and the firms' characteristics are independently distributed.

The first part of the assumption implies that $\varepsilon_i(g)$ and $\varepsilon_j(g)$ are independent, and $\varepsilon_i(g)$ and $\varepsilon_i(g')$ are also independent. The assumption seems strong, but this type of independence assumption has been extensively used in discrete choice models. I impose the first assumption only for computational simplicity, and as I will explain later I can allow error terms to be correlated across g to some extent. Allowing the more–flexible error-term structure may be computation feasible, but I adhere to the current structure to make the expression and computation as simple as possible, and use other specifications for robustness checks. The last part of the assumption concerns independence between the error terms and observable characteristics as in the empirical entry game literature, e.g., Bresnahan and Reiss (1991), Berry (1992) and Ciliberto and Tamer (2009).

Now I discuss the way I construct inequality conditions from the necessary conditions of pairwise stability. I explain it using the following simple example. Consider a small market with four firms–A, B, and C– and suppose the econometrician observes the following network in this market as in Figure 8. Note that my unit of observation is a market rather than a

[40] The partial-identification approach and moment inequality estimators are current areas of research in econometrics. For a general survey of this approach, see, e.g., Tamer (2010). Recent applications of this approach include Ho (2009), Ciliberto and Tamer (2009), and Uetake and Watanabe (2012).

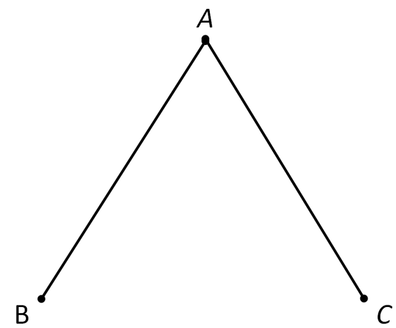

Figure 8: An Example of 3 Firms Networks

firm, and identity of each firm does not matter for my analysis. I use terms "firms A, B, and C" only for expositional simplicity.

Since I assume that the data-generating process corresponds to one of the pairwise-stable networks, deleting link AC, for example, reduces the payoffs of both firms A and C, i.e.,

$$\begin{array}{r} u_A(g)+\varepsilon_A(g) \geq u_A(g-AC)+\varepsilon_A(g-AC) \\ u_C(g)+\varepsilon_C(g) \geq u_C(g-AC)+\varepsilon_C(g-AC). \end{array} \quad \text{and} \tag{1}$$

Similarly, by deleting another existing links AB, I have two other inequalities as follows:

$$\begin{array}{r} u_A(g)+\varepsilon_A(g) \geq u_A(g-AB)+\varepsilon_A(g-AB) \\ u_B(g)+\varepsilon_B(g) \geq u_B(g-AB)+\varepsilon_B(g-AB). \end{array} \quad \text{and} \tag{2}$$

The two inequalities mean that deleting link AB reduces the payoffs of both firms A and B.

Now, add a link between firms B and C. Then, pairwise stability requires that either the payoff of firm B or C (or both) is decreased in the new network $g+BC$, i.e.,

$$\begin{array}{r} u_B(g)+\varepsilon_B(g) \geq u_B(g+BC)+\varepsilon_B(g+BC) \\ u_C(g)+\varepsilon_C(g) \geq u_C(g+BC)+\varepsilon_C(g+BC). \end{array} \quad \text{or} \tag{3}$$

In this case, the econometrician does not know which inequality is satisfied. Hence, I alternatively consider the following equivalent equality condition:

$$1\begin{bmatrix} u_B(g)+\varepsilon_B(g) < u_B(g+BC)+\varepsilon_B(g+BC) \;\; \text{and} \\ u_C(g)+\varepsilon_C(g) < u_C(g+BC)+\varepsilon_C(g+BC) \end{bmatrix}=0.$$

Since the unit of observation is the market (or the network) in my analysis, I take the summation of the inequalities (1) and (2) to get a market-level condition;

$$\sum_{(i,j)\in C(g)} u_i(g) + \varepsilon_i(g) \geq \sum_{(i,j)\in C(g)} u_i(g-ij) + \varepsilon_i(g-ij), \tag{4}$$

where $C(g)$ is the set of existing links (in Figure 8, $C(g) = \{\{AB\}, \{AC\}\}$). Here, I slightly abuse the notation because I add up all of four inequalities rather than two. Similarly, from the equality condition (3), I obtain

$$\sum_{(i,j)\in D(g)} 1\begin{bmatrix} u_i(g) + \varepsilon_i(g) < u_i(g+ij) + \varepsilon_i(g+ij) \ \text{ and} \\ u_j(g) + \varepsilon_j(g) < u_j(g+ij) + \varepsilon_j(g+ij) \end{bmatrix} = 0, \tag{5}$$

where $D(g)$ is the set of non-existing links ($D(g) = \{\{BC\}\}$ in Figure 8). Note that for the network in Figure 8, $D(g)$ is a singleton, but generally there exist multiple non-existing links.

Because we can construct inequalities similar to inequality (4) from any network (or market) by identifying the set of existing links and by summing all of inequalities up. Also, we can construct equalities similar to equality (5) from any network by identifying the set of non-existing links and by summing all of equalities up. Hence, the fact that network g^D is realized necessarily implies the following inequality;

$$E\left[\sum_{(i,j)\in C(g^D)} u_i(g^D) + \varepsilon_i(g^D)\,|\mathbf{X}\right] \geq E\left[\sum_{(i,j)\in C(g^D)} u_i(g^D-ij) + \varepsilon_i(g^D-ij)\,|\mathbf{X}\right]. \tag{6}$$

Note that this expectation is taken across markets.[41] Similarly, the fact that network g^D is observed necessarily implies the following moment condition from equality (5);

$$E\left[\sum_{(i,j)\in D(g^D)} 1\begin{bmatrix} u_i(g^D) + \varepsilon_i(g^D) < u_i(g^D+ij) + \varepsilon_i(g^D+ij) \ \text{ and} \\ u_j(g^D) + \varepsilon_j(g^D) < u_j(g^D+ij) + \varepsilon_j(g^D+ij) \end{bmatrix} |\mathbf{X}\right] = 0. \tag{7}$$

In order to use inequality (6) for estimation, I need to compute the sum of conditional expectations of $\varepsilon_i(g^D)$ and $\varepsilon_i(g^D - ij)$. I compute these conditional expectations by doing simulations for each market.[42] To do so, first note that error term $\varepsilon_i(g^D)$ associated to the

[41] In other words, from iid draws of networks, we can compute $\sum_{(i,j)\in C(g)} (\triangle u_i(g, ij) + \varepsilon_i(g) - \varepsilon_i(g-ij))$ for each network to take expectation of this object across networks.

[42] Aguirregabiria et al. (2012) use the similar way of constructing moment inequalities in their study of a

observed network g^D must support g^D as a pairwise stable network. It means that, for the network in Figure 8, $\varepsilon_A(g)$, $\varepsilon_B(g)$ and $\varepsilon_C(g)$ must satisfy the conditions (1), (2), and (3). To construct the sample analogue of inequality (6) in this network, I draw a lot of ε_i for each player, check the conditions (1), (2), and (3), and store the draws of $\varepsilon_i(g)$ satisfying them to evaluate inequality (6).

Second, regarding the error term associated to the unobserved network, $\varepsilon_i(g^D - ij)$, note first that if g^D is pairwise stable, then $g^D - ij$ is not pairwise stable. That is because under my specification of the error term, for any g and g', $U_i(g) \neq U_i(g')$ almost everywhere. Hence, for $\varepsilon_i(g^D - ij)$, I need to check the conditions such that $g^D - ij$ is *not* pairwise stable by doing the similar simulation procedure for each market as above.

Note that the region of ε_i characterized by the necessary conditions may support another network g' as a pairwise stable network, but inequality (6) and equality (7) must hold even if I use a set of $\varepsilon_i(g)$s that satisfies only the necessary conditions, because I rely on only necessary conditions for pairwise stability to derive these moment conditions.[43]

By doing this simulation procedure for all markets and by taking an average across markets, it is possible to construct the sample analogue of the inequality constraint (6) from the necessary conditions of pairwise stability. Note that all the terms in inequality (6) are now functions of only the observable characteristics and the parameters to be estimated.

As in the previous case, in order to use equality (7) for estimation, I need to check whether the necessary conditions for pairwise stability are satisfied for each observed network g^D for the error term associated to the observed network, $\varepsilon_i(g^D)$. Also, I check the conditions such that $g^D + ij$ is not pairwise stable for $\varepsilon_i(g^D + ij)$. By doing the similar simulation procedure for each market, I can construct the sample analogue of equality (7). Thus, I can construct equality conditions from the set of non-existing links.

General Case Recall that I construct inequality constraints from the set of existing links $ij \in C(g^D)$,

$$E\left[\sum_{(i,j)\in C(g^D)} u_i(g^D) - u_i(g^D - ij) + \varepsilon_i(g^D) - \varepsilon_i(g^D - ij)|\mathbf{X}\right] \geq 0,$$

large-scale discrete choice model.

[43] If it is possible to use sufficient conditions as well, the identified set can be smaller. However, checking sufficient conditions, or finding other pairwise stable networks are computationally too costly.

and equality constraints from the set of non-existing links $ij \in D(g^D)$,

$$E\left[\sum_{(i,j)\in D(g^D)} 1\begin{bmatrix} u_i(g^D) - u_i(g^D+ij) + \varepsilon_i(g^D) - \varepsilon_i(g^D+ij) < 0 \text{ and} \\ u_j(g^D) - u_j(g^D+ij) + \varepsilon_j(g^D) - \varepsilon_j(g^D+ij) < 0 \end{bmatrix} |\mathbf{X}\right] = 0.$$

To simplify the notation, I denote an element of the inequality condition by $h_{ij}(g^D)$ and an element of the equality condition by $f_{ij}(g^D)$. Recall that both h_{ij} and f_{ij} are functions of only observable characteristics and parameter θ. Using the vector of exogenous variables $\mathbf{X}$, I can define the moment conditions by

$$\begin{matrix} E\left[\sum_{(i,j)\in C(g^D)} h_{ij}(g^D;\theta) \times r(\mathbf{X})\right] \geq 0 \\ E\left[\sum_{(i,j)\in D(g^D)} f_{ij}(g^D;\theta) \times r(\mathbf{X})\right] = 0, \end{matrix} \quad \text{and}$$

where $r(\cdot)$ is any non-decreasing function to ensure that no inequalities are reversed by the interaction with $\mathbf{X}$.

Finally, I derive the sample analogue of the moment conditions as follows:

$$\frac{1}{M}\sum_m \sum_{(i,j)\in C(g_m)} h_{ij}(g_m) \times r(\mathbf{X}_m) \quad \geq \quad 0, \tag{8}$$

$$\frac{1}{M}\sum_m \sum_{(i,j)\in D(g_m)} f_{ij}(g_m) \times r(\mathbf{X}_m) \quad = \quad 0, \tag{9}$$

where M is the number of markets in the data and $\mathbf{X}_m$ is the vector of the exogenous variables in market m. The instruments must be independent of the error terms. Following a similar construction of instruments as in Ho (2009), I use indicator variables and interactions of indicators for several market and firm characteristics. For example, I create an indicator function for an exogenous variable z_m by $1\{z_m > med\{z_m\}\}$, where $med\{z_m\}$ is the median of all z_ms.

Estimation Procedure I describe the estimation procedure using the moment conditions defined above. The procedure takes advantage of simulation techniques to numerically evaluate $h_{ij}(g)$ and $f_{ij}(g)$ for observed network g for each market in the following way.

1. Fix parameter θ. For each market $m = 1, ..., M$, simulate necessary error terms from distribution F_ε (, which is known up to a parameter).

2. For each market m, compute $h_{ij}(g_m)$ and $f_{ij}(g_m)$ for each existing link and non-existing link, respectively, by using simulation draws to obtain the sample analogues of the moment conditions (8) and (9).

3. Use the moment inequality estimator by Andrews and Soares (2010).

The specific functions I exploit to construct the test statistics in Andrews and Soares (2010) are $S = S_1$ and $\varphi_j = \varphi_j^{(4)}$ with the number of bootstrapping $R = 1000$. Because it is not possible to report more than three dimensional confidence sets, I compute the min and max of the 95%-confidence set projected on each dimension and report the results in the next section (see Appendix 7.1 for details).

Some Remarks on the Estimation of the Strategic Network Formation Model I make a few remarks about the estimation at this point. First, note that this estimation procedure is computationally feasible unless I consider very large networks. Recall that, for each market m, I draw the error terms for each firm. Using these simulation draws, computing $h_{ij}(g)$ and $f_{ij}(g)$ requires me to check at most $2 \times |C(g)|$ inequalities and $|D(g)|$ equalities for observed network g, where $|C(g)| + |D(g)| = N_m(N_m - 1)/2$. Because the average number of venture capital firms in a market is about 7, the number of inequalities and equalities I need to check is at most 63,[44] which is significantly smaller than the number of possible network configurations, 2^{21}.

Second, i.i.d. assumption of $\varepsilon_i(g)$ can be relaxed to some extent. An alternative specification of the error term structure is that each firm i receives additively separable shocks from each firm that firm i is directly or indirectly connected. In Case 1, for example, firm A receives two shocks, ε_{AB} and ε_{AC}, instead of $\varepsilon_A(g)$. Once I specify the error term in this form, $\varepsilon_{AB} + \varepsilon_{AC}$ is no longer independent of ε_{AB} even if ε_{AB} and ε_{AC} are independent. However, what I need to do is to check inequalities link by link in computing $h_{ij}(g)$ and $f_{ij}(g)$.[45] The order of computation cost is still way smaller than an exponential rate, though I need to draw more error terms than the original specification.[46]

[44] It is necessary to check the conditions for each player, but it is possible to do that at once by using creating a large matrix. I use Matlab, which can operate matrix algebra quickly.

[45] Of course, if we specify the error term in the additively separable way, the expression of this conditional expectation will be a little bit different. However, we can simulate the conditional expectation corresponding to that case from necessary conditions.

[46] Given relatively small networks in my data, allowing this error term structure is not computationally infeasible.

Lastly, note that accommodating market-level unobserved shock to the payoff function is trivial. To do so, I can simulate such market-level shock and follow the same procedure above. Adding such a market level fixed effect may improve the fit of the model.

Specification of the Pair-Level Payoff Function So far I have not specified any functional form for the pair-level payoff function, $u(x_i, x_j, x_{ij}, z; \theta)$. In this section, I discuss the specific functional form that I will use for the estimation and the identification. In particular, I define the pair-level payoff function as follows:

$$u(x_i, x_j, x_{ij}, z; \theta) = \beta_0 + x_j'\beta_1 + (x_i - x_j)'\beta_2(x_i - x_j) + x_{ij}'\beta_3 + z'\beta_4,$$

where β_1 is the vector of parameters that measures the effect of partner firm j's characteristics, β_2 is the vector of parameters that captures the effect of the difference in the characteristics of x_i and x_j, the vector of parameters β_3 measures the effects of the pair-specific variables x_{ij}, and β_4 is the vector of parameters that measure the effects of the market characteristics. This specification is also used in Christakis et al. (2010).

Of these parameters, I am especially interested in β_2. This parameter, which captures the tendency of firms to form ties with firms with similar characteristics (referred to as *homophily* in the network literature), has been found to be pervasive in social and economic networks. If the first element of β_2 is positive, for example, it means that the bigger the difference between $x_i^{(1)}$ and $x_j^{(1)}$, the bigger firm i' payoff. Hence, I can regard the network as anti-homophilous in the first element of x_i. On the other hand, if the second element of β_2 is negative, it means that the bigger the difference between $x_i^{(2)}$ and $x_j^{(2)}$, the smaller the payoff, implying the network being homophilous in the second element of x_i.

4.2 Estimation of the Outcome Equation

The second part of my structural model concerns the outcome equation, which relates network structure with investment performance of each venture capital firm. I first provide the specification of the outcome equation, and then discuss the estimation of the full model.

My specification of the outcome equation follows Hochberg, Ljungvist and Lu (2007). For each venture capital firm i in market m, let us specify the outcome equation as follows:

$$IPOrate_{i,m} = \mathbf{x}_{i,m}'\boldsymbol{\alpha}_1 + \mathbf{W}_{i,m}^{g\prime}\boldsymbol{\alpha}_2 + \mathbf{z}_m'\boldsymbol{\alpha}_3 + u_i, \qquad (10)$$

where $IPOrate_{i,m}$ is the outcome of firm i in market m as defined in Section 2,[47] $\mathbf{x}_{i,m}$ is a vector of firm i's characteristics, $\mathbf{W}^g_{i,m}$ is a vector of the network statistics of firm i such as degree and betweenness, and $\mathbf{z}_m$ consists of market m characteristics. Note that the network statistics $\mathbf{W}^g_{i,m}$ are computed from the observed network g, e.g., the degree of firm i is denoted as $\sum_{j \in I} g_{ij}$. Parameters $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$ measure the effects of firm, network and market characteristics on performance, respectively.[48]

My specification of the outcome equation (10) is based on the model in Hochberg et al. (2007), though there are a few differences between my model and theirs. First, Hochberg et al. (2007) study fund-level investment performance, while my paper focuses on firm-level investment performance. Second, I control for market-level variables, such as MSA-level GDP, population and number of venture capital firms in the market, for which Hochberg et al. (2007) do not account for.[49]

Now, I specify the structure of error terms. In the model, I have two types of unobserved heterogeneity. First, I include firm i's unobserved heterogeneity regarding the preference over network g, $\varepsilon_i(g)$. I also consider an unobserved shock to firm i's investment performance, u_i. In the venture capital industry, it is likely that these two error terms are correlated, because venture capital firms with a higher monitoring ability, for example, are well connected because other firms wish to invest with such venture capital firms. Also, the investment performance of those venture capital firms is likely better because they can monitor start-up companies well. Such correlation makes the estimation of equation (10) by simple OLS biased, because u_i and $\mathbf{W}^g_i$ are correlated through the correlation of $\varepsilon_i(g)$ and u_i.

For tractability, I assume that the joint distribution of $(\varepsilon_i(g), u_i)$ is independent for different firms and follows a bivariate normal distribution as follows:

$$\begin{pmatrix} \varepsilon_i(g) \\ u_i \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1+\rho^2 \end{pmatrix}\right). \tag{11}$$

This specification of the error terms is similar to Sorensen (2007), but normality is not

[47] Recall that we construct $IPOrate_{i,m}$ and $\mathbf{W}^g_{i,m}$ from the observations of the same time period because we do not consider any dynamics during a time window. Potentially, that could be a concern, but we find qualitatively similar results under different definitions of the time window–such as, shorter or longer than the original five year time window.

[48] In our data, a significant fraction of venture capital firms have either a 0 or 1 $IPOrate$. Hence, without taking that fact into account, our estimates would be biased even after controlling for the endogeneity of the networks. We describe how to do this in the Appendix.

[49] Therefore, our estimates of the coefficients on the network statistics can be interpreted as the effect of each network statistic on the IPO rate *conditional on* the same size of the network. Normalized variables give qualitatively similar results.

essential for the estimation and the identification of my model. It is possible to have other specifications. The variance of $\varepsilon_i(g)$ is assumed to be 1, and the variance of u_i is set to $1+\rho^2$. These variances normalize the outcome equation and utility functions. Finally, ρ is a parameter to be estimated, which measures the correlation between two error terms.

Note that the structure of error terms in equation (11) allows us to decompose u_i such that

$$u_i = \rho\varepsilon_i(g) + \xi_i,$$

where $(\varepsilon_i(g), \xi_i)$ is independent for different firms and its joint distribution is a bivariate standard Normal distribution as follows:

$$\begin{pmatrix} \varepsilon_i(g) \\ \xi_i \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

Hence, the outcome equation becomes

$$IPOrate_{i,m} = \mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + \rho\varepsilon_i(g) + \xi_i.$$

The estimation of the outcome equation uses the moment conditions based on ξ_i, which is independently distributed with the exogenous characteristics. The idea of correcting the selection bias in the outcome equation is to use the draws of $\varepsilon_i(g)$ that support observed network g as a pairwise stable network, but multiplicity of pairwise stable networks makes it difficult. Before explaining how to construct moment conditions, let us define set T as the set of $\varepsilon_i(g)$ in which $\varepsilon_i(g)$ supports network g as a pairwise stable network *and* in which g is selected, i.e.,

$$T = \{\varepsilon_i(g) : \{g \text{ is stable } \cap \ g \text{ is selected}\}\}.$$

In this case, the researcher knows exactly the equilibrium selection mechanism. I also define set R as the set of $\varepsilon_i(g)$ in which g is supported as a pairwise stable network, i.e.,

$$R = \{\varepsilon_i(g) : g \text{ is stable}\} = T \cup \{\varepsilon_i(g) : g \text{ is stable, but not selected}\}.$$

Set T is derived from the necessary and sufficient conditions, while set R is derived from only the necessary conditions. Although set T is not observable, note that, in the estimation of the network formation model, I have simulated a lot of $\varepsilon_i(g)$ that is located in set R. Hence, I can estimate where set R is located if the number of simulation draws is large enough.[50] By definition, it is easy to see that $T \subseteq R$, which implies that

$$\inf T \geq \inf R \text{ and } \sup T \leq \sup R. \tag{12}$$

If the researcher knew the equilibrium selection mechanism perfectly, the researcher can exactly correct the selection bias. Hence, I obtain

$$E[IPOrate_{i,m} - (\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + \rho\varepsilon_i(g))|\varepsilon_i(g) \in T] = 0,$$

where dependence on exogenous characteristics $\mathbf{X}$ is omitted. Unfortunately, the researcher does not know the equilibrium selection mechanism, so it is not possible to use this moment condition directly to estimate the outcome equation. Especially, I cannot compute $E[\rho\varepsilon_i(g)|\varepsilon_i(g) \in T]$. However, if the researcher is given parameter ρ, the following inequalities must hold; (for simplicity, suppose $\rho > 0$)

$$\rho \times \inf T \leq E[\rho\varepsilon_i(g)|\varepsilon_i(g) \in T] \leq \rho \times \sup T. \tag{13}$$

Combining these inequalities (12) and (13), the following inequalities also hold;

$$\rho \times \inf R \leq E[\rho\varepsilon_i(g)|\varepsilon_i(g) \in T] \leq \rho \times \sup R.$$

Using these inequalities, I can construct two moment inequalities for the outcome equation as follows;

$$\begin{aligned} E[IPOrate_{i,m} - (\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + \rho \times \inf R)|\mathbf{X}] &\geq 0, \\ E[IPOrate_{i,m} - (\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + \rho \times \sup R)|\mathbf{X}] &\leq 0. \end{aligned}$$

If $\rho < 0$, then I alternatively have

$$\begin{aligned} E[IPOrate_{i,m} - (\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + \rho \times \sup R)|\mathbf{X}] &\geq 0, \\ E[IPOrate_{i,m} - (\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + \rho \times \inf R)|\mathbf{X}] &\leq 0. \end{aligned}$$

The researcher does not know whether $\rho > 0$ or not *ex ante*, but it is possible to construct

[50] In my implementation, I produced $10,000$ draws.

the following inequalities;

$$E\left[IPOrate_{i,m} - \left(\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + 1\{\rho > 0\}(\rho \times \inf R) + 1\{\rho < 0\}(\rho \times \sup R)\right)|\mathbf{X}\right] \geq 0,$$
$$E\left[IPOrate_{i,m} - \left(\mathbf{x}'_{i,m}\boldsymbol{\alpha}_1 + \mathbf{W}^{g\prime}_{i,m}\boldsymbol{\alpha}_2 + \mathbf{z}'_m\boldsymbol{\alpha}_3 + 1\{\rho < 0\}(\rho \times \inf R) + 1\{\rho > 0\}(\rho \times \sup R)\right)|\mathbf{X}\right] \leq 0.$$

Thus, although I do not know the equilibrium selection mechanism, it is possible to construct moment inequalities for estimating the outcome equation. Note that $\inf R$ and $\sup R$ are both functions of the observable characteristics and the parameters.

4.3 Identification

In this section, I informally discuss how to achieve point identification of the payoff function and the outcome equation. Recall that my estimation strategy is robust to lack of point identification, but it is still useful to discuss point identification for understanding which variation of the data can provide information about the identified set. I begin with the identification of the network formation model, and then the outcome equation.

My identification strategy for the strategic network formation model is based on an exclusion restriction.[51] In particular, I require the existence of a pair-specific exogenous variable with large support. More precisely, it is necessary to have a variable that affects only the two firms involved in a particular pair, but does not affect any other combination of firms. Moreover, I assume that the effect of the excluded variable on the payoff is known to be negative. In my application of venture capital networks, I use the geographic distance between the headquarters of firms i and j, denoted by d_{ij}, as the variable satisfying this condition. Since frequent personal communication is an important role of venture capital firms, it is not unreasonable to assume that firms are less likely to form ties as the distance between them increases.

The identification argument proceeds in two steps using the exclusion restriction discussed above. First, I drive d_{kl} to take the extreme values on its support for all $(k,l) \neq (i,j)$ so that $u(x_k, x_l, x_{kl}, z; \theta) \to -\infty$ and $u(x_l, x_k, x_{kl}, z; \theta) \to -\infty$, where d_{kl} is an element of x_{kl}. Note that this procedure does not change firms i's and j's payoffs, $u(x_i, x_j, x_{ij}, z; \theta)$ and $u(x_j, x_i, x_{ij}, z; \theta)$, at all. In other words, no pair of firms $(k,l) \neq (i,j)$ has the incentive to form a link at all, but only firms i and j have the possibility of forming a link. I am able to do this because it is known that d_{kl} has a negative impact on only firm k's and firm l's

[51] In the literature of empirical entry games, Tamer (2003) and Ciliberto and Tamer (2009) use a similar exclusion restriction to identify the parameters.

payoffs without affecting other firms' payoffs.[52]

Now, given that no pair of firms *except* firms (i, j) has an incentive to form any link, independent variation in g_{ij} and (x_i, x_j, x_{ij}, z) point-identify parameters in $u(x_i, x_j, x_{ij}, z; \theta)$. To see this, observe that $g_{ij} = 1$ if $u(x_i, x_j, x_{ij}, z; \theta) + \varepsilon_i(ij) > 0$ and $u(x_j, x_i, x_{ij}, z; \theta) + \varepsilon_j(ij) > 0$, and $g_{ij} = 0$ otherwise. Hence, the probability model is reduced to

$$\begin{aligned} \Pr(g_{ij} = 0 | x_i, x_j, x_{ij}, z) &= 1 - \Pr(u(x_i, x_j, x_{ij}, z; \theta) + \varepsilon_i(ij) > 0 \text{ and } u(x_j, x_i, x_{ij}, z; \theta) + \varepsilon_j(ij) > 0) \\ &= \Gamma\left(-u(x_i, x_j, x_{ij}, z; \theta), -u(x_j, x_i, x_{ij}, z; \theta)\right), \end{aligned}$$

where Γ is the joint distribution of $\varepsilon_i(ij)$ and $\varepsilon_j(ij)$.[53]

Note that this model is equivalent to a partially observed bivariate Probit model as in Poirier (1980) under the assumption that ε_i and ε_j jointly follow an independent bivariate Normal distribution. More generally, if I do not specify any distribution assumption over ε_i and ε_j, the model is equivalent to the double index models studied by Ichimura and Lee (1991) or Lewbel (2008).[54] Hence, I can follow the identification arguments in the literature on double– (or more generally multi–) index models.

In fact, the model is reduced to the double– (multiple–) index model studied by Ichimura and Lee (1991), since my specification of the pair-specific payoff, $u(x_i, x_j, x_{ij}, z)$, is linear in parameters.[55] A sufficient condition of the identification is $x_i \neq x_j$. Intuitively speaking, if $x_i = x_j$, then two indices in Γ move in exactly the same way. Then, it is not possible to identify $u(x_i, x_j, x_{ij}, z)$. However, if $x_i \neq x_j$, then $u(x_i, x_j, x_{ij}, z)$ and $u(x_j, x_i, x_{ij}, z)$ move differently, and I can identify the pair-level payoff function, $u(x_i, x_j, x_{ij}, z)$. Once I identify $u(x_i, x_j, x_{ij}, z)$ for all i and j, the identification of the discount factor, δ, is straightforward: the variation in the characteristics, $\mathbf{X}$, and the variation in the probability of having a link identify δ.

[52] This argument is called "identification at infinity," which has been extensively used in econometrics or empirical industrial organization. See, e.g., Heckman (1990) and Andews and Schafgans (1998).

[53] Note that when there are only two firms in the market, this also holds without any exclusion restriction. It is possible to identify $u(x_i, x_j, x_{ij}, z)$ from variation in markets with two players. In my dataset, there are sufficient number of markets with two firms, the pair-level payoff is identified from those markets as well.

[54] Ichimura and Lee (1991) study the case where all indices are linear in parameters. Lewbel (2008) considers the identification of a double- index model with one of the indices being linear in parameters but the other being a certain nonparametric function.

[55] Hence, we identify the distribution of ε_i nonparameterically, but the payoff function only parametrically. However, it is generally very difficult to identify double index models where both the distribution of ε_i and indices are nonparametric functions. If one wants to nonparametrically identify the payoff function $u(x_i, x_j, x_{ij}, z)$, a parametric assumption on the distribution of ε_i is necessary.

As the distance between any two firms in a market is finite in nature, the large support condition of d_{ij} seems to be too strong. This is not problematic, however, in my application because my estimation strategy is robust to the lack of point identification. Thus, to the extent that d_{ij} has sufficiently large support (even if it cannot be positive infinite in the data), the identified set would become sufficiently informative.

Finally, I briefly discuss (point) identification of the outcome equation. In general, the identification of the outcome equation with a sample correction equation as in Heckman (1990) requires an exclusion restriction. This is because if both the outcome equation and selection equation contain exactly the same covariates, then nonparametric identification of the outcome equation is not available. In my case, as in Sorensen (2007), the characteristics of other firms in the market present a source of exogenous variation. This is because those characteristics are independent of a given firm's unobserved heterogeneity, such as its monitoring ability, while they do affect the connection of the observed network.[56]

5 Results

5.1 Estimates of the Network Formation Model

In this subsection, I describe my estimates of the strategic network formation model. First, recall that my specification of the payoff function is as follows:

$$\begin{aligned} U_i(g) &= \sum_{j:d(i,j;g)\leq \bar{d}} \delta^{d(i,j;g)-1} u(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_{ij}, \mathbf{z}; \theta) + \varepsilon_i(g), \\ u(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_{ij}, \mathbf{z}; \theta) &= \beta_0 + \mathbf{x}_j'[\beta_{1,size}, \beta_{1,ex}] + (\mathbf{x}_i - \mathbf{x}_j)'[\beta_{2,size}, \beta_{2,ex}](\mathbf{x}_i - \mathbf{x}_j) + \\ &\quad \mathbf{x}_{ij}'[\beta_{3,dis2}, \beta_{3,dis1}, \beta_{3,same}] + \mathbf{z}'[\beta_{4,pop}, \beta_{4,gdp}, \beta_{4,num}], \end{aligned}$$

where $\mathbf{x}_j$ consists of partner firm j's logarithm of asset under management ($\$1M$) and logarithm of total investment experience. $\mathbf{x}_i - \mathbf{x}_j$ is the difference between firm i's and firm j's asset size and experience. $\mathbf{x}_{ij}$ is composed of the square of the (physical) distance between firm i and j, the distance between firms i and j, and the indicator variable that takes one if firms i and j are an expertise of the same industry. Finally, $\mathbf{z}$ consists of market character-

[56] In the sample selection models, correctly specifying the selection equation is crucial. Underlying strategic interactions among firms entail highly non-linear relationship between the endogenous network statistics and the exogenous characteristics. Our structural model of strategic network formation captures such nonlinearity and corrects the endogeneity appropriately.

Variable	95% Confidence Interval
Constant	
β_0	$[-2.780, -1.752]$
Partner's Characteristics	
$\beta_{1,size}$	$[-0.815, -0.552]$
$\beta_{1,ex}$	$[0.234, 0.377]$
Difference	
$\beta_{2,size}$	$[-0.114, -0.072]$
$\beta_{2,ex}$	$[0.017, 0.047]$
Pair Specific	
$\beta_{3,dis2}$	$[-0.003, -0.002]$
$\beta_{3,dis1}$	$[-0.023, -0.019]$
$\beta_{3,ind}$	$[-0.106, -0.085]$
Market Characteristics	
$\beta_{4,pop}$	$[-0.007, -0.005]$
$\beta_{4,gdp}$	$[0.006, 0.008]$
$\beta_{4,num}$	$[0.061, 0.076]$
Discount Factor	
δ	$[0.400, 0.764]$

Table 4: Confidence Interval of Parameters in the Network Formation Model – We report 95% confidence intervals using Andrews and Soares (2010).

istics, such as logarithm of population (thousand), logarithm of MSA-level GDP ($\$1M$) and logarithm of the number of venture capital firms in the market.

In Table 4, I report the 95% confidence interval of each parameter.[57] I start the discussion with coefficients on partner j's characteristics. The coefficient on partner j's asset size is estimated to be $\beta_{1,size} = [-0.815, -0.552]$, implying that venture capital firms with smaller asset size are more likely to be linked with other venture capital firms. This may be because firms with larger asset size prefer investing by themselves rather than forming syndicates. On the other hand, the coefficient on experience is estimated to be $\beta_{1,ex} = [0.234, 0.377]$. That is, venture capital firms with greater total investment experience are more likely to be linked. Because investment experience is considered a direct measure of a venture capitalist's ability, positive effects of investment experience indicate venture capitalists' preference for higher ability.

My next discussion is about estimates of parameters for the differences of characteristics

[57] Recall that my estimates of the coefficients are only partially identified because I estimate the model by a moment-inequality estimator. Hence, it is not possible to report point estimates of the parameters.

between firm i and firm j. The estimate of the parameter for the difference in the asset size is $\beta_{2,size} = [-0.114, -0.072]$. It is negative and statistically significant. This implies that firms prefer being connected with firms with similar asset size, i.e., a firm with larger asset size is more likely to be tied with a firm with larger asset size, and vice versa. Hence, the network of venture capital firms tend to show *homophily* property in terms of asset size. As for the investment experience, the effects of difference is estimated to be $\beta_{2,ex} = [0.017, 0.047]$. My estimate implies that firms prefer partners with dissimilar investment experience. In other words, the network of venture capital firms is *anti-homophily* in terms of investment experience. My findings regarding the differences of characteristics, $x_i - x_j$, are consistent with Hochberg et al.'s (2011) findings.

Next, I discuss the coefficients on pair-specific characteristics, which are reported in eleventh to thirteenth raws of the Table. Among these parameters, I am interested in the effect of the same industry indicator variable, $\beta_{3,ind}$, which is estimated to be $[-0.106, -0.085]$. That is, venture capital firms are more likely to be connected with firms with different industry expertise. Hence, the network tends to be *anti-homophilous* with respect to industry expertise. This is contrary to Hochberg et al.'s (2011) findings that the difference of industry expertise has insignificant effects.

The estimates of the parameters for market characteristics show that the effect of population is negative, while the effect of MSA-level GDP is positive. Because GDP is a measure of the economic activity of the local market, the positive effect of GDP implies that venture capital firms are more likely to be connected in economically active markets. The estimate of the parameter for the number of firms in the market is also estimated to be positive and statistically significant. As the number of firms in the market measures competitiveness of the market, my estimate indicates that firms are more connected in more competitive markets. Lastly, the discount factor, δ, is estimated to be $[0.400, 0.764]$, which indicates that the importance of indirect partners is about 40 to 75% of a direct partner and this is an important determinant of network formation in the venture capital industry.

5.2 Estimates of the Outcome Equation

In Table 5 I present parameter estimates of the outcome equation.[58] Recall that I estimate the outcome equation together with strategic network formation, and the endogeneity issue is taken into account. Note also that I do not report parameter estimates for the network

[58] The parameter estimates of the outcome equation are still preliminary. The numbers are subject to change. I will update the tables in this section once I finalize the estimation.

statistics in Table 5, but I do so in the following table. The effects of firm characteristics on the IPO rate are estimated to be both positive and significant, i.e., $\alpha_{1,size} = [0.059, 0.074]$ and $\alpha_{1,ex} = [0.023, 0.029]$. Hence, a 1% increase in asset size leads to a 6 to 7% increase in the IPO rate, and a 1% increase in experience raises the IPO rate by 2%.

As for the market characteristics, I control for the logarithm of the MSA-level population (thousand), the logarithm of MSA GDP ($\$1M$), and the logarithm of the number of venture capital firms in the market. The coefficient on the population is estimated to be negative, while the coefficient on GDP is estimated to be positive and statistically significant. The effect of market competition is measured by the number of venture capital firms in the market, and it is estimated to be negative and statistically significant ($\alpha_{3,num} = [-0.125, -0.094]$). This implies that as the market competition for finding promising start-up companies becomes intense, investment performance decreases. Finally, the correlation of the unobserved heterogeneities between the IPO rate and a firm's payoff is estimated to be positive and statistically significant. This correlation creates an endogeneity bias in the outcome equation.

In the last row of the Table, I report the estimate of the correlation between u_i and $\varepsilon_i(g)$. This is a measure of the selection on unobserved factors. If ρ is estimate to be zero, then two error terms are independent. When an unobserved variable positively affects both firms' payoffs and outcomes, selection on unobservables causes positive correlation between two error terms. I find that $\rho = [0.546, 0.694]$, which is positive and statistically significant. This implies it is possible to reject the null hypothesis that selection on unobservables does not arise.

The Effects of Network Characteristics and a Comparison of the Models with and without Bias Correction In Table 6, I report the estimates of the parameters associated with network characteristics: degree and betweenness. The effect of the degree is estimated to be positive and significant: having one additional direct partner leads to about a 1% increase in the IPO rate. I also control for betweenness of each firm, and find that its effect on the IPO rate is estimated to be non-negative and significant. Therefore, the better position the venture capital firms occupies in the network, the better is its investment performance, even after controlling for the endogeneity of the network.

In order to examine the effect of unobserved heterogeneity on the estimates of the outcome equation, I now compare the estimates with and without taking endogeneity bias into account. In the right panel of Table 6, I show the estimates of the coefficients on the network characteristics when I estimate the outcome equation separately from the network formation

Variable	95% Confidence Interval
Constant	
α_0	$[-0.861, -0.689]$
Firm's Characteristics	
$\alpha_{1,size}$	$[0.059, 0.074]$
$\alpha_{1,ex}$	$[0.023, 0.029]$
Market Characteristics	
$\alpha_{3,pop}$	$[-0.034, -0.020]$
$\alpha_{3,gdp}$	$[0.705, 0.934]$
$\alpha_{3,num}$	$[-0.125, -0.094]$
Correlation	
ρ	$[0.546, 0.694]$

Table 5: Confidence Interval of the Parameters in the Outcome Equation–We report 95% confidence intervals using Andrews and Soares (2010). Estimates of the coefficients on the network characteristics are reported in Table 6.

Variable	Full Model	Original Model
Degree	$[0.009, 0.012]$	$[0.017, 0.049]$
Betweeness	$[0.000, 0.003]$	$[0.002, 0.005]$

Table 6: Comparison of the 95%–Confidence Intervals of the Parameters of the Network Characteristics in the Outcome Equation

model. I find that both network characteristics are *overestimated* in absolute value unless I account for the fact that the network characteristics are endogenous. Hence, if I do not take the endogeneity of the network structure into account, the estimated effects of the network characteristics involve not only the causal effect of the network structure but also the selection effect, whereby those who obtain a higher $\varepsilon_i(g)$ are likely to achieve a higher IPO rate. Given that the average IPO rate is around 18%, the degree of overestimation is economically significant.

5.3 Counterfactual Simulation

Finally, I conduct a policy experiment. The policy I consider is government's direct investment by creating a venture capital firm. This type of government involvement is actually carried out in Canada or some European countries. More precisely, I exogenously add a new entrant with particular characteristics to each market and then find all pairwise stable networks under the new market structure including both the incumbent firms and the new entrant.[59] I consider four different types of the entrant; i) small asset size and small investment experience; ii) small asset size and large investment experience; iii) large asset size and small investment experience; and iv) large asset size and large investment experience. Once I find all pairwise stable networks, I compute each venture capital firm's degree centrality and IPO rate for each pairwise stable network, and report some statistics of the changes of degree centrality and IPO rate across markets.[60]

I report results of the counterfactual experiments in Table 7. In the left panel of the table, I report the average, minimum, and maximum changes of degree centrality. All numbers are the average across markets. More precisely, I compute the average, minimum, and maximum changes of the degree centrality across all players for each stable network and for each market, and then take the average of each statistic across stable networks and markets. In the right panel, I report the average, minimum, and maximum changes of IPO rate across all markets, which are computed in the same way.[61] Note that the change of IPO rate accounts for the negative effect of competition by having one more firm on each firm's IPO rate[62] and the

[59] Since our network formation model assumes that the number of firms is exogenously given, we cannot study whether a potential entrant has incentive to enter the market. Incorporating the entry decision into the model is left for a future research.

[60] In order to reduce the computational burden, we focus on the small markets with less than 6 firms in this counterfactual.

[61] Hence, we assume that each pairwise stable network is chosen by the same probability.

[62] This is because our estimate of the coefficient $\alpha_{3,num}$ is negative.

New Entrant	Change of Degree			Change of IPO rate		
(Size, Experience)	Mean	Min	Max	Mean	Min	Max
(Small, Small)	0.46	−0.20	2.13	−0.11	−0.23	0.00
(Large, Small)	0.51	−0.18	2.56	0.03	−0.08	0.12
(Small, Large)	0.33	−0.23	2.34	−0.05	−0.17	0.04
(Large, Large)	0.55	−0.15	2.65	0.09	−0.02	0.18

Table 7: Counterfactual Experiment–Change of the degree and the IPO rate

effect of the additional entrant's characteristics.

I start discussion about the counterfactual experiments from the change of degree centrality. In all four cases, an additional entrant leads to an increase of each player's degree centrality on average. However, the effect of the additional firm varies depending on the characteristics of the new entrant. I find that the effect is the largest when the additional firm has large asset size and large investment experience, and the smallest when the additional firm has small asset size, but has large experience.

In the right panel, I compute the change of the IPO rate to examine how the change of the equilibrium network structure translates to the change of investment performance. I find, in the first case, that the change of IPO rate is negative in any stable network. Comparing the second case and the third case, my results indicate that IPO rate is improved more when a firm with large asset size and small experience is added than when a firm with small asset size and great experience is added. Lastly, IPO rate is likely to increase by from −2% to 18%, the average of which is 9% when I add a firm with large size and large experience.

Given that my estimates of the coefficients on asset size and investment experience are positive, it is natural that the IPO rate is increased when a firm with large asset size and/or large experience enters, but adding one more entrant also has an indirect effect on the IPO rate through the change of the network structure. For example, the largest change of the IPO rate is 0% even if a firm with small asset size and small experience is added. In this case, the indirect effect through the change of the network structure just offsets the negative impact of having an additional firm.

6 Conclusion

In this paper, I study strategic network formation and the effect of network structure on the investment performance in the venture capital industry. The contributions of my paper are as follows: i) I propose an estimation strategy of strategic network formation models that is computationally feasible and is robust to the multiplicity of pairwise-stable networks, and ii) I address the potential bias of the outcome equation due to the endogeneity of the network structure by jointly estimating the network formation model and the outcome equation in which I allow for the correlation between unobserved heterogeneity in the payoff function and in the outcome equation.

I find that the network of venture capital firms tends to show a homophily property in terms of asset size, but an anti-homophily property in terms of both investment experience and industry expertise. The estimates of the outcome equation show that there exists a selection bias if I do not take the endogeneity of the networks into account. In particular, the effects of the network structure on investment performance are significantly overestimated. I also conduct a counterfactual experiment to see the effect of a new entrant on the network structure and the IPO rate.

Finally, there are many issues left for future research. First, I can conduct some counterfactual policy experiments using the estimates obtained from my structural model. A possible counterfactual experiment would be investigating what a socially efficient network looks like, and how the IPO rate would be affected under this type of efficient network structure. Second, the method I propose in this paper can be generally applied to many industries. The outcome equation, for example, can be replaced by the production function equation, allowing us to infer the effect of the network structure on firm productivity. Lastly, I do not consider any industry dynamics of the venture capital industry in this paper. A potential benefit of forming a link today would be facilitating syndicate formation in a future projects (see, e.g., Ljungqvist, Marston and Wilhelm, 2007). I study the dynamics of network formation in an ongoing project.

References

[1] Aguirregabiria, V., R. Clark, and H. Wang (2012)“Diversification of Geographic Risk in Retail Bank Networks: Evidence from Bank Expansion after the Riegle-Neal Act,” mimeo.

[2] Andrews, D. and M. Schafgans (1998) "Semiparametric Estimation of the Intercept of a Sample Selection Model," *Review of Economic Studies*, Vol. 65, pp. 487-517.

[3] Andrews, D. and G. Soares (2010) "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selections," *Econometrica*, Vol. 78, pp. 119-157.

[4] Bajari, P., C. L. Benkard, and J. Levin (2007) "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, Vol. 75, pp. 1331-70.

[5] Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010) "Estimating Static Models of Strategic Interactions," *Journal of Business and Economic Statistics*, Vol. 28, pp. 469-82.

[6] Bala, V. and S. Goyal (2000), "A Noncooperative Model of Network Formation," *Econometrica*, Vol. 68, pp.1181-1229.

[7] Beresteanu, A., F. Molinari, and I. Molchanov (2008) "Sharp Identification Regions in Games," mimeo.

[8] Berry, S. (1992) "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, Vol. 60, pp. 889–917.

[9] Bhagwat, Vineet (2011) "Manager Networks and Investment Syndication: Evidence from Venture Capital," mimeo.

[10] Brander, J., Amit, R., and W. Antweiler (2002) "Venture-Capital Syndication: Improved Venture Selection vs. the Value-Added Hypothesis," *Journal of Economics and Management Strategy*, Vol. 11, pp. 423-52.

[11] Bresnahan, T. and P. Reiss (1991) "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, Vol. 99, pp. 977–1009.

[12] Cestone, G., J. Lerner, and L. White (2007) "The Design of Syndicates in Venture Capital," mimeo.

[13] Chandrasekhar, A., and M. O. Jackson (2012) "Tractable and Consistent Random Graph Models," mimeo.

[14] Chandrasekhar, A., and R. Lewis (2011) "Econometrics of Sampled Networks," mimeo.

[15] Chakrabarti, S., and R. P. Gilles (2007) "Network Potentials," *Review of Economic Design*, Vol. 11, pp. 13-52.

[16] Christakis, N., J. Fowler, G. Imbens, and K. Kalianaraman (2010) "An Empirical Model of Strategic Network Formation," mimeo.

[17] Ciliberto, F. and E. Tamer (2009) "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, Vol. 77, pp. 1791-1828.

[18] Cohen, L., A. Frazzini, and C. Malloy (2008) "The Small World of Investing: Board Connections and Mutual Fund Returns," *Journal of Political Economy*, Vol. 116, pp. 951-979.

[19] Conley, T. G. and C. R. Udry (2010) "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, Vol. 100, pp. 35-69.

[20] Currarini, S., M. O. Jackson, and P. Pin (2009) "An Economics Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, Vol. 77, pp. 1003-45.

[21] Eisenhardt, K. and C. Schoonhoven (1996) "Resource-Based View of Strategic Alliance Formation," *Organization Science*, Vol. 7, pp. 136-50.

[22] Ellickson, P. and S. Misra (2012) "Enriching Interactions: Incorporating Outcome Data into Static Discrete Games," *Quantitative Marketing and Economics*, Vol. 10, pp. 1-26.

[23] Fox, J. (2010a) "Identification in Matching Games," *Quantitative Economics*, Vol. 1, pp. 203-54.

[24] Fox, J. (2010b) "Estimating Matching Games with Transfers," mimeo.

[25] Galichon, A. and M. Henry (2011) "Set Identification in Models with Multiple Equilibria," *Review of Economic Studies*, Vol. 78, pp. 1264-1298.

[26] Geyer, C. and E. Thompson (1992) "Constrained Monte Carlo Maximum Likelihood for Dependent Data," *Journal of the Royal Statistical Society*, Vol. 54, pp. 657-699.

[27] Goldsmith-Pinkham, P. and G. W. Imbens (2011) "Social Networks and the Identification of Peer Effects," mimeo.

[28] Gompers, P. A. (1996) "Optimal Investment, Monitoring, and the Staging of Venture Capital," *Journal of Finance*, Vol. 50, pp.1461-90.

[29] Gompers, P. and J. Lerner (1999) *The Venture Capital Cycle*, MIT Press.

[30] Gordon, N. and B. Knight (2009) "A Spatial Merger Estimator with an Application to School District Consolidation," *Journal of Public Economics*, Vol. 93, pp. 752-65.

[31] Heckman, J. (1990) "Varieties of Selection Bias," *American Economic Review*, Vol. 80, pp. 313-18.

[32] Hellman, T (2012) "On the Existence and Uniqueness of Pairwise Stable Networks," forthcoming in *International Journal of Game Theory*.

[33] Hellman, T and M. Puri (2002) "Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence," *Journal of Finance*, Vol. 57, pp. 169-97.

[34] Ho, K. (2009) "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, Vol. 99, pp. 393-430.

[35] Hochberg, Y. (2011) "Venture Capital and Corporate Governance in the Newly Public Firm," forthcoming in *Review of Finance*.

[36] Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2007) "Whom You Know Matters: Venture Capital Networks and Investment Performance," *Journal of Finance*, Vol. 62, pp.251-301.

[37] Hochberg, Y., L. Lindsay, and M. Westerfield (2011) "Partner Selection in Co-Investment Networks: Evidence from Venture Capital," mimeo.

[38] Hochberg, Y., M. Mazzeo, and R. McDevitt (2011) "Specialization and Competition in the Venture Capital Industry," mimeo.

[39] Hochberg, Y. and J. Rauh (2012) "Syndication Networks and the Spatial Distribution of Venture Capital Investments," mimeo.

[40] Ichimura, H. and L. Lee (1991) "Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation," *Non-Parametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press.

[41] Jackson, M (2008) *Social and Economic Networks*, Princeton University Press.

[42] Jackson, M and A. van den Nouweland (2005) "Strongly Stable Networks," *Games and Economics Behavior*, Vol. 51, pp. 420-44.

[43] Jackson, M and B. Rogers (2007) "Meeting Strangers and Friends of Friends: How Random Are Social Networks?," *American Economic Review*, Vol. 97, pp. 890-915.

[44] Jackson, M. and A. Watts (2001) "The Existence of Pairwise Stable Networks," *Seoul Journal of Economics*, Vol. 14, pp. 299-321.

[45] Jackson, M. and A. Wolinsky (1996) "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, Vol. 71, pp.44-74.

[46] Kaplan, S. and A. Schoar (2005) "Private Equity Performance: Returns, Persistence and Capital Flows," *Journal of Finance*, Vol. 60, pp.1791-1823.

[47] Kinnan, C. and R. Townsend (2011), "Kinship and Financial Networks, Formal Financial Access and Risk Reduction," mimeo.

[48] Kuemmerle, W., K. Kobayashi, and C. Ellis (2000) "*JAFCO America Ventures, Inc.: Building a Venture Capital Firm*," Boston, MA: Harvard Business School Publishing.

[49] Lerner, J. (1994) "The Syndication of Venture Capital Investment," *Financial Management*, Vol. 23, pp. 16-27.

[50] Lerner, J. (1995) "Venture Capitalists and the Oversight of Private Firms," *Journal of Finance*, Vol. 50, pp. 301-18.

[51] Lewbel, A. (2008) "Semiparametric Double Index Model Identification and Estimation," mimeo.

[52] Lindsey, L. (2008) "Blurring Firm Boundaries: The Role of Venture Capital in Strategic Alliances," *Journal of Finance*, Vol. 63, pp. 1137-68.

[53] Ljungqvist, A., F. C. Marston, and W. J. Wilhelm (2009) "Scaling the Hierarchy: How and Why Investment Banks Compete for Syndicate Co-Management Appointments," *Review of Financial Studies*, Vol. 22, pp. 3977-4007.

[54] Manski, C. F. (1993) "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, Vol. 60, pp. 531-542.

[55] Mazzeo, M. (2002a) "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, Vol. 33, pp. 1-22.

[56] Mazzeo, M. (2002b) "Competitive Outcomes in Product-Differentiated Oligopoly," *Review of Economics and Statistics*, Vol. 84, pp. 716-728.

[57] McPherson, M., L. Smith-Lovin and J. M. Cook (2001) "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, Vol. 27, pp. 415-444.

[58] Mele, A. (2011) "A Structural Model of Segregation in Social Networks," mimeo.

[59] Myerson, R. (1991), *Game Theory: Analysis of Conflict*, Harvard University Press.

[60] Pakes, A. (2010) "Alternative Models for Moment Inequalities," *Econometrica*, Vol. 78, pp. 1783-1822.

[61] Pakes, A., J. Porter, K. Ho, and J. Ishii (2011) "Moment Inequalities and Their Application," mimeo.

[62] Poirier, D. (1980) "Partial Observability in Bivariate Probit Models," *Journal of Econometrics*, Vol. 12, pp. 209-17.

[63] Robinson, D. and O. Stuart (2007) "Network Effects in the Governance of Strategic Alliances," *Journal of Law, Economics, and Organization*, Vol. 23, pp. 242-73.

[64] Sah, R. K., and J. E. Stiglitz (1986) "The Architecture of Economic Systems: Hierachies and Polyarchies," *American Economic Reiview*, pp. 716-27.

[65] Snijders, T. (2002) "Markov Chain Monte Carlo Estimation of Exponential Random Graph Models," *Journal of Social Structure*, Vol. 3.

[66] Sorenson, O. and T. Stuart (2000) "Syndication Networks and the Spatial Distribution of Venture Capital Investments," *American Journal of Sociology*, Vol. 106, pp. 1546-86.

[67] Tamer, E. (2003) "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, Vol. 70, pp. 147-165.

[68] Tamer, E. (2010) "Partial Identification Approach in Econometrics," *Annual Review of Economics*, Vol. 2, pp. 167-95.

[69] Uetake, K. and Y. Watanabe (2012a) "A Note on Estimation of Two-Sided Matching Models," mimeo.

[70] Uetake, K. and Y. Watanabe (2012b) "Entry by Merger: Estimates from a Two-Sided Matching Model with Externalities," mimeo.

[71] Weese, E. (2011) "Political Mergers as Coalition Formation," mimeo.

[72] Wright, M. and A. Lockett (2003) "The Structure and Management of Alliances: Syndication in the Venture Capital Industry," *Journal of Management Studies*, Vol. 40, pp. 2073-2102.

7 Appendix

7.1 Computation of Confidence Interval

The model has 21 parameters including both parameters in the strategic network formation model and the outcome equation, and the confidence set, which I denote as CS, is a 21dimensional object. As it is not possible to present a 21-dimensional object in a convenient way, I present the min and max of the CS along each dimension in Table. In the following, I explain how I obtained the min and the max of the CS along each dimension.

Following the notation of Andrews and Soares (2010), a parameter value θ is included in CS if $T_n(\theta) \leq \widehat{c}_n(\theta, 1-\alpha)$ where $T_n(\theta)$ is the test statistic and $\widehat{c}_n(\theta, 1-\alpha)$ is the critical value. Denoting the j-th element of θ by θ^j, I will report $\underline{\theta}^j = \min\{\theta^j | \theta \in CS\}$ and $\overline{\theta}^j = \max\{\theta^j | \theta \in CS\}$. Though computing CS directly is extremely costly given that the CS has 21 dimensions, I can compute $\underline{\theta}^j$ within manageable time by solving the following constrained optimization problem for each of j-th dimension;

$$\begin{aligned} &\min_{\theta} \theta^j \\ &s.t.\ T_n(\theta) \leq \widehat{c}_n(\theta, 1-\alpha) \end{aligned}$$

where $\theta^j$is the j-th element of θ. By maximizing instead of minimizing θ^j, I can obtain $\overline{\theta}^j$. I repeat this for $j = 1, ..., 21$ and report the optimizer in Tables in Section 5.

7.2 Estimation of the Outcome Equation When the Outcome is Censored

In my empirical application, the performance measure of venture capital firms, i.e., IPO rate, is not distributed smoothly. In particular, a significant fraction of firms has IPO rate of either 0 or 1. Hence, estimates might be biased unless I take such fact into account. To do so, I alternatively consider the following model of the outcome equation:

$$\begin{aligned} y_i &= \begin{cases} 0 & \text{if } y_i^* \leq 0, \\ y_i^* & \text{if } y_i^* \in (0, a), \\ 1 & \text{if } y_i^* \geq a, \end{cases} \\ y_i^* &= \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{W}_i^{g\prime}\boldsymbol{\gamma} + \rho\varepsilon_i(g) + \xi_i, \end{aligned}$$

where y_i is the outcome measure observed in the data, y_i^* is a latent variable which measures unobserved productivity determining each firm's IPO rate, and a is a threshold parameter for observing IPO rate of 1, which I also estimate. I construct a moment equality of this model as follows. Taking conditional expectation, I have

$$\begin{aligned} E[y_i|X] &= \Pr(y_i^* \leq 0)E[y_i|\mathbf{X}, y_i^* \leq 0] + \Pr(y_i^* \in (0,1))E[y_i|\mathbf{X}, y_i^* \in (0,1)] \\ &\quad + \Pr(y_i^* \geq 1)E[y_i|\mathbf{X}, y_i^* \geq 1] \\ &= 1 - \Phi\left(\frac{a - \mathbf{x}_i'\boldsymbol{\beta} - \mathbf{W}_i^{g\prime}\boldsymbol{\gamma}}{\sqrt{1+\rho^2}}\right) \\ &\quad + \left(\Phi\left(\frac{a - \mathbf{x}_i'\boldsymbol{\beta} - \mathbf{W}_i^{g\prime}\boldsymbol{\gamma}}{\sqrt{1+\rho^2}}\right) - \Phi\left(\frac{-\mathbf{x}_i'\boldsymbol{\beta} - \mathbf{W}_i^{g\prime}\boldsymbol{\gamma}}{\sqrt{1+\rho^2}}\right)\right) \\ &\quad \times \left(\mathbf{x}_i'\boldsymbol{\beta} + \mathbf{W}_i^{g\prime}\boldsymbol{\gamma} + E\left[\rho\varepsilon_i(g) + \xi_i|\mathbf{X}, y_i^* \in (0,1)\right]\right), \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard Normal distribution. I add this moment to the moment conditions derived from the network formation game. Note that I can easily compute $E\left[\rho\varepsilon_i(g) + \xi_i|\mathbf{X}, y_i^* \in (0,1)\right]$ by the similar way that I have simulated $E\left[\varepsilon_i(g)|\mathbf{X}\right]$ by simulating $\varepsilon_i(g)$ and ξ_i.

7.3 Robustness Checks

In this appendix, I discuss the robustness of my estimation results in Section 5. In particular, I check whether my estimation results are robust to different definitions of a time window that venture capital firms play a network formation game in a geographic market. In the main text of the paper, I define a market by a MSA and 5 year window. I consider the following 3 different definitions of a time window; i) first 3 years of each five-year window (i.e., 1980-82, 85-87, ...), ii) first 5 years of each ten-year window (i.e., 1980-84, 1990-94, 2000-2004), and iii) first and last 10 years in my sample (1980-1989, 1996-2005). The specification

Variable	Robust 1 Confidence Interval	Robust 2 Confidence Interval	Robust 3 Confidence Interval
Constant			
β_0	$[-2.521, -1.621]$	$[-2.067, -1.654]$	$[-2.077, -1.532]$
Partner's Characteristics			
$\beta_{1,size}$	$[-0.709, -0.574]$	$[-0.716, -0.573]$	$[-0.473, -0.128]$
$\beta_{1,ex}$	$[0.304, 0.372]$	$[0.301, 0.376]$	$[0.355, 0.524]$
Difference			
$\beta_{2,size}$	$[-0.097, -0.078]$	$[-0.097, -0.035]$	$[-0.087, -0.023]$
$\beta_{2,ex}$	$[0.037, 0.045]$	$[0.036, 0.046]$	$[0.036, 0.062]$
Pair Specific			
$\beta_{3,dis2}$	$[-0.003, -0.002]$	$[-0.003, -0.002]$	$[-0.003, -0.002]$
$\beta_{3,dis1}$	$[-0.024, -0.020]$	$[-0.039, -0.019]$	$[-0.023, -0.004]$
$\beta_{3,ind}$	$[-0.113, -0.091]$	$[-0.101, -0.071]$	$[-0.108, -0.087]$
Market Characteristics			
$\beta_{4,pop}$	$[-0.014, -0.006]$	$[-0.008, -0.006]$	$[-0.007, -0.006]$
$\beta_{4,gdp}$	$[0.009, 0.011]$	$[0.009, 0.011]$	$[0.006, 0.008]$
$\beta_{4,num}$	$[0.029, 0.080]$	$[0.065, 0.218]$	$[0.081, 0.148]$
Discount Factor			
δ	$[0.505, 0.630]$	$[0.632, 0.968]$	$[0.688, 0.860]$

Table 8: Robustness Checks – Confidence Intervals under Different Definitions of a Geographic Market

of the payoff function is the same as in the main text.

In Table 8, I show the estimates of the parameters in the strategic network formation model. All three columns show that the estimates are qualitatively similar regardless of the definition of a market. Therefore, I consider my findings regarding the structure of venture capital networks and the effects of network structure on their investment performance remain to be valid.