Abstract

Venture Capital Communities

Syndicates account for two-thirds of the capital invested by venture capital firms. Through syndications, venture capitalists form several non-exclusive, partially overlapping partnerships with other VC firms. We study the structure of these inter-VC alliances. We show that the repeated partnerships leads to agglomeration of VC firms into "communities" or soft-border conglomerates whose members are probabilistically more likely to partner with each other than with outsiders. We characterize the number and composition of communities and their economic effects. Communities exhibit subtle composition effects with heterogeneity on the dimensions of size, influence, and geography but homogeneity in industry and stage focus. These effects are consistent with resource complementarity theories of organizational boundaries as well as theories in which syndicate members rely on and value each other for evaluation, screening, and risk-sharing. Community membership is associated with positive economic outcomes for portfolio firms. Firms sourcing capital from community VCs are more likely to exit and do so sooner. Our results are consistent with models in which VCs learn by doing and the effectiveness of learning depends on the nature of a VC's syndicate partners.

Key words

Venture Capital, Syndication, Community Detection, Social Interactions

JEL classification

G20, G24

1 Introduction

Venture capital (VC) firms are financial intermediaries that provide capital to young entrepreneurial firms. They raise capital from wealthy individuals or institutional investors such as pension funds and university endowments to invest in young and risky ventures with high upside. The VC industry has grown significantly since the first limited partnership was formed in 1958. According to the National Venture Capital Association, there are over 56,000 VC cash-for-equity deals for $429 billion in the U.S. between 1995 and 2009. Venture capital has spawned successful firms such as Apple Computers, Cisco, and Microsoft.

Gompers and Lerner (2001, 2004) classify venture capital activities into three phases of the "venture cycle." These phases comprise fund-raising by VC firms, investment in portfolio companies, and exit. We focus on the second stage of the cycle, the VC investment process. VC firms tend to invest in young firms with few tangible assets and unproven business models, making the investments highly risky. VC investments are also resource intensive. As Gompers and Lerner (2004) write, a venture capitalist will often conduct more than 100 reference checks prior to investing. After investing, VCs serve on the boards of their portfolio companies, conduct site visits to monitor performance, provide advice on growth strategies, help recruit key personnel, professionalize the firms or find strategic partners. Thus, VC investments are risky and resource intensive, demanding considerable efforts both in ex-ante screening and ex-post investment effort in providing strategic direction to portfolio companies.[1]

VC firms use several strategies to manage the risks and resource demands placed by the investment process. For instance, VC contracts routinely include security design features

[1] See Lerner (1995) for directorships, Bygrave and Timmons (1992) or Gorman and Sahlman (1989) for advising, Hellmann and Puri (2002) for professionalization, and Lindsey (2008) for strategic partnering roles of VC firms. Casamatta (2003) models the dual advising-financing function.

such as priority and staging or covenants to mitigate potential agency or holdup problems.[2] An important element of a venture capitalist's strategy is syndication, or co-investing in portfolio firms together with other VC firms. Syndicated deals comprise a significant portion of VC financing. For instance, in the U.S. venture capital market, syndicated rounds account for 44% of the rounds financed and 66% of VC investment proceeds.[3] Syndication is also important when viewed from the VC firm's perspective. For instance, in the time period between 1980 and 1999, only 5% of U.S. VC firms never syndicate and these are small, peripheral players. 95% of VCs enter into at least one syndicate and often do so with multiple non-overlapping partners.

A syndicated round can be viewed as a collaborative effort between the venture capitalists who finance the round. These collaborations are not exclusive, so VC firms that join together in one syndicate are not bound to work together in future deals. In fact, top venture capitalists can enter into syndicate partnerships with hundreds of VC firms and many venture capitalists enter into and manage multiple syndications at the same time. These patterns suggest that structure of alliances formed by venture capitalists, with multiple relationships with several partners at the same time, is likely more complex than suggested by static models of contracting between two parties.

We examine the structure of the partnerships formed by venture capitalists through the syndication process. We show that VC firms do not choose syndicate partners at random. Rather, they tend to prefer some syndicate partners over others, resulting in agglomeration of VC firms into spatial clusters that we term as VC *communities*. We examine community formation in the VC industry and ask three interrelated questions. First, is there evidence

[2]See Neher (1999) or Cornelli and Yosha (2003) on security design and Kaplan and Stromberg (2003, 2004), Robinson and Stuart (2007) or Robinson and Sensoy (2011) for evidence on VC contracts.

[3]For evidence in non-US markets, see Lockett and Wright (2001), Brander, Amit, and Antweiler (2002), or Hopp and Rieder (2010).

of community formation? Second, what is their composition? In particular, is there heterogeneity, in the sense that VCs with diverse skills and resources form alliances with each other? Alternatively, is there homophily so that VCs with similar skill sets partner with each other, consistent with the view that VCs with similar functional capabilities rely on each other to screen and certify investments? Finally, is sourcing capital from a community VC beneficial for a portfolio company seeking venture capital?

Briefly, we find extensive and robust evidence of community formation in a sample spanning 20 years of VC syndication data. Communities exhibit subtle composition effects with both homogeneity and heterogeneity on different dimensions, as we explain later. Finally, portfolio firms sourcing capital from community VCs are more likely to exit and do so faster. The findings echo the view of the recent literature on strategic alliances between firms (e.g., Robinson and Stuart, 2007; Robinson, 2008). As the literature emphasizes, strategic alliances involve simultaneous, multilateral relationships that likely go beyond simple models of two-party bilateral contracting. We provide empirical evidence of such effects. While our results support the resource complementarity view of strategic alliance formation, they are also consistent with models such as Sorensen (2008) in which (VC) firms learn by doing and the effectiveness of learning depends on the nature of the firm's alliance partners.

Before discussing our methods and results, we briefly review examples of VC partnerships and consider the economic forces that motivate community formation. The propensity to prefer some partners over others is illustrated by the partnerships of J. P. Morgan Ventures. Between 1980 and 1999, J. P. Morgan Ventures co-invests with 640 different venture capital firms. Figure 1 displays the frequency distribution of J. P. Morgan's partners. The distribution has thick left mass, indicating that some partners are preferred over others. The distribution also has a long and thin right tail, indicating that the list of J. P. Morgan's syndicate partners is extensive. Figure 2 plots the distribution of the top 20 partners for J. P. Morgan Partners, Matrix Partners, Sequoia Capital, and Kleiner Perkins. over the 1980-1999 period. The patterns are similar to those in Figure 1.

While not displayed in Figures 1 and 2, we also find that different VC firms can have different – but not necessarily mutually exclusive – preferred partners. For instance, the top 5 partners of J. P. Morgan are Kleiner Perkins, Oak Investment Partners, J. F. Shea, Bay Partners, and Mayfield Fund. The top 5 partners for Kleiner Perkins are Mayfield Fund, J. P. Morgan, Institutional Venture Partners, New Enterprise Associates and Sequoia Capital. Thus, community formation is a *probabilistic* rather than a deterministic propensity to prefer some partners, so community VCs can syndicate outside their native communities. For instance J. P. Morgan prefers some partners but enters into syndicates with several others. The existence, the nature, and the economic consequences of community formation in the overall VC sample are the questions we pursue here.

As economic motivation, we consider why venture capitalists may self-organize into communities. Success in venture capital demands skills in selecting good investments and skills in maturing the portfolio companies (Sorensen (2007); Das, Jo and Kim (2011)). Some of the skill set and knowledge to be successful is undoubtedly endowed. However, the existing stock of skills and knowledge must be renewed or refined and new skills learnt because the businesses funded by VC capital tend to be immature and have unproven business models. For the same reason, much of the learning in the VC business is hands-on, through the act of investing (Goldfarb, Kirsch, and Miller (2007), Sorensen (2008)). Learning improves current outcomes and also improves future decision-making, as the VC firm can better identify, evaluate, and develop future opportunities in related sectors. VC community formation is consistent with the joint hypothesis that venture capitalists learn through investing with their syndicate partners and that learning is enhanced when syndicate partners are familiar.

The learning hypotheses for community formation has two elements. One, venture capitalists learn from their syndicate partners, and two, that syndication with familiar partners enhances learning. Some part of a venture capitalist's learning comes from its own screening and ongoing involvement with a portfolio company. However, syndicate partners also contribute to skill development. The partners conduct their own screening prior to investment. As Gompers and Lerner (2004) write, venture capitalists will often not invest unless acceptable syndicate partners agree that the project is desirable. Syndicate partners also participate in the ongoing evaluation and monitoring of an investment ex-post. The ex-ante and ex-post activities of partners generate valuable informational signals that a venture capitalist can learn from.

A second element of the learning hypothesis is that syndication with familiar partners can enhance the usefulness of the information flows from partners. Familiarity can matter because it facilitates flow of informal knowledge due to better understanding of partners' norms and processes (Gertler (1995); Porter (2000)). Sociologists emphasize a similar point. For example, in an influential article, Granovetter (1985) points out that agents place far greater confidence in information flows from trusted sources than from unfamiliar ones. Thus, venture capitalists are more likely to place faith in – and benefit from – the assessments of syndicate partners when the partners are familiar. Familiarity can arise temporally from a venture capitalist's past dealings with the potential partner. It can also arise spatially from second-hand knowledge when the potential partner has transacted with a VC's network of former partners.[4] The general point, as Centola (2010) writes, is that clustering is advantageous when social reinforcement matters.

[4]For example, the VC firm Matrix Partners writes on its website "... The best way to get in touch with our team is through an introduction from someone you know in our network." http://matrixpartners.com/site/about_partnering-with-matrix, accessed May 3, 2011.

The formation of communities is also an implication of incomplete contracting theories. Venture capitalists operate in environments of high information asymmetry not only about their portfolio companies but also about potential syndicate partners and what they bring to the table. The suspicion that potential partners may free ride or ex-post hold up a venture can cause VC firms to underinvest effort in their syndicated investments (Grossman and Hart (1986); Hart and Moore (1990)). Familiarity can mitigate these problems through multiple channels. First, it can diminish the scale of the incomplete contracting problems because familiar partners face less asymmetric information about each other. Alternatively or additionally, familiarity can build trust, which leads to equilibrium strategies in which agents treat partners better and reciprocally expect (and receive) better treatment (Guiso, Sapienza, and Zingales (2004); Bottazzi, Da Rin and Hellmann (2011)). Similar effects arise in economic models of reciprocity or models of social interactions where interactions can result in multiplier effects.[5] The bottom line is that if familiarity mitigates the problems arising out of incomplete contracting, it leads to a greater likelihood of syndication with familiar partners and thus clustering of venture capitalists into communities.

Besides examining community formation, we also test the performance consequences associated with community membership. These tests help sort out the alternative perspective that communities may form but are economically neutral. For instance, VC firms may cluster into communities simply because of a behavioral affinity for the familiar. Alternatively, VC firms may choose familiar partners to lower routine administrative and paperwork costs. To sort out these less material reasons for community formation from deeper motivations such as learning, we examine the ex-post performance of community VC funded ventures.

Our main sample comprises 39,725 unique VC investment rounds made in the U.S. be-

[5]See Granovetter (1985) and Glaeser, Sacerdote and Scheinkman (1996) for a discussion of the economic effects of social interactions. See Rabin (1993) and Fehr and Schmidt (2005) for reciprocity in economics and Cai (2010) for reciprocity in loan syndications.

tween 1980 and 1999 in 15,455 portfolio firms. This sample begins after the Employee Retirement Income Security Act, which led to the institutionalization of the VC market. The sample ends in 1999, which allows sufficient time periods for judging the performance of the VC-funded portfolio firms. We identify VC communities based on rolling windows of 5 years between years $t-5$ and $t-1$ and use this information to analyze the performance of VC investments made in year t.

From a computational perspective, identifying communities is essentially a clustering problem, although a computationally difficult one to accommodate the real world features of VC data. The number of community clusters is neither known nor fixed across time periods. Nor do we constrain the size of each community. In fact, each community can comprise a different number of VC firms. Moreover, community boundaries are fuzzy as community VC firms also syndicate outside the community. We employ the technique of modularity optimization, which effectively partitions VC firms into groups such that the groups are tight knit internally but have looser connections outside and use the fast walk-trap method suggested by Pons and Latapy (2005), to optimize modularity.

We detect several communities in every five year sub-period in our sample. The number of communities varies from 12 to 35 in the individual five-year periods. Thus, like Lindsey (2008), we too find a "blurred boundary" effect. Her study explains how venture capitalists can blur boundaries between portfolio *firms*. Here, communities blur boundaries between *venture capitalists*. We find remarkable persistence among community VCs. About 75 percent of community VCs continue to be part of a community after five years. We examine whether sourcing funds from community VCs is beneficial to the company receiving the funding. We follow the VC literature (e.g., Brander, Amit and Antweiler (2002); Lindsey (2008); Sorensen (2008)) and define success as exit either by an IPO or by merger with another company. Community membership is significant in probit models predicting 10-year exits or in Cox hazard models modeling the time to exit.

Our results are robust to a broad set of controls. These include measures of VC firm influence used in the literature such as VC age, centrality, industry and year fixed effects, geographical clustering of portfolio companies and VC firms, ownership type of VC firms and their stage and industry focus. In particular, our results are also robust to controls for whether a deal is syndicated or not. The fact that syndicated deals are more successful is well known.[6] We too find a syndication effect. Our main point, however, is that the structure and composition of syndicates also matters beyond just the syndication of the deal. Specifically, syndicates that include community VCs perform better. The broader point made by these results is that while syndication itself matters, the *interactions* it gives rise to also have incremental performance effects.

The final part of our paper examines the composition of VC communities. Here, we ask what type of VC firms join together into communities. We consider two views of community formation. At one end of the spectrum is the view that communities are formed to aggregate heterogeneous skills, effectively offering a one-stop shop to cater to a wide range of needs of young startups. Under this view, communities may be conglomerates in disguise, with soft borders in lieu of rigid organizational lines. The alternative "specialization" viewpoint suggests that venture capitalists with similar functional capabilities form communities. This could arise because of a well-known behavioral propensity for homophily (McPherson, Smith-Lovin and Cook (2001)) or because similarity may reinforce the effect of familiarity in social interactions. We let the data speak to the two possibilities by comparing the within-community variation in characteristics relative to within-community variation for the same

[6] See Brander, Amit and Antweiler (2002), Lerner (1994), Cestone, Lerner and White (2006), Sorensen (2007) and Das, Jo and Kim (2011).

characteristics for randomly formed communities.

We find some support for the soft conglomerate perspective. There is similarity among VCs within a community on the dimension of focus, suggesting that focused firms tend to syndicate together. This finding suggests that community members learn more from syndicate partners with concentrated expertise. On dimensions such as VC influence, assets under management, and portfolio company location, communities are heterogeneous. The latter findings suggest that VCs syndicate to extend their geographical reach and access deal flow available from small partners. The results suggest a nuanced view of the "birds of a feather flock together" effect of McPherson, Smith-Lovin and Cook (2001). Community members exhibit commonality in some though not all dimensions. The results appear to be more consistent with economic forces rather than passive behavioral propinquity towards the familiar driving the selection of VC partners.

The rest of the paper is organized as follows. In the next section, we describe the problem of community detection, our computational approach, and its relation to the broader literature in networks. Section 3 discusses the VC data. Section 4 presents several results on performance and community composition. This section detects and describes communities, affirms that syndicated investments perform better, shows that investments by community VCs perform better than non-community investments, even after allowing for syndication and other control variables. Communities are also characterized as being based on similarity in some dimensions and variety in others. Section 5 concludes.

2 Communities and Community Detection

2.1 Communities

The idea of community formation may be motivated by the observation that many naturally occurring complex systems actually comprise coherent subsystems ("communities") of densely connected members who interact for a functional purpose. In early work, Simon (1962) argues that community structures describe many systems in the behavioral sciences. However, research now establishes that Simon's description is apt in many other disciplines. Community detection and analysis is a thriving multidisciplinary area of research. One stream of research focuses on improving computational techniques for community detection. Another applies community detection to characterize and understand physical, biological, and social phenomena. The goal is to uncover community structures embedded in larger groups and use them to understand the functional forces underlying the larger entities.

An important application of community detection is in uncovering and understanding functional biological modules. Examples include metabolic networks of cellular organisms (Ravasz et al (2002); Duch and Arenas (2005)), and protein-protein interactions to identify protein complexes that propagate or perform specific functions (Guimera and Amaral (2005); Gao, Sun, and Song (2009); Lewis et al (2010)). Recent work in brain imaging has also uncovered community structures in the human brain. Inter-community connections appear to weaken in older people. These insights can help better understand age-related changes in brain functioning (Wu et al (2011)). In biology, community detection has been used to examine the compartmentalization of food chain webs. Understanding these structures gives insights on the stability and robustness of ecosystems when unanticipated shocks endanger species (Dunne (2006); Girvan and Newman (2002)).

In political science, community detection is used to uncover political preferences from voting patterns. Community structures reveal that political preferences can transcend traditional party lines (Porter et al (2007)). In an interesting experiment, Zachary (1977) records ex-ante social interactions between individuals in a karate club. He reports that the ex-ante interactions strongly predict ex-post community formation. The evidence on bottlenose dolphins in Lusseau (2003) suggests that communities are an evolutionary mechanism against isolation that can occur when a member is subject to random attack.[7] Other research on community structures includes mobile phone and online networks (Porter, Onnela, and Mucha (2009)), air transportation networks (Guimera et al (2005)), word adjacency in linguistics and cognitive sciences (Newman (2006)), and collaborations between scientists (Newman (2001); Duch and Arenas (2005)).

Fortunato (2009) presents a relatively recent and thorough survey of the community detection literature and its open challenges. Fortunato points out that the literature has progressed on the computational issues to the point where many methods yield similar results in practice. However, in his view, there are fewer insights on the functional roles of communities or their quantitative effect on outcomes of interest. Fortunato suggests that this is a key challenge in the literature.[8] It is also an area in which we offer progress, at least in the setting of venture capital. We detect communities and tie community formation to functional economic effects, VC exits via mergers or IPOs.[9]

[7] The lines demarcating communities in both Zachary (1977) and Lusseau (2003) are sharp and the community structures are well motivated. Thus, their datasets are now the standards for benchmarking the quality of community detection algorithms. For additional datasets, see Mark Newman's website http://www-personal.umich.edu/ mejn/netdata/. For an analysis of word clusters in product descriptions, see Hoberg and Phillips (2010).

[8] As Fortunato concludes "... What shall we do with communities? What can they tell us about a system?" He writes that "... This is the main question beneath the whole endeavor."

[9] In our study, communities organize to improve investment outcomes for firms, and by extension, for themselves. This is what one might expect of economically motivated agents. In living organisms, different community modules may serve different functional purposes.

2.2 Community Detection Algorithms

Detecting VC communities usually begins with a network graph of ties between VC firms. The network graph is represented by an adjacency matrix A, with the VCs on the rows and columns. In our setting, an element $A(i, j)$ of the matrix A represents the numbers of rounds that two VCs i and j finance together in a joint syndicate. Thus, A is a weighted matrix such that more intense transactional partnerships lead to greater weights.

The diagonal element of the adjacency matrix is zero. While this is standard in the networks literature, the assumption has an economic basis and meaning. The underlying economic assumption is that venture capitalists get no benefit of learning from community partners when they deal with themselves. We model relationships between VCs as being symmetric. The economic argument is that benefits flow to all members of a syndicate, for instance because all VC syndicate members learn from financing a portfolio firm. Thus, we model a weighted adjacency matrix that is symmetric about the main diagonal. This is an undirected graph.

The final output of the community detection techniques is a partition that identifies what community each VC firm belongs to. Some firms may not belong to any community. To assess the overall quality of the partition that generates communities, we compute a modularity score Q for the partition (see, e.g., Newman (2006)):

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{d_i \times d_j}{2m} \right] \cdot \delta(i, j) \tag{1}$$

In equation (1), A_{ij} is the (i, j)-th entry in the adjacency matrix, i.e., the number of syndicate transactions in which VC firm i and j jointly participated, $d_i = \sum_j A_{ij}$ is the total number of rounds that VC firm i participated in with other VCs (or, the degree of i) and $m = \frac{1}{2} \sum_{ij} A_{ij}$

is the sum of all edge weights in matrix A. The function $\delta(i,j)$ is Kronecker's delta. It takes the value of 1.0 if the nodes i and j are from the same community, and is zero otherwise. In equation (1), modularity Q takes values from -1 to $+1$. When $Q > 0$ it means that the number of connections within communities exceeds that between communities.

Appendix A gives a simple example for which we compute Q. While modularity optimization is the most popular technique in recent applications (Fortunato (2009)), it suffers from known drawbacks. For instance, Fortunato and Barthelemy (2007) show that it cannot identify tiny communities. We require a minimum community size of three members. We also require that the diameter of a community from end-to-end not exceed one-fourth that of the entire network. As it turns out this constraint was never binding for our data set.

The computational algorithm to pick the best partition is far from straightforward. As Fortunato (2009) discusses (see his Section III), community detection is an **NP**-hard problem for which there are no known exact solutions beyond very small systems. For large scale datasets, community detection algorithms are of three types. *Graph partitioning* places VC firms into groups of equal size, such that the number of connections between communities is minimized. We do not use this approach as there is no economic reason to believe that VC communities should have equal size. *Partitional Clustering* presets the number of communities and minimizes a loss function to detect them. We do not use this approach as we do not have a reasonable economic basis to presuppose an arbitrary number of communities. *Hierarchical Clustering*, our approach, starts with a few large communities and breaks these down into smaller ones based on the density of connections within and outside groups.

The ideas in these broad classes of algorithms have been relaxed, extended, and implemented in many different ways. Leskovec, Kang and Mahoney (2010) compare several of them. Community detection algorithms can also be classified into "agglomerative" and "divisive." The divisive class of community detection algorithms is a top-down approach. It starts by assuming the entire graph is one community. It then breaks down the graph into smaller units. An example of such an algorithm is the "fast-greedy" algorithm of Girvan and Newman (2002). Divisive algorithms have a tendency to produce communities that are often too large especially when there is not an extremely strong community structure.

Agglomerative algorithms, like the "walktrap" algorithm we use, begin by assuming all nodes are separate communities. Nodes are then collected into communities that form a partition on the graph. This is a bottom-up approach that builds larger communities from smaller ones. Among the quickest algorithms of this nature are dynamic methods based on random walks. The essential intuition of these algorithms is that if a random walk enters a strong community, it is likely to spend a long time inside it before finding a way out. Setting the maximal number of steps in the random walk is necessary to implement the algorithm (Pons and Latapy (2005)).

2.3 Community Detection and Social Network Analysis

Community analysis is part of a growing literature on the role of social connections in economics and finance that extends pairwise relationships to higher-level group structure. One strand of this literature focuses only on the pairwise connections between individuals arising out of common educational alma mater or employers, often exploiting the Boardex database. Cohen, Frazzini, and Malloy (2008a, 2008b) show that such educational connections result in economically valuable information flows, a point also made by Shue (2011). Hwang and Kim (2009) show that pairwise employment connections between boards and CEOs affect CEO compensation, while Ishii and Xuan (2009) study how employment connections impact M&A activity.

Other studies examine the entire "social network" constructed out of the pairwise connections. The essential idea of these papers is that agents can derive economic benefits not only from their direct connections but also indirectly through the connections of the agents they are connected to. Thus, the literature focuses on the aggregate connectedness or the "rolodex" of an individual. Hochberg, Ljungqvist and Lu (2007) find that portfolio firms with more central VC investors are more likely to successfully exit via IPO or merger. Engelberg, Gao and Parsons (2000) find that CEO centrality is related to total compensation.

While we include and control for centrality, we emphasize that the community metric has a rather different flavor and economic motivation. Centrality is a construct for the aggregate influence of an individual in a network. Influence comes from having many connections, or (recursively) being connected to particularly well connected individuals, or possessing particularly critical connections. These connections often arise out of the individual's personal skill or resources. Community membership, on the other hand, is a *group* attribute that is motivated by the role of interactions. Neither is a subset of the other.

Communities are also different from a construct called "clique." A clique is a subset of nodes in a network that are all connected to each other or within a given distance from each other (called an n-clan or n-clique) but no node outside the clique is connected to all nodes inside the clique.[10] Cliques are self-contained and self-referential clusters. Thus, it is too restrictive to apply to the VC context, where boundaries are soft and alliances are probabilistic propensities. Clique formation can be detrimental because lack of interaction across cliques impedes information flows. In the VC setting, interaction across communities may be discouraged for competitive reasons, yet to some extent may be beneficial as information exchange leads to better decisions by communities.

[10]Sub-graphs of diameter n are also known as n-clubs. See the definitions in Mokken (1979).

3 Data

We use venture-backed investment rounds' data obtained from Thomson Financial's Venture Economics database. Recent studies using the data include Kaplan and Schoar (2005) and Lindsey (2008). We analyze VC investments made from 1980 to 1999. We start in 1980, around the time the VC industry started growing rapidly (see, e.g, Figure 1 in Gompers and Lerner, 2001). Our dataset ends in 1999 to allow at least 10 years from investment to outcome. We drop the cases in which the database does not disclose a VC firm name, or lists the VC firm as an angel, individual or management. We only consider domestic investments by U.S.-based VC funds in non-buyout deals.

We sample IPO firms using data from Thomson Financial's SDC Platinum. We match companies by their cusip identifiers, cross-check the matches against actual names, and further hand-match the names with those in the Venture Economics database. 1,470 ventures in our sample exited via IPOs. We also obtain M&A data from Thomson Financial's SDC Mergers and Acquisitions database. We conduct similar hand-matching of portfolio company names in the Venture Economics database. We find that there are 3,545 exits via mergers in our sample.

Table 1 gives descriptive statistics for our sample at the level of an individual venture capitalist. There are a total of 1,962 unique VC firms in the sample period. On average, a VC firm invests $595 million (median = $110 million) in about 22 portfolio firms and 48 rounds. The average investment money raised per round is $19.47 million (median = $10.56 million). The total funds raised by a VC amount to about $128 million. Three in every four deals of an average venture capitalist are syndicated. One-third of each VC firm's deals are for early stage firms. The mean age of each VC at the time of its last investment in our sample is a little less than 10 years. There are 127 Metropolitan Statistical Areas (MSAs) covered in our data set. There are 14 VC firms per MSA on average.

4 Results

4.1 Community Detection

We use rolling 5-year windows to identify communities. Thus, the first community is based on VC investments from 1980 to 1984, the second community is based on 1981-1985 investments, and so on. We choose a window length of 5 years for community detection. This is a compromise between allowing a long time period to permit community formation and detecting it, and using excessively long periods that may contain stale information.

The community detection algorithm identifies a large number of communities, varying from a minimum of 12 in 1987-91 to a maximum of 35 communities in 1995-99. In each window, several VC firms do not belong to communities. Between 81 and 183 VC firms, representing about 20% of the VCs active in any time period belong to communities. The median community has 13 members.[11] Figures 3–6 depict communities for four non-overlapping 5-year windows, viz., 1980–1984, 1985–1989, 1990–1994, and 1995-1999, with members of the largest three communities shown in different colors. The upper plots in each figure show the entire VC network. To present a less cluttered view of the network, the lower figure plots the largest community embedded within all communities of at least 5 members. We see that connections within the largest community are much greater than connections across communities, thereby visually affirming the definition of a community. In Figure 3 all large communities are connected to one another, but in Figures 4 and 5 there are satellite communities that are large but disconnected from all other communities. Figure 6 shows satellite

[11]We note that the community/non-community status of a VC firm is not fixed for the whole sample period. A VC firm may be part of a community during one 5-year window but is not necessarily a member of any community in all subsequent windows.

communities in the upper plot, but the largest communities are well connected to the rest of the communities. In the lower plot, all large communities are connected, but a few peripheral ones at the edge of the network are relatively isolated.

Table 2 shows two sample communities that illustrate community formation. These samples are drawn from earlier time periods and identify communities in which Stanford University (arbitrarily chosen) was a member. These two communities comprise storied Silicon Valley VC firms well-known to practitioners. The size and composition of the communities is not identical from period to period. In one, there are 17 VCs as against 15 in the other. While the choice of these community examples was driven by Stanford University's community membership, there were two other VCs who were also in both communitites, namely Sequoia Capital and Mohr Davidow Ventures, two large VC firms. The remaining 26 VCs belonged in either the first or the second of the two communities.

A VC who belongs in a community with another specific VC need not share its community membership with the other in the next period. In fact, the VC may or may not be a part of a community in another period. On average, a VC who is ever in a community continues to be in a community for 88 percent of the years it is in our sample. Table 3 provides information on community VCs from each rolling window on their continued community membership in the next 1, 3 and 5 year rolling windows. Community participation tends to be sticky. An average of 90 percent of community VCs continue to be part of a community in the next rolling window. Three out of four community VCs continue to be in a community after 5 years. The results indicate significant stability among VCs who become part of a community.

We next consider stability in the *composition* of communities. We consider 5-year rolling windows to allow for sufficient time for communities to evolve and change their composition. It is not surprising that there is some change in the nature of one's preferred partners over time. For instance, one's circle of friends today is different from one 10 years ago. Quantifying the changes in group composition is, however, non-trivial. A community of three VCs could stay unchanged in the next rolling window. It could also break up in many different ways. Each of the 3 members could go their separate ways, or a pair could be together in a community the next year but not the third VC. The possible combinations multiply with the size of the original community.

To quantify community stability, we use the Jaccard index. For any pair of sets, the Jaccard index is defined as the number of members in the intersection of the two sets divided by the number of members in the union of the sets. In other words, if A and B are a pair of VC communities, the Jaccard index for the two is given by $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$. Because members from each VC community could be spread across multiple communities in the next rolling window, we also need to measure the overlap between a community in a given period and all communities in the next window. Let $A^t = A_1, A_2, ..., A_m$ be the set of m communities at time t, and $B^{t+1} = B_1, B_2, ..., B_n$ be the set of n communities at time $t+1$. For each community A_i, we determine a composite Jaccard measure, $JC(A_i, B) = Mean_j(J(A_i, B_j)|J(A_i, B_j) > 0)$ for all $j = 1, 2, ..., n$. The value of the measure would depend on the number of communities in the next period and how the community members are dispersed in the next period. For instance, if there was one community both in periods t and $t+1$, and all the period t members remained in the same community next period, $JC = 1$.

Table 4 presents results for the composite Jaccard measure, averaged for each community in each rolling window, when paired with communities in the adjacent rolling window. In order to determine a benchmark against which to evaluate the stability of communities, we generate random communities. We mimic the number of communities in each 5-year rolling window as well as the size of each community in our sample. We determine the composite Jaccard measure for these random communities and bootstrap the communities to generate an empirical distribution of the Jaccard index for random communities. For each rolling window, the Jaccard measure of our community is greater than that of bootstrapped communities, at the 1% level of significance. Thus, communities identified in our data are significantly more stable than would occur by chance.

4.2 Univariate Comparisons

Based on the communities identified over a 5-year window, we analyze the performance of VC investment in the following year. Given that the first window is 1980-84, our performance analysis starts from year 1985. Table 5 describes descriptive statistics for our sample organized by investment round since 1985. Syndicated rounds have multiple participating venture capitalists with differing characteristics. For instance, young venture capitalists may join together with older ones in a round. We use the maximum of the VC-level variables to construct round-level VC variables. Thus, a corporate VC dummy takes the value 1.0 if at least one VC in a round of financing is a corporate VC, and zero otherwise. Panel A shows that 15,220 (45%) of the 33,924 investment rounds have at least one community VC firm. 33% of Round 1 financings have at least one community venture capitalist. 44% of the subsequent financing rounds are community rounds.

Syndication is common in VC investment. In our sample, 14,897 out of 33,924 rounds (about 44%) are syndicated and these account for 66% of proceeds. 10,056 out of 14,897 or 67% of syndications are community rounds. Early stage rounds account for about a third of the sample. 45% of these are community rounds. Close to one-half of the investment rounds is in the geographical clusters in California or Massachusetts, reflecting a concentration of VC investments in these states and their representation in VC databases (Kaplan, Sensoy and

Stromberg (2002)). Community VC-based financing rounds account for about 60% of these rounds. 3,372 rounds have corporate VC participation and 7,586 rounds have a financial VC firm. In each case, communities account for 58% of the financing rounds.

Venture Economics classifies VC portfolio firms into 10 industries. We report these data in Panel B. The software industry with 20% accounts for the largest share of financing rounds in our sample, followed by medical or health firms, communications and media, and internet firms. Interestingly, community VC is more likely for the more risky and complex business models characteristic of software businesses and less likely for consumer product or industrial businesses. The finding indicates that VC firms rely more on familiar partners in riskier industries. We control for such variation by incorporating industry fixed effects in the multivariate analyses.

Panel C in Table 5 describes key characteristics across rounds. There is greater investment in rounds with a community VC ($48 million) than in rounds with no community VCs ($29 million). Besides higher investment per round, community rounds tend to have more VC firms than rounds with no community VC. This may reflect the greater representation of community VC firms in syndicated rounds. However, even within the subsample of syndicated rounds, rounds with a community VC tend to have 4 VCs on average compared to 3 in rounds with no community VC. This pattern holds for early stage rounds and initial financing rounds.

4.3 Performance

We test whether a portfolio firm sourcing capital from a community VC experiences better ex-post performance. Following the VC literature (e.g., Lindsey (2008)), our primary measure of success is exit via merger or IPO. To the extent that a venture's success represents a signal of an investor's success, exits can also proxy for VC performance.

We lag the community variable relative to the window over which we identify community to predict performance. For instance, we construct communities based on VC syndication patterns between 1980 and 1984. We use these data to classify investments in 1985 as coming from community or non-community VCs. Likewise, the next window for community construction is 1981 to 1985 and the community classification is applied to VC investments in 1986. This approach follows the strategy of Hochberg, Ljungqvist, and Lu (2007). As they discuss, the lag structure results in a conservative structure where past 5-year syndication patterns predict outcomes over windows of several years into the future.

In terms of performance at the round level, Panel D of Table 5 indicates that 12,604 (or 37%) of financing rounds exit. IPOs account for less than a third of these and about 11% of rounds financed. In community rounds, 14% exit through IPOs and 29% exit through mergers compared to 9% and 24% for non-community VC rounds, respectively. We find a similar pattern when considering exits classified by the number of portfolio companies rather than number of rounds of financing. 13% of companies sourcing funds from a community VC firm at least once have IPO exits compared with 7% of companies who never have community VC financing. As an alternative measure of success, we consider a round to be successful if a venture raises at least one round of financing subsequently within the next five years. A higher proportion of all rounds with a community VC were successful (78%) compared to rounds with no community VC (65%).

4.4 Multivariate Specifications

This section considers two specifications for investing success. Following, e.g., Hellmann and Puri (2002), we consider the time to exit using a Cox proportional hazards specification.

Because the Cox model allows a flexible non-parametric baseline hazard, it is a popular choice for modeling duration. In the Cox model, we report the exponentiated hazards ratio. A ratio greater than 1.0 for a variable indicates that the variable increases the time to exit, while a ratio less than 1.0 indicates that the variable lowers the time to exit. Following Lindsey (2008), we also consider a probit model in which success is defined as an exit either through an M&A transaction or an IPO within 10 years of the investment round.

4.4.1 Explanatory Variables

Our primary interest is how sourcing funds from a community VC is related to exit. The key variable of interest is the *Community Dummy*, which takes value 1.0 if the round has at least one community VC, and zero otherwise. We include several controls that are suggested by the recent VC literature (e.g., Lindsey, 2008).

Agency problems and information asymmetry are more likely when portfolio companies are in the early stages of their life cycle. These problems may adversely affect performance. Accordingly, we include the variable *Early Stage*, which takes value of 1.0 if the financing is in an early stage round, and zero otherwise. The literature in economics suggests that there is geographical clustering or agglomeration that conveys economic benefits to firms located in geographic clusters (e.g., Porter (1998); Glaeser (2010)). In the context of venture financing, well-known geographic clusters are in California (CA) and Massachusetts (MA). We include a geographic cluster dummy variable that takes the value 1.0 if a portfolio company is located in either CA or MA, and zero otherwise.

We include controls for the characteristics of venture capital firms participating in a financing round. In particular, we control for whether a VC in a financing round is a corporate VC arm or not. Following Hellmann, Lindsey, and Puri (2008), VC arms of financial institutions may have systematically different success rates. Thus, we also control for financial institution venture capitalists. A long stream of research going back to at least Lerner (1994) finds that syndication is a key determinant of success. Accordingly, we include a control for whether a round is syndicated or not.

A number of papers in the VC literature stress the role of VC experience and skill.[12] For instance, Kaplan and Schoar (2005) identify the importance of experience in VC fund performance. Sorensen (2007) points to the greater likelihood of an IPO of a portfolio company that is funded by a more experienced VC. We control for a VC's skill in maturing its portfolio company using *IPO Rate*, or the rate at which it is able to take its portfolio companies public.[13] Hochberg, Ljungqvist and Lu (2007) find that a VC firm's connectedness often subsumes traditional measures of VC experience in explaining performance. Thus, we include the lead VC's eigenvalue centrality based on investments from $t-1$ through $t-5$. Following Lindsey (2008), we define *Experience* as the average age of the participating VCs as of the year before the financing round.[14]

We consider two more measures of VC experience. Given the particular challenges associated with early stage financing, a VC with experience in early stage may be considered to be different in terms of value and skills than an investor without such experience. Such differences in investment focus could also affect company performance. We define *Early Stage Focus* as the proportion of companies that the participating VCs invested at an early stage until the year prior to the financing round. Similarly, each industry presents its own

---

[12]For a recent review, see Krishnan and Masulis (2011).

[13]In calculating the IPO rate, we follow Krishnan and Masulis (2011) who find strong evidence that the number of completed IPOs in a VC's portfolio over the prior 3 calendar years relative to the number of companies it actively invested in is a predictor of portfolio company performance.

[14]Our definition modifies Lindsey's definition on two fronts. First, we consider age based on the VC firm's founding year rather than its entry into Venture Economics. Second, we consider a VC's experience based on time periods prior to the financing round in question.

challenges. Skills and expertise required for a biotechnology company can be different from those necessary for investing in a software product company. We define *Industry Focus* as the proportion of companies funded by the participating VCs in the same industry as the portfolio company until the year prior to the financing round.

Finally, as in the context of portfolio companies, there may be benefits of agglomeration for VC firms too which may impact portfolio company performance. We include a geographical cluster control for the VCs, which takes the value 1.0 if at least one of the participating VCs is located in California or Massachusetts, and zero otherwise. All our specifications include fixed effects for the industry that the portfolio company belongs to and the year of the financing round.

4.4.2 Estimates

Table 6 reports the Cox and probit estimates. In the Cox model (i.e., specification (1)), we find that the hazard ratio for the *Community Dummy* is greater than one, at 1.11, and is significant at the 1% level. The estimate shows that having a community VC in a financing round shortens the time to exit by 11%.

Among the controls, both the company-level variables are statistically significant. The coefficient for *Early Stage* is less than one, suggesting that early stage deals may take longer to mature and exit. Companies in geographical clusters of California and Massachusetts are likely to exit sooner, perhaps due to improved resource flows and better decision-making arising out of agglomeration (Porter (1998, 2000); Glaeser (2010)). Ownership of VC firms matters. In particular, a financing round with at least one corporate VC or financial institution VC is likely to experience speedier exit.

We find that syndicated ventures tend to exit faster. A VC firms' reputation for taking its companies public, measured by the *IPO Rate*, is not statistically significant in explaining speed of exit. As in Hochberg, Ljungqvist, and Lu (2007), we find that a more centrally networked VC facilitates faster exit, though at the 10% level of significance. VC experience, in terms of their age at the time of financing, early stage focus or specific industry focus, is not statistically significant. However, funding from VCs who lie within the California-Massachusetts cluster facilitates quicker exit for portfolio companies. The important finding in Table 6 is that VC community is significant even after including these controls.

Specification (2) in Table 6 reports probit estimates that model the probability of exit within 10 years. Most of the results from the Cox model go through in the probit specification. One difference is that the IPO rate is significant at the 10% level in the probit model but not in the hazards model. However, the community variable continues to remain significant and is associated with a greater likelihood of exit. We also estimate but do not report univariate specifications with community dummy alone and partial specifications that include it with subsets of controls. We note that the VC community variable is significant in these models as well with a similar or higher exponentiated hazards ratio.

4.5 Performance By Round

In this section, we consider follow-on financing as a measure of success. Follow-on rounds of financing involve reassessment of the portfolio company. New investors are often brought in, incumbent VCs increase their investment, and both sets of investors have the opportunity to re-evaluate and reconsider the progress of the portfolio firm. Thus, attracting follow-on funding can be viewed as an alternative metric of success. Cochrane (2005) suggests that round-by-round financing data can be used to construct VC performance metrics.

We rely on Venture Economics codes to specify the round number. These data are not without noise. In some instances, the first available round of financing available in the database may not be round number one and round numbers may be missing between rounds. We take a conservative approach. We only consider those rounds that are identifiably numbered and do not have missing data for subsequent rounds when one exists. These criteria reduce the sample of first three rounds from 22,683 to 22,271 rounds.[15]

Table 7 shows the round-by-round results with both the Cox and the probit specifications. Community VC accelerates the progression to a future round of financing in the earlier rounds (rounds 1 and 2) but not in the later round (round 3). Community VCs appear to matter less when the firms are more mature in their life cycle.

Among the control variables, the coefficient on the early stage dummy variable is positive and statistically significant. One interpretation of this finding is that staged financing is more prevalent at the early stage firms given the greater informational issues with these firms (Cornelli and Yosha (2003)). Thus, VC firms manage early stage financing through more frequent injections of smaller amounts of capital.

Neither a corporate VC nor a financial institution VC helps with subsequent financing rounds in the initial stages. In fact, financial institution VCs have a negative and significant effect, perhaps because VC arms of financial firms are structurally different (Hellmann, Lindsey and Puri (2008)). As before, syndicated rounds have a higher chance of obtaining future funding and do so sooner. VCs' reputation for taking portfolio companies public (IPO rate) seems to have an adverse effect on subsequent funding after round 2. Eigenvector centrality is significant in round 1 in the Cox specification at 10% significant level but not in the probit specification. However, it matters in round 3 under both specifications, suggesting that centrality and community play complementary roles. Perhaps thick rolodexes are more

[15]The 412 rounds we lose due to missing sequential round numbers are spread evenly through the sample period and in both early and non-early stages.

critical in later stages when it provides firms access to a broader set of resources such as personnel or strategic contacts.

Portfolio companies located in California or Massachusetts experience a higher likelihood of next-round financing except in round 2. VC firms belonging to the geographical clusters in California and Massachusetts have a positive effect in earlier round but not in later rounds. VC experience, in terms of the participating VCs' age, does not help and even impedes progress in initial rounds. Early stage focus is associated with a greater likelihood of and faster follow-on financing or exit. Thus, firms that declare specialization in early stage ventures appear to accelerate a firm's progress to a next round of financing. In each of the specifications, the industry focus on the participating VCs is statistically insignificant. In any event, the key result is that community VC remains significant in facilitating follow-on financing, particularly in the initial rounds.

4.6 Community Composition

Our previous results indicate that venture capitalists form communities with preferred partners. However, they say little about the composition of individual communities, or the types of partners that they prefer. We address this issue next.

In principle, communities could consist of venture capitalists with similarity or diversity in attributes. The case for diverse attributes rests on the view that VC investing requires skills along multiple dimensions to assess investments and to manage them to maturity. Partners with heterogeneous attributes can extend the skill set of an individual venture capitalist or expand its investment possibilities. On the flip side, communities could also be homogeneous. The behavioral literature (e.g., McPherson, Smith-Lovin, and Cook, 2001) suggests that like tend to affiliate with like. In the VC context, Gompers and Lerner (2004) argue that VC firms will often not syndicate unless the investment is vetted by a partner they trust. To the extent VC firms can better assess other VCs in similar space, VC communities may tend to be homogeneous. Cestone, Lerner, and White (2006) argue that vetting is most useful to a VC when the partner is of similar caliber. If syndication manifests the need for validation by partners, it is plausible that similar types of VC firms form communities, especially when sorted by VC skill or experience in a relevant sector. Whether heterogeneity or homogeneity in attributes dominates comes down to an empirical question. Table 8 reports the results.

Panel A in Table 8 reports the average characteristics of VC community members. This table sheds light on the types of *venture capitalists* that cluster into communities. The first column of results reports the mean characteristic for our sample of communities. The second column reports the mean characteristics for bootstrapped communities. These are generated by randomly picking a random set of VCs and assigning them to communities, with the number of communities and their size fixed to the actual number of communities identified by our algorithm. The third column reports the p-value for the mean characteristics based on the simulated distribution of the characteristics for the bootstrapped communities. We find that older, larger, prestigious (with high centrality) VC firms concentrated across states tend to form communities. Community members tend to be more focused in terms of industries and location of portfolio companies but not in terms of stage of investment.

Panel B reports the variation in characteristics *within* the community compared to the variation for bootstrapped communities. This panel sheds light on the type of *communities* formed by venture capitalists. Communities are more concentrated in terms of industry and stage focus, consistent with the idea that focused firms are more likely to deal with and learn from each other. On the other hand, communities have greater variation in assets under management, in VC influence, and diversity in the portfolio company state. These results suggest that communities permit VCs to extend their reach in generating new investments.

Our results provide a partial reconciliation of the contradictory findings reported by Du (2009) and Hochberg, Lindsey, and Westerfield (2011). Du reports that venture capitalists are more likely to syndicate with partners while Hochberg et al. find that VCs syndicate with dissimilar partners. Our results suggest a somewhat nuanced view of the homophily versus diversity debate. We find that syndicate partnerships tend to exhibit commonality in some dimensions but not others. The results thus reconcile and support contradictory economic forces: the second opinion hypothesis of Cestone, Lerner, and White (2007) that pushes similar VC firms to partner together and the resource complementarity view of the strategic alliance literature in which partnerships incentivize the flow of unique resources (e.g., Robinson and Stuart, 2007; Robinson, 2008). More generally, the results reinforce the point made by Harrison and Klein (2007) that diversity is a multidimensional construct that is not summarized as a single scalar variable. The larger point made by our results is that characteristics-based similarities seem to be based on economic roots rather than passive behavioral propinquity for the familiar.

5 Conclusion

Syndication is a pervasive feature of venture capital financing. Over the course of its life, a venture capital firm is likely to form syndicates several times and do so with different partners. However, not all VC firms associate with each other in syndicates. Nor are the syndicate partnerships formed randomly. Instead, VC firms tend to exhibit associative properties in which they tend to syndicate more with some partners than with others. This leads to clusters that we term venture capital *communities*.

We examine community formation in the venture capital industry and characterize its economic consequences. We employ flexible community detection techniques that accommodate real world features of the VC market including flexibility in the number, size, and composition of communities and communities with porous borders. Using 20 years of venture financing data, we find robust evidence of VC community formation throughout our sample time period. Communities do not pool heterogeneous venture capitalists into soft conglomerates but consist of functionally similar VCs. Community VCs are also associated with positive economic outcomes. Firms that source capital from community VCs are more likely to experience successful exit, after controlling for other plausible explanatory variables such as syndication, that also plays an important role. The evidence is most consistent with the view that collaborations play a key economic role in the VC market, and that, repeated interactions with familiar partners add further value to venture capitalists and their portfolio firms.

Our study contributes to the growing literature on social networks in finance and economics. Social networks are usually viewed as being important because they endow positions of influence to central individuals on a network. For instance, a well connected CEO can benefit herself or her firm because more connections provide greater access to network resources. Our study emphasizes a complementary point: networking is also beneficial because it facilitates social interactions. These interactions can create value. Interactions can enhance learning, improve soft information flows, or foster trust and reciprocity, leading to better economic outcomes. These benefits matter especially in environments of risk, uncertainty, and asymmetric information that characterize the VC market. The performance effects of communities can be viewed as a manifestation of the benefits of interactions.

We make two related points on this issue. First, our paper suggests that VC skill is not entirely endowed. Some part of it is learnt or upgraded through the syndication process, as suggested by Sorensen (2008). Second, the VC literature has long emphasized that syndication is a critical determinant of investment outcome. While we confirm this finding, our additional point is that the structure and composition of syndicates also matter. More broadly, syndication has benefits beyond its value for a specific deal or a firm. Syndication also gives rise to interactions that can benefit future investments by syndicate members.

While our study focuses on community formation in venture capital, it is also interesting to examine community formation in other contexts. For example, repeated collaborations can create communities in other areas such as syndication in investment banking, underwriting, or lending. An interesting question is whether communities in these other areas are associated with better economic outcomes or whether they are motivated by economically neutral behavioral preferences for the familiar or transaction cost motives to lower administrative or operational costs. More generally, it is interesting to understand whether community formation arising out of evolutionary processes in biology and the natural sciences are also mimicked in settings where agents interact for economic benefits.

A Calculating Modularity

In order to offer the reader a better sense of how modularity is computed in different settings, we provide a simple example here, and discuss the different interpretations of modularity that are possible. The calculations here are based on the measure developed in Newman (2006). Since we used the `igraph` package in R, we will present the code that may be used with the package to compute modularity.

Consider a network of five nodes $\{A, B, C, D, E\}$, where the edge weights are as follows: $A : B = 6$, $A : C = 5$, $B : C = 2$, $C : D = 2$, and $D : E = 10$. Assume that a community detection algorithm assigns $\{A, B, C\}$ to one community and $\{D, E\}$ to another, i.e., only two communities. The adjacency matrix for this graph is

$$\{A_{ij}\} = \begin{bmatrix} 0 & 6 & 5 & 0 & 0 \\ 6 & 0 & 2 & 0 & 0 \\ 5 & 2 & 0 & 2 & 0 \\ 0 & 0 & 2 & 0 & 10 \\ 0 & 0 & 0 & 10 & 0 \end{bmatrix}$$

The Kronecker delta matrix that delineates the communities will be

$$\{\delta_{ij}\} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

The modularity score is

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{d_i \times d_j}{2m} \right] \cdot \delta_{ij} \tag{2}$$

where $m = \frac{1}{2} \sum_{ij} A_{ij} = \frac{1}{2} \sum_i d_i$ is the sum of edge weights in the graph, A_{ij} is the (i, j)-th entry in the adjacency matrix, i.e., the weight of the edge between nodes i and j, and $d_i = \sum_j A_{ij}$ is the degree of node i. The function δ_{ij} is Kronecker's delta and takes value 1 when the nodes i and j are from the same community, else takes value zero. The core of the formula comprises the modularity matrix $\left[A_{ij} - \frac{d_i \times d_j}{2m} \right]$ which gives a score that increases when the number of connections within a community exceeds the expected proportion of connections if they are assigned at random depending on the degree of each node. The score takes a value ranging from -1 to $+1$ as it is normalized by dividing by $2m$. When

$Q > 0$ it means that the number of connections within communities exceeds that between communities. The program code that takes in the adjacency matrix and delta matrix is as follows:

```
#MODULARITY
Amodularity = function(A,delta) {
    n = length(A[1,])
    d = matrix(0,n,1)
    for (j in 1:n) { d[j] = sum(A[j,]) }
        m = 0.5*sum(d)
        Q = 0
        for (i in 1:n) {
            for (j in 1:n) {
                Q = Q + (A[i,j] - d[i]*d[j]/(2*m))*delta[i,j]
            }
        }
        Q = Q/(2*m)
    }
```

We use the R programming language to compute modularity using a canned function, and we will show that we get the same result as the formula provided in the function above. First, we enter the two matrices and then call the function shown above:

```
> A = matrix(c(0,6,5,0,0,6,0,2,0,0,5,2,0,2,0,0,0,2,0,10,0,0,0,10,0),5,5)
> delta = matrix(c(1,1,1,0,0,1,1,1,0,0,1,1,1,0,0,0,0,0,1,1,0,0,0,1,1),5,5)
> print(Amodularity(A,delta))
[1] 0.4128
```

We now repeat the same analysis using the R package. Our exposition here will also show how the walktrap algorithm is used to detect communities, and then using these communities, how modularity is computed. Our first step is to convert the adjacency matrix into a graph for use by the community detection algorithm.

```
> g = graph.adjacency(A,mode="undirected",weighted=TRUE,diag=FALSE)
```

We then pass this graph to the walktrap algorithm:

```
> wtc=walktrap.community(g,modularity=TRUE,weights=E(g)$weight)
> res=community.to.membership(g,wtc$merges,steps=3)
```

```
> print(res)
$membership
[1] 0 0 0 1 1

$csize
[1] 3 2
```

We see that the algorithm has assigned the first three nodes to one community and the next two to another (look at the membership variable above). The sizes of the communities are shown in the size variable above. We now proceed to compute the modularity

```
> print(modularity(g,res$membership,weights=E(g)$weight))
[1] 0.4128
```

This confirms the value we obtained from the calculation using our implementation of the formula.

Modularity can also be computed using a graph where edge weights are unweighted. In this case, we have the following adjacency matrix

```
> A
     [,1] [,2] [,3] [,4] [,5]
[1,]    0    1    1    0    0
[2,]    1    0    1    0    0
[3,]    1    1    0    1    0
[4,]    0    0    1    0    1
[5,]    0    0    0    1    0
```

Using our function, we get

```
> print(Amodularity(A,delta))
[1] 0.22
```

We can generate the same result using R:

```
> g = graph.adjacency(A,mode="undirected",diag=FALSE)
> wtc = walktrap.community(g)
> res=community.to.membership(g,wtc$merges,steps=3)
> print(res)
$membership
```

```
[1] 1 1 1 0 0

$csize
[1] 2 3

> print(modularity(g,res$membership))
[1] 0.22
```

A final variation on these modularity calculations is to use a Kronecker delta matrix that has diagonal elements of zero. In the paper we use the first approach presented in this Appendix.

B Variable Definitions

Variable	Description
Dummy Variables	
Community	Equals 1.0 if there is at least one community VC in the financing round and zero otherwise
Early Stage	Equals 1.0 if the round is an early stage financing and zero otherwise.
Company Geographical Cluster	Equals 1.0 if the portfolio company funded by the VC is in the state of California or Massachusetts and zero otherwise.
Corporate VC	Equals 1.0 if there is at least one venture capitalist who is the corporate VC arm of a firm.
FI VC	Equals 1.0 if there is at least one financial institution VC in the round
Syndicated	Equals 1.0 if the round is syndicated, zero otherwise
VC Geographical Cluster	Equals 1.0 if at least one participating VC is in the state of CA or MA
Other Variables	
IPO Rate	natural log of one plus the average of each participating VC's ratio of IPOs to number of portfolio companies in the last three years prior to the financing round
Centrality	lead VC's eigenvector centrality, normalized for the sample in each specification
Experience	natural log of one plus the average age, in years, of the participating VCs from their founding until the year prior to the financing round
Early Stage Focus	natural log of one plus the proportion of companies that the participating VCs invested at an early stage until the year prior to the financing round
Industry Focus	natural log of one plus the proportion of companies funded by the participating VCs in the same industry as the portfolio company until the year prior to the financing round.

References

Allen, J.W., Phillips, G.M. (2000). Corporate equity ownership in the U.S.: Product market alliances, investment and contracting costs, *Journal of Finance* 55, 2791-2815.

Bottazzi, L., Da Rin, M., Hellmann, T. (2011). The importance of trust for investment: Evidence from venture capital. *Working Paper*.

Brander, J. A., Amit, R., Antweiler, W. (2002). Venture-capital syndication: Improved venture selection vs. the value-added hypothesis, *Journal of Economics and Management Strategy*, v11, 423-452.

Bygrave, W., Timmons, J. (1992). Venture Capital at the crossroads, Harvard Business School Press.

Cai, Jian (2010). Competition or Collaboration? The Reciprocity Effect in Loan Syndication, *Federal Reserve Bank of Cleveland Working Paper* 09–09R.

Casamatta, Catherine (2003). Financing and advising: Optimal financial contracts with venture capital, *Journal of Finance* 58, 2059-2085.

Centola, Damon (2010). The spread of behavior in an online social network experiment, *Science* 329, 1194–1197.

Cestone, Giacinta., Lerner, Josh, White, Lucy (2006). The design of syndicates in venture capital, *Harvard Business School Working Paper*.

Cochrane, John (2005). The risk and return of venture capital. *Journal of Financial Economics* 75, 3-52.

Cohen, Lauren, Frazzini, Andrea, Malloy, Christopher (2008a). The small world of investing: Board connections and mutual fund returns, *Journal of Political Economy* 116, 951–979.

Cohen, Lauren, Frazzini, Andrea, Malloy, Christopher (2008b). Sell-Side school ties, forthcoming, *Journal of Finance*.

Cornelli, F., Yosha, O. (2003). Stage financing and the role of convertible securities, *Review of Economic Studies* 70, 1-32.

Das, Sanjiv., Jo, Hoje, Kim, Yongtae (2011). Polishing diamonds in the rough: The sources of syndicated venture performance, *Journal of Financial Intermediation*, 20(2), 199-230.

Du, Qianqian (2009). Birds of a feather or celebrating differences? The formation and impact of venture capital syndication, *University of British Columbia Working Paper*.

Duch, J., Arenas, A. (2005). Community detection using extremal optimization, *Physics Review E* 72, 027104.

Dunne, Jennifer (2006). The network structure of food webs, Chapter 5 in *Ecological Structure: Linking Networks to Dynamics in Food Webs*, Edited by M. Pascual and J. A. Dunne, Oxford University Press.

Engelberg, Joseph., Gao, Pengjie, Parsons, Christopher (2000). The value of a rolodex: CEO pay and personal networks, Working Paper, University of North Carolina at Chapel Hill.

Fehr, E., Schmidt, K. (2005). The economics of fairness, reciprocity and altruism, in *Handbook of Altruism, Gift Giving and Reciprocity*, North Holland.

Fortunato, S., Barthelemy, M. (2007). Resolution limit in community detection, *Proceedings of the National Academy of Science of the USA* 104 (1), 36-41.

Fortunato, S. (2009). Community detection in graphs, *arXiv*:0906.0612v1 [physics.soc-ph].

Gao, L., Sun, P.G., Song, J. (2009). Clustering algorithms for detecting functional modules in protein interaction networks, *J Bioinform Comput Biol.* 7(1), 217-42..

Gertler, M.S. (1995). Being there: proximity, organization and culture in the development and adoption of advanced manufacturing technologies, *Economic Geography* 7(1), 1-26.

Girvan, M., Newman, M. (2002). Community structure in social and biological networks, *Proc. of the National Academy of Science* 99(12), 7821–7826.

Glaeser, E., ed., (2010). Agglomeration Economics, University of Chicago Press.

Glaeser, E., Sacerdote, B., Scheinkman, J. (1996). Crime and social interactions, *Quarterly Journal of Economics* 111, 507-548.

Goldfarb, Brent, Kirsch, David, Miller, David, (2007). Was there too little entry in the dot com era?, *Journal of Financial Economics* 86(1), 100-144.

Gomes-Casseres, B., Hagedoorn, J., Jaffe, A. (2006). Do alliances promote knowledge flows?, *Journal of Financial Economics* 80, 5-33.

Gompers, P., Lerner, J. (2001). The venture capital revolution, *Journal of Economic Perspectives* 15(2), 45-62.

Gompers, P., Lerner, J., (2004). The Venture Capital Cycle, MIT Press.

Gorman, M., Sahlman, W. (1989). What do venture capitalists do? *Journal of Business Venturing* 4, 231-248.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness, *American Journal of Sociology* 91(3), 481-510.

Grossman, S., Hart, O. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* 94(4), 691-719.

Guimera, R., Amaral, L.A.N. (2005). Functional cartography of complex metabolic networks, *Nature* 433, 895-900.

Guimera, R., Mossa, S., Turtschi, A., Amaral, L.A.N. (2005). The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles, *Proceedings of the National Academy of Science* 102, 7794-7799.

Guiso, L., Sapienza, P., Zingales, L. (2004). The role of social capital in financial development, *American Economic Review* 94, 526-556.

Harrison, D., Klein, K. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organization, *Academy of Management Review* 32(4), 1199-1228.

Hart, O., Moore, J. (1990). Property rights and the nature of the firm, *Journal of Political Economy* 98(6), 1119-1158.

Hellmann, T. J., Lindsey, L., Puri, M. (2008). Building relationships early: Banks in venture capital, *Review of Financial Studies* 21(2), 513-541.

Hellmann, T. J., Puri, M. (2000). The interaction between product market and financing strategy: The role of venture capital, *Review of Financial Studies* 13, 959-984.

Hellmann, T. J., Puri, M. (2002). Venture capital and the professionalization of start-up firms: Empirical evidence, *Journal of Finance* 57, 169-197.

Hoberg, G., Phillips, G. (2010). Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis, *Review of Financial Studies* 23(10), 3773–3811.

Hochberg, Y., Ljungqvist, A., Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance, *Journal of Finance* 62(1), 251-301.

Hochberg, Y., Lindsey, L., Westerfield, M. (2011). Inter-firm Economic Ties: Evidence from Venture Capital, *Northwestern University Working Paper.*

Hopp, C., Rieder, F. (2010). What drives venture capital syndication?, *Applied Economics*, 1-14.

Hwang, B., Kim, S. (2009). It pays to have friends, *Journal of Financial Economics* 93, 138-158.

Ishii, J.L., Xuan, Y. (2009). Acquirer-Target social ties and merger outcomes, *Working Paper*, SSRN: http://ssrn.com/abstract=1361106.

Kaplan, S. N., Schoar, A. (2005). Private equity performance: Returns, persistence and capital flows, *Journal of Finance* 60, 1791-1823.

Kaplan, S. N., Sensoy, B., Stromberg, P. (2002). How well do venture capital databases reflect actual investments?, *Working paper*, University of Chicago.

Kaplan, S. N., Stromberg, P. (2003). Financial contracting theory meets the real world: Evidence from venture capital contracts, *Review of Economic Studies* 70, 281-316.

Kaplan, S. N., Stromberg, P. (2004). Characteristics, contracts and actions: Evidence from venture capital analyses, *Journal of Finance* 59, 2177-2210.

Krishnan, C. N. V., Masulis, R. W. (2011). Venture capital reputation, in Douglas J. Cummings, ed., *Handbook on Entrepreneurial Finance, Venture Capital and Private Equity*, Oxford University Press.

Lerner, J. (1994). The syndication of venture capital investments. *Financial Management* 23, 1627.

Lerner, J. (1995). Venture capitalists and the oversight of private firms, *Journal of Finance* 50 (1), 302-318

Lerner, J., Merges, R.P. (1998). The control of technology alliances: An empirical analysis of the biotechnology industry, *Journal of Industrial Economics* 46, 125-156.

Leskovec, J., Kang, K.J., Mahoney, M.W. (2010). Empirical comparison of algorithms for network community detection, *ACM WWW International Conference on World Wide Web.*

Lindsey, L. A. (2008). Blurring boundaries: The role of venture capital in strategic alliances, *Journal of Finance* 63(3), 1137-1168.

Lewis, A.C., Jones, N.S., Porter, M.A., Deane, C.M. (2010). The function of communities in protein interaction networks at multiple scales, *BMC Syst Biology* 4(1), 100.

Lockett, A., Wright, M. (2001). The syndication of venture capital investments, *International Journal of Management Sciences* 29, 375-390.

Lusseau, D. (2003). The emergent properties of a dolphin social network, *Proceedings of the Royal Society of London B* 271 S6: 477–481.

McPherson, M., Smith-Lovin, L., Cook, J. (2001). Birds of a feather: Homophily in social networks, *Annual Review of Sociology* 27, 415-444.

Merton, R.E. (1987). A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.

Mokken, R.J. (1979). Cliques, clubs and clans, *Quality and Quantity* 13, 161–173.

Neher, D. V. (1999). Staged financing: An agency perspective, *Review of Economic Studies* 66, 255-274.

Newman, M. (2001). Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality, *Physical Review E* 64, 016132.

Newman, M. (2004). Fast algorithm for detecting community in networks, *Physical Review E* 69 (6), 066133.

Newman, M. (2006). Modularity and community structure in networks, *Proc. of the National Academy of Science* 103 (23), 8577-8582.

Pons, Pascal., Latapy, Matthieu, (2005). Computing communities in large networks using random walks, *Journal of Graph Algorithms and Applications* 10, 191–218.

Porter, M.E. (1998). Clusters and the new economy, *Harvard Business Review* 76(6), 77-90.

Porter, M.E. (2000). Location, competition and economic development: Local clusters in a global economy, *Economic Development Quarterly* 14(1), 15-34.

Porter, Mason., Mucha, Peter, Newman, Mark, Friend, A. J. (2007). Community structure in the United States House of Representatives, *Physica A: Statistical Mechanics and its Applications* 386(1), 413–438.

Porter, Mason., Onnela, Jukka-Pekka, Mucha, Peter, (2009). Communities in networks, *Notices of the AMS* 56(9), 1082–1166.

Rabin, M. (1993). Incorporating fairness into game theory and economics, *American Economic Review* 83, 1281-1302.

Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, N., Barabasi, A.L. (2002). Hierarchical organization of modularity in metabolic networks, *Science* 297 (5586), 1551.

Robinson, D. (2008). Strategic alliances and the boundaries of the firm, *Review of Financial Studies* 21(2), 649-681.

Robinson, D., Sensoy, B. (2011). Private equity in the 21st century: cash flows, performance, and contract terms from 1984-2010, *Working Paper*, Ohio State University.

Robinson, D., Stuart, T. (2007). Financial contracting in biotech strategic alliances, *Journal of Law and Economics* 50(3), 559-596.

Sahlman, W.A. (1990). The structure and governance of venture capital organizations, *Journal of Financial Economics* 27, 473-521.

Shue, Kelly (2011). Executive networks and firm policies: Evidence from the random assignment of MBA peers, *Harvard University Department of Economics Working Paper*.

Simon, Herbert (1962). The architecture of complexity, *Proceedings of the American Philosophical Society* 106 (6), 467–482.

Sorensen, Morten (2007). How smart is smart money? A Two-sided matching model of venture capital, *Journal of Finance* 62, 2725-2762.

Sorensen, Morten (2008). Learning by investing: evidence from venture capital, *Columbia University Working Paper*.

Sorenson, O., Stuart, T. (2001). Syndication networks and spatial distribution of venture capital investment, *American Journal of Sociology* 106, 1546-1588.

Sorenson, O., Stuart, T. (2008). Bringing the context back in: Settings and the search for venture partners in venture capital investment networks, *Administrative Science Quarterly* 53, 266–294.

Wu, K., Taki, Y., Sato, K., Kinomura, S., Goto, R., Okada, K., Kawashima, R., He, Y., Evans, A. C. and Fukuda, H. (2011). Age-related changes in topological organization of structural brain networks in healthy individuals, *Human Brain Mapping* 32, doi: 10.1002/hbm.21232.

Zachary, Wayne (1977). An information flow model for conflict and fission in small groups, *Journal of Anthropological Research* 33, 452–473.

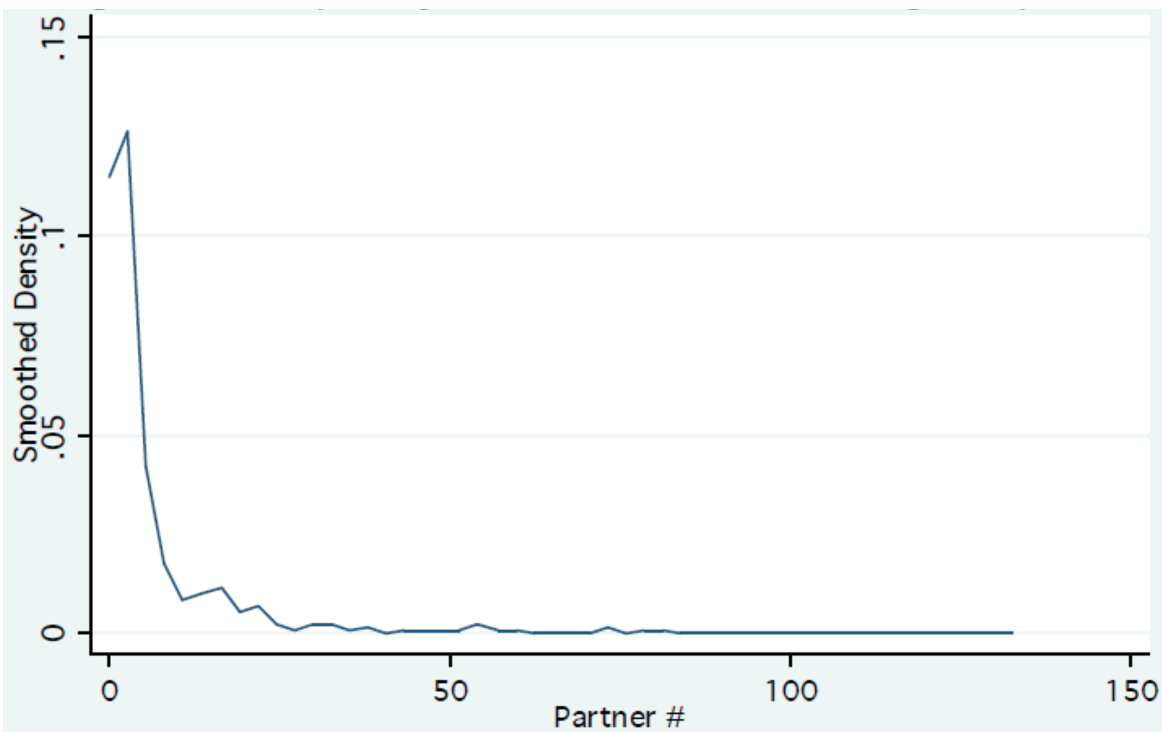

Figure 1: Frequency distribution of J.P. Morgan's partners.

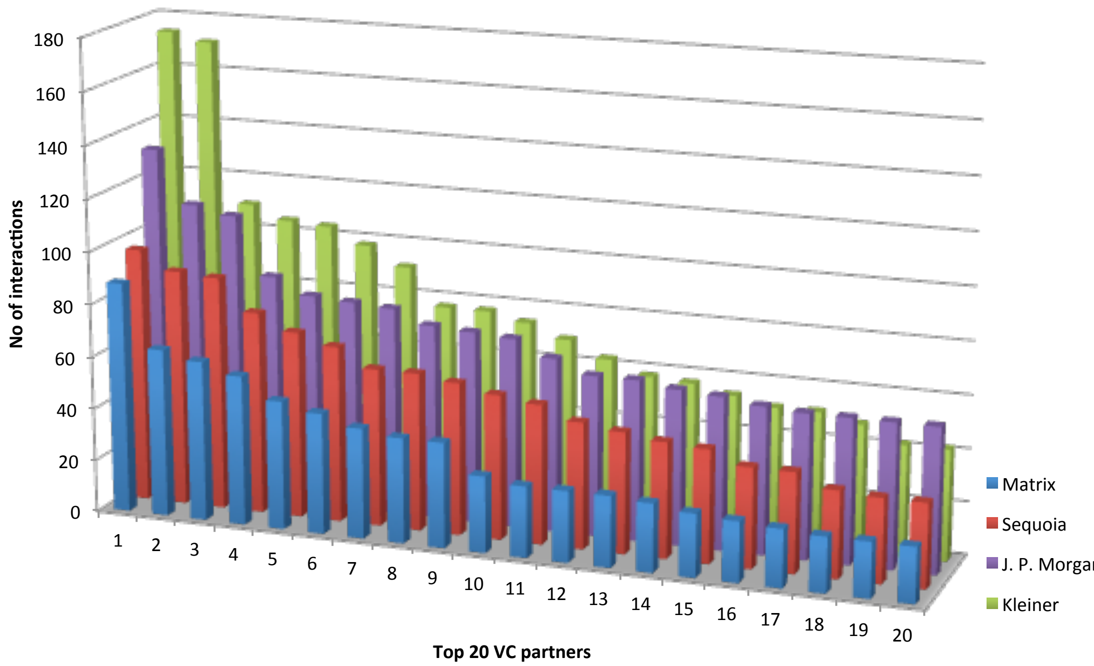

Figure 2: Distribution of the number of interactions of four top firms with their top 20 collaborators.

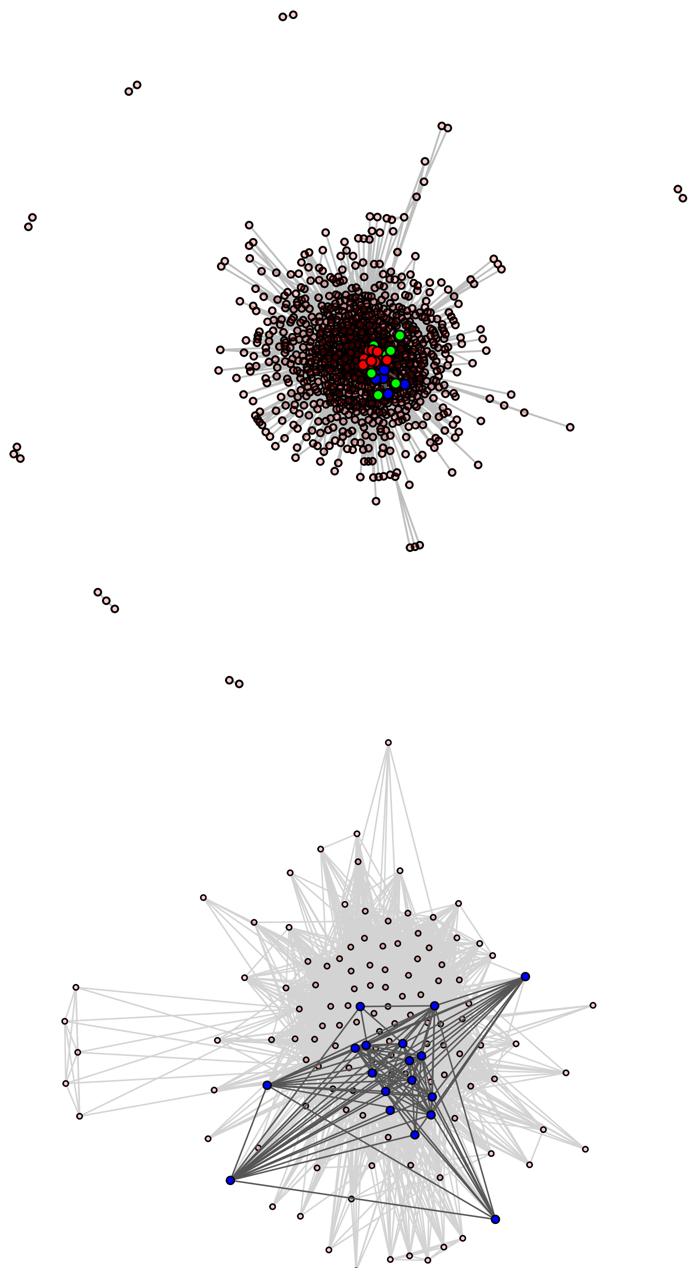

Figure 3: Network graph for connected VCs (1980–84). The upper plot shows the network of all VCs in communities (1180 in all), and blue, green, and red nodes in the center of the network are the VCs in the top three largest communities, respectively. The lower plot shows the network comprised only of the 134 VCs who are members of the 18 communities that have at least five VCs. The darker nodes in the lower plot show the VCs in the largest community.

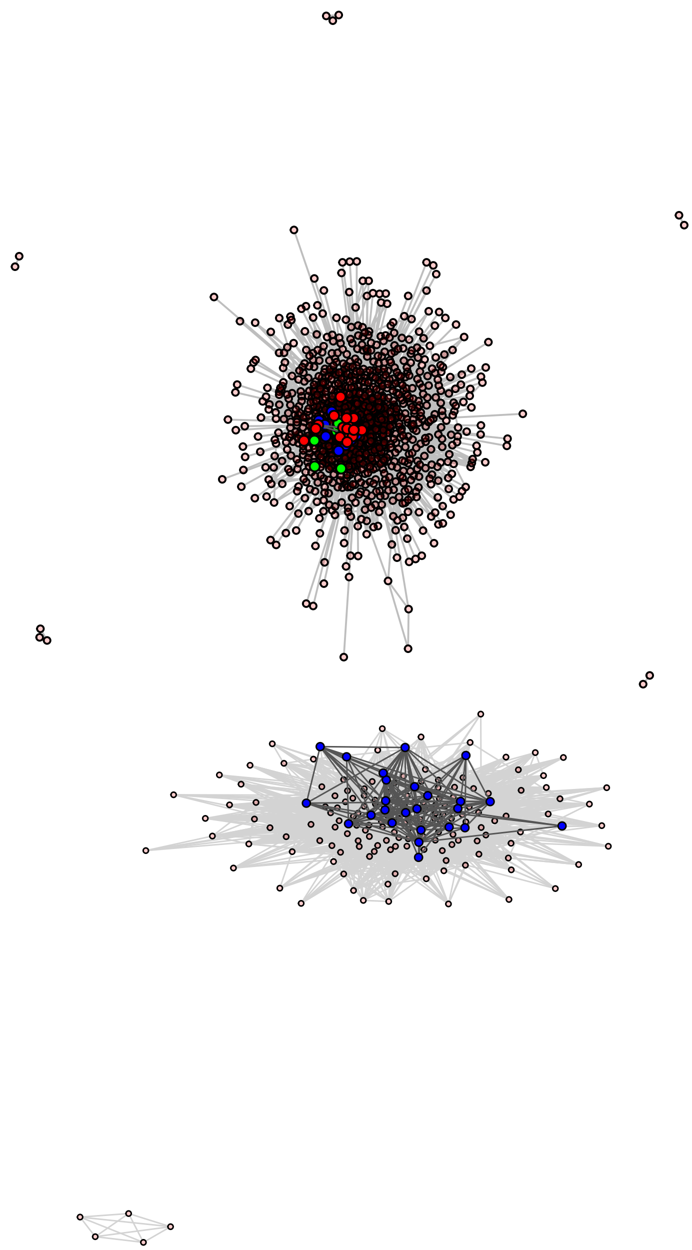

Figure 4: Network graph for connected VCs (1985–89). The upper plot shows the network of all VCs in communities (1295 in all), and blue, green, and red nodes in the center of the network are the VCs in the top three largest communities, respectively. The lower plot shows the network comprised only of the 180 VCs who are members of the 18 communities that have at least five VCs. The darker nodes in the lower plot show the VCs in the largest community. Note the single satellite community at the bottom of the lower plot. Such a community has low centrality.

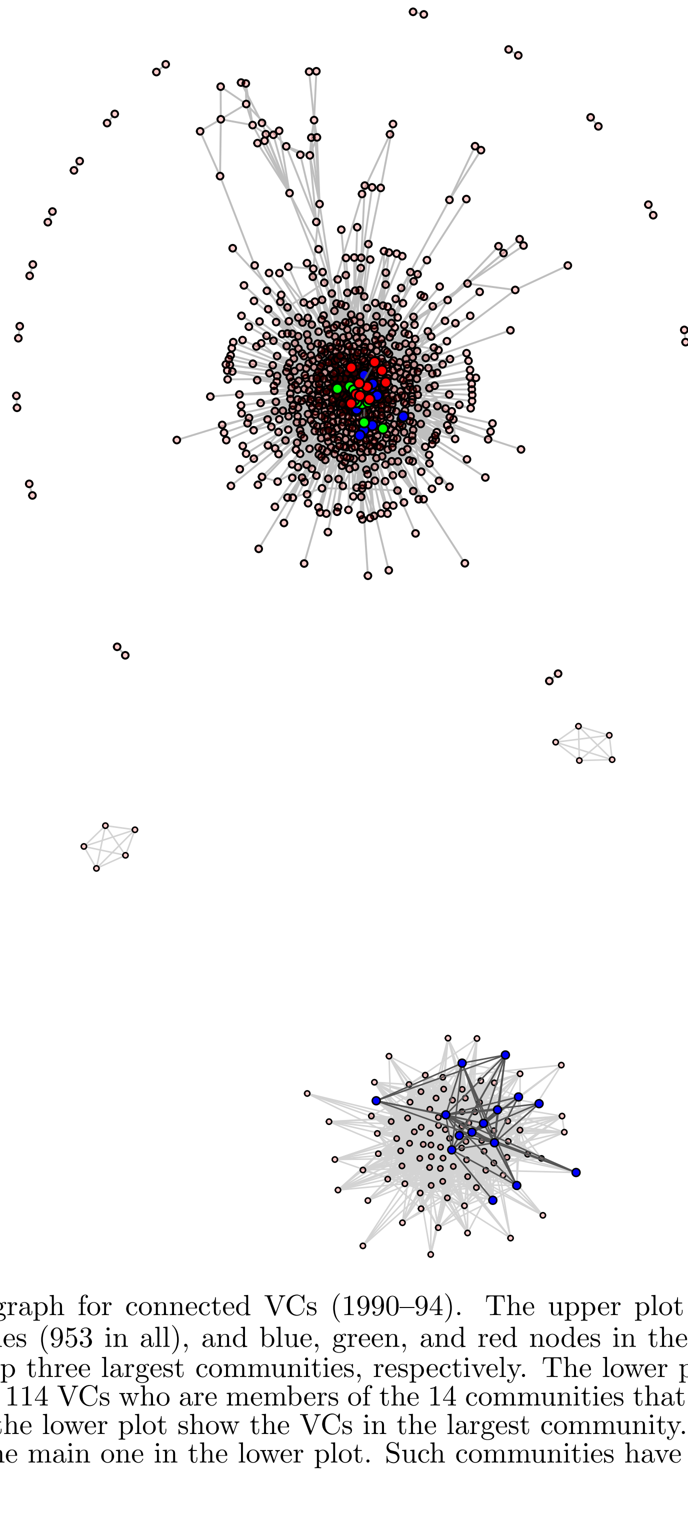

Figure 5: Network graph for connected VCs (1990–94). The upper plot shows the network of all VCs in communities (953 in all), and blue, green, and red nodes in the center of the network are the VCs in the top three largest communities, respectively. The lower plot shows the network comprised only of the 114 VCs who are members of the 14 communities that have at least five VCs. The darker nodes in the lower plot show the VCs in the largest community. Note the two satellite communities above the main one in the lower plot. Such communities have low centraity.

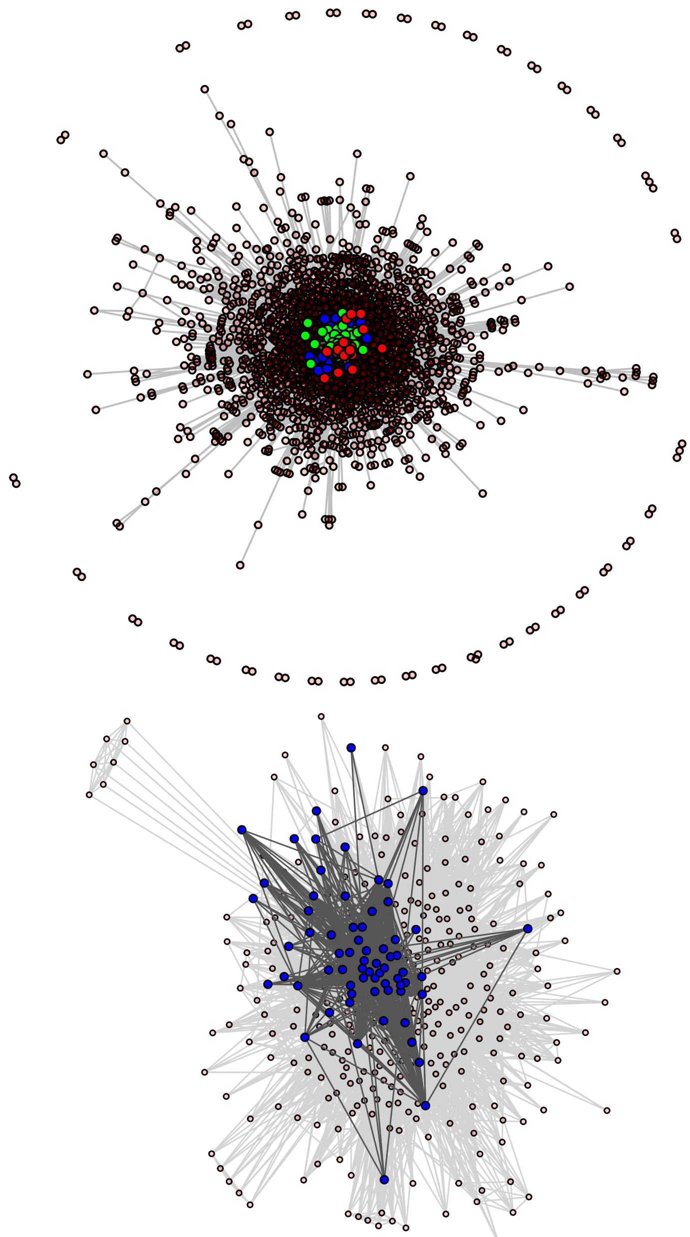

Figure 6: Network graph for connected VCs (1995–99). The upper plot shows the network of all VCs in communities (2772 in all), and blue, green, and red nodes in the center of the network are the VCs in the top three largest communities, respectively. The lower plot shows the network comprised only of the 379 VCs who are members of the 35 communities that have at least five VCs. The darker nodes in the lower plot show the VCs in the largest community.

Table 1: Venture Capitalists in our sample. This table provides descriptive statistics of the 1,962 unique U.S.-based VCs in our database over the entire 20-year period, from 1980 to 1999. Data are from Venture Economics and exclude non-US investments, angel investors, and VC firms focusing on buyouts. Size is the sum of the capital under management in all funds that were active during 1980-1999. Total investment is the sum of a VC's investments over this time period. % Deals Syndicated is the fraction of all rounds that a VC invested in that were syndicated. % early stage deals is the fraction of a VC's deals that are in the early stage. Age is defined as the difference in the year of the VC's last investment in the sample period and the VC firm's founding date.

Variables:	Mean	Median	# Observations
# Rounds	47.98	9.00	1,962
# Companies	21.64	7.00	1,962
Size ($ mm)	128.01	17.50	1,552
Total Investment ($ mm)	595.14	110.47	1,945
Investment per round ($ mm)	19.47	10.56	1,945
% Deals Syndicated	73.62	80.90	1,962
% Early Stage Deals	35.95	33.33	1,962
Age	9.59	6.00	1,950
# VC firms per MSA	14.24	3.00	127

Table 2: Sample Communities. This table details venture capitalists that belong to two sample communities, one each for 1985-1989 and 1990-1994. We chose the communities that had Stanford University's VC arm.

Sample community from the period 1985–89:
(1) Technology Venture Investors, (2) Associated Venture Investors (AKA: AVI Capital), (3) Bryan & Edwards, (4) Pacific Venture Partners, (5) Sequoia Capital, (6) Suez Ventures (FKA: Indosuez Ventures), (7) Partech International, (8) **Stanford University, (9) Asset Management Company Venture Capital, (10) Arthur Rock & Co., (11) Mohr Davidow Ventures, (12) OSCCO Ventures, (13) Draper Fisher Jurvetson (FKA: Draper Associates), (14) MedVenture Associates (AKA: MVA), (15) GT Technology Fund, (16) New Zealand Insurance, (17) Nippon Investment & Finance Co Ltd.**

Sample community from the period 1990–94:
(1) Mohr Davidow Ventures, (2) **Stanford University, (3) Kleiner Perkins Caufield & Byers, (4) Mayfield Fund, (5) Delphi Ventures, (6) Sequoia Capital, (7) Berkeley International Capital Corp., (8) W.S. Investments, (9) Avalon Ventures, (10) Technology Investment Fund, Inc., (11) Frazier Healthcare and Technology Ventures(FKA Frazier & Co), (12) Vertex Management Pte, Ltd. (AKA: Vertex Venture Holdings), (13) Integral Capital Partners, (14) Silicon Graphics, Inc., (15) Trinity Capital Partners.**

Table 3: Stability of community participation. The table provides data on the number of community VCs in each 5-year window. The variable "After one (three, five)) year" shows the proportion of community VCs in a window who continued to be in a community one, three and five years hence.

Window	# Community VCs	After 1 year	After 3 years	After 5 years
1980-1984	134	0.90	0.85	0.77
1981-1985	153	0.96	0.90	0.80
1982-1986	180	0.93	0.80	0.72
1983-1987	177	0.96	0.87	0.77
1984-1988	205	0.87	0.78	0.67
1985-1989	180	0.92	0.83	0.71
1986-1990	169	0.88	0.76	0.69
1987-1991	125	0.88	0.79	0.77
1988-1992	130	0.93	0.78	0.75
1989-1993	111	0.86	0.77	0.71
1990-1994	114	0.89	0.80	0.77
1991-1995	112	0.82	0.80	
1992-1996	146	0.93	0.89	
1993-1997	173	0.90		
1994-1998	246	0.94		
1995-1999	379			

Table 4: Stability of communities. For every pair of adjacent rolling windows, we generate the Jaccard similarity index for every community pair, one community from each of the two rolling windows. The index is defined as the ratio of the size of the intersection set to the size of the union set. We report a composite measure for each pair of adjacent rolling window, based on the mean Jaccard similarity index conditional on the index being positive. We compare these composite measures of communities with those of random communities generated through bootstrapping based on matching community sizes and number of communities in each 5-year rolling window. The last column shows the p-values testing the equality of the composite measure for the community and bootstrapped community. ***, **, and * denote 1%, 5% and 10% significance, respectively.

Window 1	Window 2	Community	Bootstrapped Community	p-value
1980-1984	1981-1985	0.188	0.064	0.01***
1981-1985	1982-1986	0.175	0.060	0.01***
1982-1986	1983-1987	0.182	0.056	0.01***
1983-1987	1984-1988	0.217	0.058	0.01***
1984-1988	1985-1989	0.141	0.055	0.01***
1985-1989	1986-1990	0.177	0.052	0.01***
1986-1990	1987-1991	0.155	0.052	0.01***
1987-1991	1988-1992	0.155	0.050	0.01***
1988-1992	1989-1993	0.252	0.055	0.01***
1989-1993	1990-1994	0.123	0.062	0.01***
1990-1994	1991-1995	0.246	0.065	0.01***
1991-1995	1992-1996	0.143	0.055	0.01***
1992-1996	1993-1997	0.128	0.042	0.01***
1993-1997	1994-1998	0.135	0.041	0.01***
1994-1998	1995-1999	0.109	0.042	0.01***

Table 5: Descriptive statistics for 33,924 rounds in 13,541 unique portfolio companies from 1985-1999. A round is a community round if at least one VC firm participating in it comes from a VC community. Communities are detected using a walk trap algorithm applied to syndicated deals over five year windows rolled forward one year at a time. The sample comprises VC deals obtained from Venture Economics excluding buyouts, angel investments and non-US deals. Industry classifications are as per Venture Economics. Exit data are obtained by matching with Thomson Financial IPO and M&A databases.

Variable	Total	Community Round	Not Community Round
	Panel A: Counts By Round		
# Deals	33,924	15,220	18,704
—Round 1	11,018	3581	7437
—Round 2	6881	3015	3866
—Round 3	4784	2410	2374
Syndicated	14,897	10,056	4841
Early stage	12,118	5472	6646
Geographical Cluster	16,270	9607	6663
Rounds with			
—Geographical Cluster VC	19,678	12,140	7538
—Corporate VC	3372	1923	1449
—FI VC	7586	4415	3171
	Panel B: Percentage By Venture Economics Industry		
—Biotech	6.8	7.3	6.3
—Commu&Media	12.1	13.3	11.1
—Hardware	7.3	9.0	6.0
—Software	19.8	22.7	17.5
—Cons Pdts	7.8	5.3	9.9
—Indus, Energy	5.9	3.4	8.0
—Internet	11.0	11.9	10.3
—Medical	13.7	15.0	12.7
—Others	8.5	4.4	11.9
—Semicond, Electricals	7.0	7.9	6.3
	Panel C: Round Statistics		
Proceeds ($ million)	38 (13)	48 (21)	29 (9)
# VCs	2.08 (1)	2.89 (2)	1.42 (1)
—in syndicated rounds	3.46 (3)	3.85 (3)	2.64 (2)
—in early stage rounds	1.93 (1)	2.53 (2)	1.43 (1)
—in round 1	1.54 (1)	2.03 (2)	1.31 (1)
—in round 2	2.00 (1)	2.70 (2)	1.45 (1)
—in round 3	2.38 (2)	3.23 (3)	1.52 (1)
	PANEL D: Exit		
Rounds with			
—IPO exits	3828	2071	1757
—M&A exits	8794	4363	4431
—Follow-on funding	23,972	11,903	12,069

Table 6: Time to exit and probability of exit. Specification (1) reports the estimates of a Cox proportional hazards model. The dependent variable is the number of days from financing to the earlier of exit or April 30, 2010. Specification (2) reports the estimates of a probit model in which the dependent variable is 1.0 if there is a successful exit (IPO or merger) within 10 years of the investment round and 0 otherwise. See Appendix B for a description of the independent variables. The sample comprises VC deals obtained from Venture Economics excluding buyouts, angel investments and non-US deals. All specifications include year and industry fixed effects, which are not reported for brevity. Both the specifications are overall significant at 1%. t-statistics based on robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	Cox	Probit
	(1)	(2)
Community	1.110	0.068
	(3.96)***	(3.44)***
Early Stage	0.907	-0.047
	(-4.63)***	(-2.98)***
Company Geographical Cluster	1.057	0.033
	(2.54)**	(2.05)**
Corporate VC	1.327	0.193
	(8.71)***	(7.65)***
FI VC	1.078	0.050
	(3.02)***	(2.70)***
Syndicated	1.319	0.219
	(11.99)***	(12.66)***
IPO Rate	1.078	0.071
	(1.34)	(1.65)*
Centrality	1.022	0.019
	(1.90)*	(2.19)**
VC Geographic Cluster	1.044	0.030
	(1.75)*	(1.68)*
Experience	0.985	-0.014
	(-1.14)	(-1.46)
Early Stage Focus	1.050	0.015
	(0.57)	(0.25)
Industry Focus	0.910	-0.049
	(-1.15)	(-0.80)
# Observations	30769	32362

Table 7: Success through next round financing or exit. Specifications (1)-(3) report the estimates of a Cox proportional hazards model. The dependent variable is the number of days from financing to the earliest of the next financing round, exit, or April 30, 2010. Specifications (4)-(6) report the estimates of a probit model in which the dependent variable is 1.0 if there is a successful exit (IPO or merger) or financing round within 10 years of the investment round and 0 otherwise. See Appendix B for a description of the independent variables. All specifications include year and industry fixed effects, which are not reported for brevity. The sample comprises VC deals obtained from Venture Economics excluding buyouts, angel investments and non-US deals. t-statistics based on robust standard errors are in parentheses. All specifications are overall significant at the 1% level. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	Cox			Probit		
	Round1	Round2	Round3	Round1	Round2	Round3
	(1)	(2)	(3)	(4)	(5)	(6)
Community	1.163	1.139	1.058	0.169	0.222	0.113
	(4.53)***	(3.40)***	(1.29)	(4.04)***	(4.02)***	(1.67)*
Early Stage	1.332	1.263	1.194	0.261	0.229	0.233
	(10.38)***	(7.98)***	(4.84)***	(8.30)***	(5.45)***	(4.08)***
Company Geographical Cluster	1.078	1.002	1.098	0.083	0.046	0.140
	(2.68)***	(0.08)	(2.50)**	(2.48)**	(1.02)	(2.59)***
Corporate VC	0.891	0.974	0.992	-0.126	-0.023	0.106
	(-2.32)**	(-0.58)	(-0.17)	(-2.06)**	(-0.31)	(1.24)
FI VC	0.889	0.927	0.908	-0.142	-0.126	-0.067
	(-3.64)***	(-2.14)**	(-2.47)**	(-3.78)***	(-2.49)**	(-1.06)
Syndicated	1.468	1.293	1.384	0.528	0.505	0.581
	(13.89)***	(7.63)***	(7.87)***	(14.70)***	(10.39)***	(9.48)***
IPO Rate	1.242	0.766	0.977	-0.098	-0.303	-0.051
	(0.86)	(-2.90)***	(-0.20)	(-1.34)	(-2.77)***	(-0.34)
Centrality	1.027	1.022	1.051	0.019	0.039	0.093
	(1.82)*	(1.41)	(2.72)***	(0.92)	(1.39)	(2.56)***
VC Geographical Cluster	1.062	0.974	0.891	0.090	-0.009	-0.098
	(2.06)**	(-0.74)	(-2.67)***	(2.66)***	(-0.18)	(-1.64)*
Experience	0.952	0.978	0.944	-0.064	-0.065	-0.045
	(-3.32)***	(-1.23)	(-2.38)**	(-3.88)***	(-2.73)***	(-1.46)
Early Stage Focus	1.386	1.867	2.005	0.183	0.639	0.552
	(3.70)***	(5.28)***	(4.29)***	(1.80)*	(4.06)***	(2.72)***
Industry Focus	1.148	0.978	1.003	0.074	-0.007	0.036
	(1.53)	(-0.19)	(0.02)	(0.71)	(-0.05)	(0.18)
# Observations	9471	6218	4320	9867	6522	4563

Table 8: Similarity of Same-Community VCs. The table compares key community characteristics with those of random communities generated through bootstrapping based on matching community sizes and number of communities in each 5-year rolling window. For each community (and bootstrapped community), we generate the mean and standard deviation of the characteristic. In Panel A and Panel B, the table presents the average value of these measures across communities (and random communities). *Age* uses the number of years between a VC's last investment in a 5-year window and the founding year of the VC firm. *Assets under management (AUM)*, in ($ million), uses the sum of a VC's active funds during each 5-year period. *Centrality* is based on each VC's eigenvector centrality determined for each 5-year rolling window. *Ownership HHI* is the Herfindahl index based on the different types of VC ownership in a community. *VC State HHI* is the Herfindahl index based on the number of VCs in each state. *Industry HHI* is the Herfindahl index based on the amount invested in each industry, while *Stage HHI* is the Herfindahl index based on the amount invested in each stage of investment. *Company State HHI* is the Herfindahl index based on the amount invested in each state by VCs. The last column shows the p-values testing the equality of the means of the community and bootstrapped community characteristics. ***, **, and * denote 1%, 5% and 10% significance, respectively.

	Community	Bootstrapped Community	p-value
	Panel A: Mean		
Age	9.51	8.60	0.01***
AUM	138.50	82.24	0.01***
Centrality	0.09	0.03	0.01***
Ownership HHI	0.44	0.43	0.25
VC State HHI, #	0.43	0.20	0.01***
Stage HHI	0.34	0.34	0.50
Industry HHI	0.25	0.21	0.01***
Company State HHI	0.37	0.27	0.01***
	Panel B: Standard Deviation		
Age	7.78	8.17	0.15
AUM	165.06	131.51	0.05**
Centrality	0.09	0.05	0.01***
Stage HHI	0.24	0.27	0.01***
Industry HHI	0.25	0.31	0.01***
Company State HHI	0.50	0.30	0.01***