Venture Capital Communities [1]

Amit Bubna
Indian School of Business
Gachibowli, Hyderabad, India 500 032

Sanjiv R. Das
Leavey School of Business
Santa Clara University, CA 95053

Nagpurnanand Prabhala
CAFRAL and University of Maryland, College Park

March 24, 2016

[1]We thank Alexandre Baptista, David Feldman, Jiekun Huang, Ozgur Ince, Vladimir Ivanov, Pete Kyle, Josh Lerner, Laura Lindsey, Robert Marquez, Vojislav Maksimovic, Manju Puri, Krishna Ramaswamy, Rajdeep Singh, Richard Smith, Anjan Thakor, Susan Woodward, Bernard Yeung, and participants at the CAF, FIRS, Midwest Finance Association, World Private Equity and TAPMI conferences, and seminar participants at Blackrock, Florida, George Washington University, Georgia State, Georgia Tech, Indian School of Business, Kellogg, Maryland, National University of Singapore, the Bay Area R User Group, Rutgers, Tulane, University of New South Wales, and Villanova for helpful comments. The authors may be reached at their respective email addresses: `amitbubna@gmail.com`, `srdas@scu.edu`, and `prabhala@umd.edu`.

Abstract

Venture Capital Communities

While venture capitalists (VCs) can select syndicate partners from a large pool, we show that many select partners from small subsets of VCs. Such preferences lead to the formation of several clusters, "VC communities," whose members are probabilistically more likely to partner with each other than others. Communities are new organizational forms that permit more resource pooling than in an arms-length market but do so without the rigidities of formal conglomerates. We use advances in computational techniques and over two decades of syndication data to identify VC communities. We characterize their size, number, spatial loci and the economic activities of community VC-funded portfolio firms. Communities comprise VCs with similar age, influence, and functional style. Community VC-funded firms, especially earlier-stage firms with limited innovation histories, display more innovation scale and novelty and mature faster, particularly when functionally similar VCs pool diverse experiences.

JEL classification: G20, G24
Key words: Venture capital, syndication, community detection, social networks, boundaries of the firm

1 Introduction

Venture capitalists raise capital from wealthy individuals and institutional investors and invest in young firms that promise high upside. According to the National Venture Capital Association data there are 76,558 VC deals for over $618 billion between 1995 and 2014 in the U.S. VC successes include many high tech firms such as Apple, Cisco, and Microsoft. See Da Rin, Hellmann, and Puri (2012) for a recent survey of VC research.

VC investing is risky. Firms financed by VCs tend to be young, and often have unproven business models. VC investing is also resource intensive, demanding considerable effort in terms of ex-ante screening and follow on advising for portfolio firms.[1] VCs use multiple strategies to manage both the risks and the resource demands of investing. One approach is to use security design features such as priority, staged financing, or contracting over control rights.[2] Another approach is to syndicate deals, or co-invest in portfolio firms alongside other VC firms. Syndicated financings account for over two-thirds of all VC deals. Only 5% of VCs never syndicate and these are small, peripheral firms.

We study the VC syndication process, focusing on new types of economic agglomerates, "venture capital communities," that form as a result of the syndication process. Our study has three aims. First, we formalize VC communities as a mathematical construct. We discuss the computational issues in identifying them and the plausible economic forces that lead to their formation as an equilibrium outcome. Two, we empirically characterize the number and nature of communities in U.S. VC data. Finally, we characterize the economic activities of communities such as the scale and novelty of innovation and firms maturation.

To motivate the analysis, we observe that while VCs have many potential partners for syndication, they exhibit strong preferences for some partners over others. For instance, Figure 1 displays the frequency distribution of the syndicate partners of four VCs. The long and thin right tail indicates that the VC has many partners. However, the thick left tail shows the VC's preference for some partners over others. J P Morgan's preferred partners include Kleiner Perkins, Oak Investment Partners, and the Mayfield Fund. Similar features

[1] See Gorman and Sahlman (1989) or Hellmann and Puri (2002).

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

are revealed in anecdotal evidence.[3]

The bottom line is that while VCs can potentially choose from a large pool of syndicate partners, they tend to exhibit strong preferences for some partners over the others. The preferential attachment of VCs to some syndicate partners over others leads to the formation of agglomerates, or clusters of VCs, that we term as VC communities. These community clusters have the property that its members are more likely to partner with each other than randomly with outside-cluster VCs. This preference is, however, probabilistic rather than deterministic so cluster members can partner with outsiders although less frequently than with each other.

As an organizational form, a community lies in between conglomerates and a pure arms-length marketplace where firms transact with each other. As community members are likely future partners, VCs have closer relationships than in a pure arms-length marketplace. However, the ties are less formal than those between entities residing inside a formal conglomerate. For instance, community members can engage more freely with outside-community entities. This description is reminiscent of the work in inter-firm alliances (Robinson and Stuart, 2007; Baker, Gibbons, and Murphy, 2008; Robinson, 2008). The analogy is apt. The construct developed here, a VC community, formally defines and operationalizes the entity that emerges from a web of multiple, simultaneous, and partially overlapping partnerships between entities. Developing the notion formally and illustrating the resulting organizational form in the VC context is indeed our primary goal.[4]

We begin by identifying communities from observed VC syndication data. The raw data are the history of partnerships between VCs in past syndications. From these data, we identify communities, which are clusters of VCs with a high propensity to do business with each other relative to what would be expected by chance. Community detection is thus a clustering technique. It is, however, different from clustering methods used in finance, e.g., Brown and Goetzmann (1997). We briefly consider the differences next.

---

[3]Fred Wilson of Union Square Ventures, a prominent investor in major social network sites such as Tumblr, Twitter and Zynga, says "... there are probably five or ten VCs who I have worked with frequently in my career and I know very well and love to work with. It's not hard to figure out who they are.." http://www.avc.com/a_vc/2009/03/coinvestors.html.

[4]In our framework, communities are disjoint sets. The analysis of non-disjoint communities, i.e., overlapping communities, with each VC in multiple communities, introduces computational and definitional complexity beyond the current literature on communities. Work on non-disjoint communities is still developing. One way to look at our results is that we capture strong associations. Weaker links are assigned elsewhere outside communities. A more continuous measure of groups in which one belongs to different groups with different levels of strengths is a complex mathematical construct that we leave for future work.

Standard clustering problems optimize the variation of characteristics within relative to across clusters. Here, the clustering problem optimizes *syndication probability* within relative to across communities. Second, we allow considerable flexibility in cluster formation. We do not pre-specify the number of clusters or cluster size. We also let different clusters be of different sizes. We allow the possibility that not all VCs may belong to clusters. The features described before represent the syndication process more realistically but result in a computationally complex problem, whose exact solutions are not known. Approximate solutions are available using algorithms such as the walk-trap method (Fortunato, 2009; Pons and Latapy (2005)).

Our data comprise U.S. VC syndications between 1980 and 2010. We follow the VC literature and work with 5-year rolling windows of syndication links between VCs. Each firm enters only once during each period. The set of syndicates and thus communities are updated in the next window so we allow for entry and exit of VCs across successive time periods. We detect communities in every time period in our sample. On average, 11% of VC firms in each period belong to communities. The median community has 8 members. We also find that a VC firm's community status is stable. A VC belonging to a community has an 88% probability of belonging to a community 5 years later.

We next consider the characteristics of VCs comprising a community. One view is that a VC's preferred partners should be similar. Similarity could reflect a version of the "birds of a feather flock together" viewpoint of McPherson, Smith-Lovin and Cook (2001), theories of contracting with private and manipulable signals (Cestone, Lerner and White (2014)) that suggest partner complementarities, or greater trust in partners who are functionally similar. On the other hand, heterogeneous partners give VCs access to broader skill sets, bring more diverse resources to help portfolio firms, or extend a VC's reach into new domains (Hochberg, Lindsey, and Westerfield (2015)).

Our tests require us to specify attributes along which there may be similarity- or dissimilarity-seeking behavior. We include proxies for VC reach and functional style. Statistical tests pose a separate challenge as we must compare clusters of varying size and number in each period and also account for similarities that may arise definitionally from the syndication process. We use appropriately conditioned simulations, as we explain later, following a familiar path suggested in the finance literature (Brown and Warner, 1985; Barber and Lyon, 1997). We find similarity-seeking behavior, specifically in partner age, influence, and functional style both in how VCs spread capital and where they do so.

We next consider where communities locate in the style space. One view of communities is that they are soft conglomerates that house a broad vector of skills to cater to a broad range of businesses. If so, communities should not spread locations in the style space. Alternatively, different communities could spread locations in the style space, perhaps so communities pool different skills and compete through differentiation. We test these two views by examining the distances between communities along different attribute dimensions. We do not find significant differences in the spatial loci of communities.

We next characterize the economic outcomes associated with community VC financed firms. The outcomes could be a result of better ex-ante selection by community VCs or their ex-post causal effects. We do not aim to disentangle these components or estimate causal effects. Rather, our goal is to characterize communities as equilibrium outcomes, novel organizational forms that are robust features of three decades of syndication data and understand their economic drivers. Thus, for instance, it is plausible that community VC funded firms have greater innovation because community membership makes VCs "smart" so they join rounds in which there is more innovation potential, potentially but not necessarily due to their own causal impact. Our analysis does not preclude the latter but does not rely on it either. Our view is that if causal inference is the goal, the structural approach of Sorensen (2007) is appropriate but estimable structural models for community formation are not known.[5] However, with regards to drivers of community formation, progress is possible and we show how. Instead of specifying a structure, which may or may not fit (unknown) moments derived from community structures, we invert observed market outcomes, communities, to understand the economics driving their formation.[6]

Our first outcome variable is patenting, which is a proxy for innovation. Patentable innovation is a key focus of young enterprises financed by VCs as reflected in the volumes and type of patents. While we recognize the limitations of patenting as a measure of innovation,

[5] Using standard IV methods (Roberts and Whited, 2011) has two issues. One is the difficulty in finding natural experiments or instruments for community formation. This is a non-trivial hurdle as even instrumenting for syndication, the input into community formation is difficult (although see Gompers et al., 2012 for an effort). Relatedly, there are long gestation period prior to observing VC outcomes. This results in a complex dynamic structure unsuited to a one-off IV approach.

[6] Equivalently, our paper treats communities as given. We characterize their economics using the rich variation in entities, economic outcomes, and heterogeneity to better understand the economic forces that lead to their formation.

it is the primary focus of received work on innovation. Our primary finding is that the participation of community VCs in a financing round is associated with an increase in both the number of patents and the citations per patent. In baseline specifications, community VCs are associated with a 2% increase in total patents and a 5% increase in average citations in these firms. In more stringent specifications, the estimates are 6% and 17%, respectively.

The patenting results are more pronounced in portfolio companies in early stages with no prior innovation, who likely face more ambiguities in their growth path and likely derive greater benefits from information flows among community VCs. The results are in line with the literature which shows that innovation tends to happen outside large conglomerates, through strategic alliances and joint ventures (Seru (2014)), which are also alliances with soft boundaries and implicit contracts, much like VC communities. Better connected VCs, those with more wide experience, similarity of VC age and connectedness within a community, matter more.

We also examine the standard outcome variable in the VC literature, i.e., exit. We find that portfolio companies with community VCs are 9% more likely to successfully exit (via IPO or acquisition) and exit 13% sooner. A key issue in large-scale VC studies is the considerable noise in the exits via M&As, which reflects a mix between failures and success stories such as Skype's acquisition by Microsoft. We thus consider multiple definitions of exit. We include a model that has only IPOs. We also estimate a competing risks specification in which M&As and IPOs are alternative forms of exit in a competing hazards framework. We find a positive relation between successful exit and getting funding from a community VC rather than non-community VC.

We proceed as follows. Section 2 formally defines communities and develops the related optimization problem and solution methods. Section 3 discusses data. In Section 4, we describe the implementation of community detection methods on the VC data and discuss the baseline results. Section 5 examines composition of VC communities. Sections 6 and 7 study innovation and exit. Section 8 concludes and suggests directions for future research. The Appendixes give sample code and defines variables used in our analysis.

2 Community Formation

Given a set of VCs, we define a community as a *group* or cluster of VCs with the property that the members of each cluster have strong propensities to syndicate with each than with random VCs. We explain the relevant mathematical construct, discuss its economics, and position our work relative to literature. A key point is that community is a *group* construct. It identifies entire groups of VCs who tend to work together. In contrast, conventional network metrics focus on an *individual* entity's connections. We return to this point later.

2.1 Mathematical Definition

We suppress time period subscripts t for compactness. Suppose that there are K communities that each comprise VCs who prefer to co-syndicate. Let there be $n_k \geq 2$ members per cluster, $k = 1, 2, 3, ...K$. We do *not* require that all n_k's be equal nor do we require that $N_k = \sum_{k=1}^{K} n_k = N$, the number of VCs. Thus, there could be $N - N_k$ singleton VCs who do not belong to any community. Mathematically, a community identifies K, n_k, and a set of indicators that assign each VC to a specific community or no community.

We specify the community detection problem formally. Formally, let c_k denote community k and $\delta_{ij}(c_k)$ be an indicator variable equal to 1 when both VC i and VC j belong to community k. We define the *propensity to syndicate* within cluster k as the actual number of syndications within a cluster minus what is expected by random assignment. The sum of in-cluster syndication propensities is modularity Q. The mathematical problem of community detection is to choose the optimal number of clusters K, the size of each cluster n_k, and the cluster membership, i.e., the set of indicator variables $\delta_{ij}(c_k)$, to maximize *modularity* Q, or

$$\begin{aligned} &\underset{\delta_{ij},c_k}{\text{maximize}} \quad Q \\ &\text{where} \qquad Q = \frac{1}{2m} \sum_k \sum_{i,j} \left[a_{ij} - \frac{d_i \times d_j}{2m} \right] \cdot \delta_{ij}(c_k) \end{aligned} \tag{1}$$

where $d_i = \sum_j a_{ij}$ is the number of syndications done by VC firm i (j) and $m = \frac{1}{2} \sum_{ij} a_{ij}$, with the factor of 2 to reflect the equality of ties between i and j and ties between j and i.

The first term in the square bracket in Eq. (1) represents the actual number of deals co-syndicated by VCs i and j. The second term in [...] represents deals expected to be co-syndicated between i and j purely by chance. Intuitively, VC i with many connections will have greater odds of syndication with *any* VC j. Thus, the numerator in the second term $d_i \times d_j$ is increasing in the number of connections of VC i. The difference between the two terms represents cluster k's propensity to in-syndicate.

Modularity Q sums the in-cluster syndication propensities across all clusters c_k. Q lies in [-1, +1]. $Q > 0$ means that intra-community ties exceed ties predicted by chance. There are no known exact solutions for Q-optimization beyond tiny systems. The problem is computationally intensive (NP-hard) given the large number of feasible partitions, number of clusters, and cluster sizes. Fast solution methods include agglomerative techniques that start by assuming all nodes are separate communities and build up clusters iteratively. The walk-trap method that we employ initiates simultaneous random walks at several nodes and takes random steps, with communities defined as clusters from which the random walks fail to exit within a fixed number of steps. (Fortunato (2009, Pons and Latapy (2005)).

2.2 Economic Motivation for Community Formation

Economic theory motivates why VCs cluster into agglomerates of preferred partners. For instance, learning-by-doing models (Goldfarb, Kirsch, and Miller (2007), Sorensen (2008)) argue that VC investing is skill-intensive and skills must be acquired through learning-by-doing. Drawing from a small set of preferred partners aids learning as the familiarity results in better understanding of partner norms and processes (Gertler (1995); Porter (2000)). Likewise, in incomplete contracting problems studied in Grossman and Hart (1986) or Hart and Moore (1990), free riding or hold-up problems between partners lowers investment. These problems are alleviated when partners are familiar, which may also help by enhancing trust and reciprocity (Guiso, Sapienza and Zingales (2004), Bottazzi, Da Rin and Hellmann (2011)). Work in sociology (Granovetter (1985)) also suggests that agents place more weight on information flows from familiar sources.

A related literature discusses the tradeoffs between familiar and unfamiliar partners. Relationships with familiar partners generate social capital (Gulati (1995)) but repeating strong relationships with no change also precludes access to new information (Granovetter (1973)). Uzzi (1997) argues that something in between where agents have both strong and weak ties is likely optimal. This description is, of course, exactly the intuition for VC communities.

Community formation may also be viewed as being a spatial analog of urban agglomeration. As Glaeser (2010) writes, the benefits of agglomeration occur when individuals live close to each other in cities, or in the case of businesses, when they co-locate in clusters. In the VC context, dealing with each other repeatedly in syndicates has similar effects. In the urban agglomeration setting, transport costs, which Glaeser interprets broadly to include transaction costs of living, exchanging ideas, and so on, are lowered in agglomerates, Likewise, a VC's costs of contracting with syndicate partners such as costs of screening, ex-post monitoring, or efforts in advising ventures, can be lower when they belong to communities.

2.3 Communities and Social Networks

Our study is related to work on social networks in finance. The primitive input into this work is a tie between two individuals or entities. One part of the literature simply focuses on the tie itself.[7] This literature takes pairwise ties as attribute variables that explain economic outcomes such as information flows, fraud, or investing success. A second portion studies the sum or aggregates of all ties of individuals. [8] Hochberg, Ljungqvist and Lu (2007) study VC centrality while Engelberg, Gao and Parsons (2013) study CEO centrality and predict outcomes such as exit or compensation.

Communities are different organizational constructs that do not imply nor are implied by pairwise ties or network measures such as centrality. Community takes as primitive input the pairwise ties between individuals, but shares little else either operationally or economically with network constructs.[9] Operationally, centrality is a raw or weighted sum of an agent's connections while community membership is a solution to the more complex problem of optimizing modularity Q in Eq. (1). Economically, the two measures capture very different economic intuitions. Centrality measures an individual's influence while community is a group construct. Community membership identifies entire *groups of agents* who tend to do business together. Centrality counts ties of an agent, not whether the agent belongs to a

[7] Cohen, Frazzini and Malloy (2010) and Cohen, Frazzini and Malloy (2012) examine ties between analysts and boards of directors of firms. Hwang and Kim (2009) and Chidambaran, Kedia, and Prabhala (2010) analyze links between CEOs and directors. Bhagwat (2013) and Gompers et al (2012) examine connections between executives employed at VC firms based on VC executive biographies. See also Bengtsson and Hsu (2010) and Hegde and Tumlinson (2014).

[8] For a textbook treatment, see `http://faculty.ucr.edu/~hanneman/nettext/`.

[9] The physical sciences literature on communities barely refers to centrality. For further discussion of these distinctions, see Sections 7.1, 7.2, and 11.6 of Newman (2010).

cluster, how many such clusters form, or characteristics of clusters, but these questions are at the heart of our study of communities.[10]

A related question is how community detection is related to analyzing ties one pair at a time. We argue that community detection is computationally complex but nevertheless is a simpler approach towards understanding pairwise tie formation. Specifically, underlying a tie formation model is a structural model that describes what set of VCs join together for a given financing opportunity. This is likely a complex model that maximizes joint production, which in turn depends on the business opportunity on hand, a VC's own experience and skill set, and complementarities. Even if these parameters can be specified, inference is complex as the error term has complicated time series and spatial structures and the number of partners per financing opportunity vary. In fact, far simpler network formation models are difficult to estimate (e.g, Currarini, Jackson, and Pin (2012)).

Reduced form i.i.d logit models of pairwise tie formation such as those of Du (2011) and Hochberg, Lindsey, and Westerfield (2015) are one approach towards understanding the complex structure. However, community detection offers an alternate approach. It starts with the *end product*, or the actual observed syndication partnerships established by VCs. These are the revealed preferences of VCs. Community detection can be thought of as inverting these observed choices to infer the drivers of VC behavioral preferences.[11]

3 Data

We match three different datasets for our analysis. We obtain data on investments made between 1980 to 2010 and reported in Thomson's Venture Economics (or VentureXpert) database. We start in 1980 as it dates the institutionalization of the VC industry (Gompers and Lerner (2001)). We end in 2010 to allow sufficient time to observe investment outcomes. We refine the initial sample from VentureXpert by dropping non-U.S. funds, buyout funds,

[10]Related applications include, for instance, work on identifying politicians who vote together (Porter, Mucha, Newman, and Friend (2007)). product word groups (Hoberg and Phillips (2010)), and collaboration networks (Newman (2001)). Consider also Figure 2, reproduced from Burdick et al (2011). The three banks with high centrality are Citigroup, J. P. Morgan, and Bank of America. None belongs to communities, which are in the left and right nodes of Figure 2.

[11]As an example, a theory of "preferential attachment" (Barabasi and Albert (1999)) suggests that well-connected nodes attract connections to even more nodes. Adding familiarity, in which repeat interactions lead to learning and greater likelihood of working together, and the characteristics of VCs and portfolio funds, potentially offers a structural model for community formation.

deals in which we cannot identify investors or if there are only individual investors such as angels or management. We do not exclude deals that involve institutions such as subsidiaries of financial institutions and technology transfer offices of universities. While these investors have different incentive schemes, there are two reasons to keep them in sample. First, the deals they finance involve traditional institutional VCs. Second, each syndicated deal, whether between institutional VCs or by institutional VCs with others, offers VCs an opportunity to interact and learn. These elements are the essence of community formation.

Table 1 gives descriptive statistics for our sample at the level of the VC firm. Our sample includes 3,275 unique VC firms. On average, a VC firm invests in 24 portfolio firms and 56 rounds. Each round involves investment of $4.25 million. Close to three-quarters of the deals made by a VC are syndicated and about one-third of the rounds are classified as early stage investments. The total funds raised by a VC amount to about $472 million (median = $60 million). The average age of each VC at the time of its last investment in our sample is a little less than 11 years. VCs' headquarters are located in 153 Metropolitan Statistical Areas (MSAs) in our sample, with an average of about 20 VCs per MSA (median of 4 VCs). The two big VC clusters in California (CA) and Massachusetts (MA) account for about 35% of the VC firms' headquarters.[12]

We study two economic outcome variables. One is patenting, which is a measure of innovation. We work with the NBER patent database created by Hall, Jaffe and Trajtenberg (2001). The patent database covers over 3 million utility patents granted by the U.S. Patent and Trademark Office (USPTO) from 1976 to 2006. Following the literature, we focus on utility patents. The database only looks at patents that were ultimately granted. If a patent was granted, the database provides name and location (country and state) of the assignee. It is possible for a patent to be assigned to more than one company. However, in many instances, the assignee information is missing. There are about 218,000 unique assignees of which about 113,000 are located in the United States, accounting for about 1.4 million patents.

We match patent assignee names with those of VC portfolio companies. We follow the process of standardizing names in both datasets, and then matching them both by name as well as location (country and state). In case there is no exact name match, we follow

[12]Some VCs may have satellite offices that we do not include in the current analysis.

a fuzzy matching algorithm, and then identify the matching pair, if any, by hand based again both on name and location. We recognize that companies may experience changes in names over time, for reasons such as M&A transactions. While VentureXpert provides both pre-merger and post-merger names of some companies, our program additionally crawls the BusinessWBusinessweekeek website for changes in names of the other companies in the VentureXpert database. We then match both these names against those in the NBER database. We end up with about 7,000 unique name matches, i.e., portfolio companies in the VentureXpert database with patent information, accounting for about 100,000 unique patents.

Finally, we study exits. VC firms can exit through mergers and acquisitions (M&A's) or through IPOs. We obtain data on IPO firms 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 VentureXpert database. 2,545 ventures in our sample exit via IPOs. We obtain M&A data from Thomson Financial's SDC M&A database. There are 5,106 exits via mergers in our sample. We examine exits until 2012.

4 Baseline Results

4.1 Community Detection and Simulated Communities

We use overlapping windows of 5-year length to detect communities. Thus, the first window uses VC investments from 1980 to 1984, the second one uses data from 1981-1985 and so on. These windows allow sufficient time to identify partner preferences but avoid excessively long periods that may contain stale information. We require a minimum community size of five members. In addition, we require that the end-to-end diameter not exceed one-fourth that of the entire network.[13] This constraint is not binding in our dataset. To mitigate repeated syndications between the same partners, we count multiple links among VCs in any portfolio company just once in each 5-year window.

We find several communities in the time periods in our dataset. VCs cluster into between 7 and 25 communities in each five year period. Across the whole sample period, the median community has 8 members. There is considerable variation in the median community size in

[13] Diameter is the longest of the shortest paths from one node to another within a community.

each 5-year window. Overall, about 11% of the active VCs belong to communities, though there is significant variation in community membership among active VCs over time.

We ask whether a VC who belongs to a community in one period belongs to *some* community in another period. Table 2 reports the results. On average, over 99% of community VCs continue to belong to a community in the next period. It could be argued that community status stability from one period to the next is driven by the overlapping nature of 5-year windows used for generating communities. Instead, we also consider a VC's community status for longer horizons, such as 3 years and 5 years later. For instance, for a VC's status 5 years later, if a VC was in a community in the 1980-1984 window, we examine whether she is again in a community in the 1985-1989 window. We find that a significant proportion, 88%, of community VCs remain community VCs five years later. The community status of a VC is stable but not invariant.[14]

4.2 Other Descriptive Statistics

Table 3 reports descriptive statistics for classified by whether rounds are financed by a community VC. 32,547 out of 73,414 rounds (about 44%) are community rounds and these account for 65% of proceeds. 22,812 out of 35,088 syndicated rounds, or 65%, are community rounds. Early stage rounds account for about a third of the sample and 41% of these are community rounds. 35,414 deals or close to one-half of the investment rounds are in the California and Massachusetts (CA/MA) geographical clusters. This pattern reflects a concentration of VC investments in these states as well as their representation in VC databases (Kaplan, Sensoy and Stromberg (2002)).

VentureXpert classifies VC firms into 10 industry categories. Panel B of Table 3 shows that the software industry accounts for the largest share of financing rounds in our sample, followed by internet firms, medical or health firms, and communications and media firms. Interestingly, community VC is least likely for consumer product or industrial businesses, which are amongst the less risky and less complex industrial sectors. The finding indicates that VC firms draw on preferred partners more when facing greater uncertainty, consistent

[14] An additional stability measure is whether VCs tend to belong to the *same community* in successive periods. However, given the number of ways in which a community could break up and re-form, individual measures of overlap are infeasible. Aggregating overlaps across all communities results in a weaker measure. We cannot reject the hypothesis that the communities are similar.

with Cestone, Lerner and White (2014).

Panel C in Table 3 describes key characteristics across rounds. There is significantly greater investment in rounds with a community VC ($15 million) than in rounds with no community VCs ($4 million). Besides higher investment per round, community rounds tend to have more VC firms than rounds with no community VC, both in syndicated and non-syndicated rounds. For instance, community VC syndicated rounds have 3.69 VCs on average compared to 2.69 VCs in rounds with no community VC. This pattern holds for early stage rounds and initial financing rounds. Panel D gives exit information, which we analyze and discuss later.

5 VC Community Composition

In this section, we examine the characteristics of community agglomerates formed by VCs and also the spatial loci of the different VC communities. We briefly motivate both the variables and the type of analysis before discussing the results.

5.1 Hypotheses and Structure of Analysis

In principle, VCs could seek partners who are similar or dissimilar to themselves. Similarity-seeking behavior among VCs could reflect a behavioral preference for interacting with people of similar backgrounds, as discussed in McPherson, Smith-Lovin and Cook (2001). Alternatively, Chung, Singh, and Lee (2000) argue that status-based homophily can prevail because it has signaling value to outsiders. Cestone, Lerner and White (2014) develop theoretical models in which partner similarity acts in complementary ways to generate the best financing decisions. In the learning and vetting hypothesis for syndicate formation (Sorensen (2008)), the evaluation and screening abilities of partners drive syndicate partner preferences. If knowledge is style-specific, there should be within-community similarity along the dimension of functional expertise of VCs.

The case for dissimilar partners is based on the benefits of diversity. For instance, funds skilled in raising capital may partner with niche focused funds with skills in specific sectors. As Hochberg, Lindsey, and Westerfield (2015) point out, complementarity-seeking behavior could also result in preferences for the dissimilar. For instance, if there are two attributes X

and Y characterizing VCs, complementarity-seeking behavior suggests that high X, low Y VCs should prefer as a partner low X, high Y VCs. The net effect is that preferred partners have high variance in both X and Y. As a concrete example, the large VC firm Kleiner, Perkins, Caufield, and Byers largely prefers to invest in the clean tech area with the smaller VC firm Foundation Capital, which has greater domain expertise in the area. Foundation Capital benefits from the fund-raising capability and the reputation of the larger Kleiner, while Kleiner accesses niche expertise that a small firm such as Foundation Capital brings.

Following the above discussion, we specify two categories of attributes. One category of attributes proxies for a VC's reach and influence. Following Hsu (2004), one is VC age, which is the difference between the VC's last investment in year t and the VC's founding year. The second proxy is the VC's assets under management, which is a direct measure of its dollar resources. Finally, following Hochberg, Lindsey, and Westerfield (2015), we consider eigenvector centrality, which measures influence based on a weighted sum of a VC's total number of connections. Greater centrality implies more influential VCs. See the first three variables in Table 4.

A second set of attributes reflects a VC's investing *style*, i.e., the set of specific asset classes or investment types that the VC focuses attention on. Placement memorandums used for fund raising articulate investing styles. While these descriptions are not legally binding, they have bite as they form the basis on which limited partners allocate capital. In fact, Coller Capital's 2008 Global Capital Barometer report finds, 84% of limited partners do not look upon changes in stated styles (or style drift) favorably.

Existing literature suggests three major dimensions of VC style: industry, stage and geography (Sorenson and Stuart (2001); Chen, Gompers, Kovner and Lerner (2010); Tian (2011)), quantified in the second set of variables in Table 4. It is common for a VC fund to identify an industry focus in the formal agreement with limited partners (Lerner et al. (2007)). Anecdotal evidence is also consistent with these dimensions of functional styles. An illustration of the three dimensions of focus is the case of SV Angel, a "seed-stage" fund with "... tentacles into New York media and the advertising world."[15]. Thus, we consider style based on a VC's past investments in different sectors and portfolio firm stages as reported by VentureXpert. For geography, we consider both the location of the portfolio companies

[15] http://techcrunch.com/2011/05/24/sv-angel-partners-with-lerer-ventures-to-cross-syndicate-valleynyc-deals.

in which the VCs invest and the location of the VC's headquarters. Investing experience in a region results in information flows about the resources in the region while the location of senior management of a VC can also identify hard and soft information about investments in a geographic area.

5.2 Test Statistics

To examine the nature of VC partner preferences across attributes, we compute the within-community variation of measurable VC attributes. We benchmark these results by comparing the variation to similar variation for simulated communities. Lower within-community variation relative to the null indicates a preference for similar VCs as syndicate partners.

For the simulations, one possibility is to replicate communities in terms of the count and size distribution of communities observed in the data and assign VCs randomly to different communities. However, the distribution may not replicate the data at the round level. Thus, we simulate the entire dataset at the financing round level. We then generate communities for each such simulated dataset. We run these simulations 100 times to generate the null distributions. Further, the simulations are *conditional* on characteristics of financing rounds. We consider three features of rounds, namely industry, stage of funding, and the region in which the portfolio company is located. We draw VCs for a given financing round without replacement from a pool of VCs who have at least one investment that meets these criteria. In the event that such a set of VCs is a null set, we match on any two criteria instead of three. Most matches do not hit the constraint. The number of VCs drawn for a financing round matches the actual syndicate size.

With regard to continuous characteristics (such as age, VC's assets under management (AUM) and centrality), the within-variation is the standard deviation of the attribute for VCs within the community, averaged across all communities. For discrete variables, we consider a style similarity measure based on the Herfindahl-Hirschman Index (HHI) by category for each VC. If VCs are functionally similar, they should have similar HHIs. The dispersion of HHIs measures the functional similarity of VCs. HHIs can be computed based on the number of transactions in each style bucket or the proceeds. We report the former but obtain similar results for the latter.[16] While HHIs vary from VC to VC, we also consider the geographic

[16]Both measures are reasonable. The key resource in a VC firm is partner time, which first order scales by the number of investments. On the other hand, the capital at risk is proportional to the proceeds invested

headquarter concentration of all VCs for a community. We compute the HHI based on the proportion of a community's VCs in each geographic location.

Besides HHI, we consider a second, more stringent, measure that incorporates both the extent of specialization of a VC *and* the specific sectors the VC specializes in. Consider, for instance, one VC with 100% focus in software and another with 100% focus in biotech. Both will have sector HHI's equal to 1.0. The community containing both VCs will (correctly) show no dispersion in HHIs as the two VCs are similarly concentrated in allocating their capital proceeds. However, the two VCs differ in the specific sectors they focus on, which is not picked up by the HHI variation. To capture the differences in the sectors receiving capital allocation, we compute standard deviation of the fraction of deals in each bucket across all VCs in a community. We average the standard deviation across all buckets within a community and then across all communities. Formally, let the fraction of assets flowing into bucket j for VC i in community k be f_{ijk}. We compute the standard deviation $\sigma_{jk} = \sum_{i=1}^{n} \frac{(f_{ijk} - \overline{f_{jk}})^2}{n-1}$. The average standard deviation across all k communities indexes similarity in functional focus.

Operationally, we use five stage variables in VentureXpert (early stage, expansion, later stage, other, and startup/seed), the 10-industry classification used in VentureXpert, and experiment with a variety of geographical clusters. We use the very granular MSA to state-based locations to the 14-region classifications (e.g., Northern California, Southern California, New England). We obtain similar results under all approaches.

5.3 Results on Within-Community Variation

Table 4 reports the average characteristics of VC community members in our sample. The average age and size of VCs in a community are no different than among VCs in simulated communities. However, community VCs have greater centrality and are less focused in terms of industry, stage, and geography of portfolio companies compared to simulated communities.

Table 5 tests hypotheses about within-community similarity and dissimilarity. We report the within-community variation measures for various attributes for both observed communities and simulated communities, and the p-value based on simulated communities. Panel A

---

in a firm.

reports the results for attributes that proxy for VC reach and influence. Preferred partner groups appear to be more homogeneous, both in terms of age and centrality, than VCs in simulated communities. The differences are economically significant and statistically significant at the 1% level. On size as a proxy for reach, within-community variation in assets under management (AUM) is marginally higher, at the 10% significance level, in observed communities than in simulated communities. Overall, we find evidence in support of within-community similarity in VC reach and influence.

Panel B in Table 5 focuses on attributes relating to functional styles. We find clear evidence of homogeneity along each dimension of style - industry, stage, and location (based on the 14-region classification given by VentureXpert which reflects relatively homogeneous clusters of operation of VCs).[17] The first three rows of the Panel show that communities tend to have lower variation in HHIs than simulated communities. Thus, generalist VCs prefer as partners other generalist VCs, while concentrated VCs prefer to syndicate with other concentrated VCs. However, from Table 4, we know that on average, communities are more likely to have generalist VCs.

The rows below the HHI statistics in Panel B report variation within communities by percentages invested in each industry, each stage, or each geographical area of portfolio companies, respectively. Here, the bar for similarity is higher. A low variation requires not only a pairing of generalists (specialists) with other generalists (specialists) but also matching on the specific areas of functional expertise. For instance, low variation in industry implies that focused VCs prefer as partners other focused VCs *and* that both have relatively similar distributions of deals across specific industry sectors. The results are in line with those based on variations in HHI. Panels C and D focus on the location and ownership of VC firms. We find no statistically significant evidence that communities draw VCs from similar geographies compared with the location of VCs in simulated communities. Similarly, we do not find that VCs within a community have similar ownership form.

One concern about our analysis based on the share of deals in each bucket may be about the large number of empty or sparsely populated style cells. In other words, a VC may have no investment in a large number of styles, whether industry, stage, or location. While this concern is largely mitigated by our practice of benchmarking relative to simulated commu-

[17] Our results are qualitatively similar when location is based on the more granular MSA classifications and the less granular level of the state.

nities, we also consider an alternative approach that focuses on well populated cells. In each time period of 5 years, we consider within-community variation in the fraction of deals in the top 4 industries, top 2 stages, and top 3 geographic areas of portfolio companies identified in each sub period. Table 6 reports similar results for variations within each bucket separately, for industry, stage, and geographic region. VCs within communities exhibit lower variation in each of the separate buckets of industry, stage, and geographic region, providing evidence of similarity among community VCs in terms of functional expertise as well.[18]

In sum, the results suggest homogeneity in partner preferences based on reach, influence, and functional styles.[19]

5.4 Spatial Loci of VC Communities

While the previous analysis focused on differences *within* communities, this section tests the differences *between* communities by examining the spatial loci of communities along the style dimensions mentioned above.

Communities comprise VCs who are likely to partner with each other. Communities could comprise VCs with specialized skills. For instance, specialist knowledge may be required to finance clean energy due to technical expertise or knowledge of sector-specific regulations. Special expertise may be needed for assessing novel therapeutic protocols for cancer treatments. The opinions and judgments of multiple VCs specializing in the same sector become useful and VCs should seek preferred partners in the same skill area. This model of differentiation suggests that communities choose different spatial locations. The differences *between* communities would then be complementary to similarity *within* communities.

The alternative viewpoint is that communities are similar to each other. This is the soft conglomerate view of communities, consistent with a view of VC syndication in which communities help pool the ability to work with each other rather than specific functional or domain knowledge in narrow areas. "Generalist" communities could arise if there is generalized management skill, as in Lucas (1978) or Maksimovic and Phillips (2002) that is

[18]The top industries, stages and geography of interest change over time. For instance, consumer products is in the top-4 in the early 1980s, but Internet industry replaces it after the 1990s. Cosine similarity measures as in Hoberg and Phillips (2010) give similar results.

[19]We also examine cosine similarity of VCs to judge the aggregate similarity between VCs across all attributes within a community and find a preference for the similar.

important to all forms of VC investing. For instance, early stage firms may have good ideas within specific functional domains but may lack the organizational, management, or financial expertise to scale the ideas and translate them into successful businesses. If this type of skill is important and scalable across a broad range of firms seeking venture financing, we should observe communities choosing similar locations in the VC attribute space.

We identify the centroid of communities along each style dimension. The centroid of a community is the vector of values that is the average of characteristics across all members of the community. For example, if we consider two characteristics, such as age and assets under management, the centroid will be a two-vector containing the average age and average assets under management over all members. We test whether the style distance between community centroids is greater than what we observe for simulated communities. Greater distance implies specialization and differentiation while spatial proximity along a style dimension implies that it is not a basis for differentiation. Table 7 reports the results. We find little evidence of specialization. There is little evidence that VCs compete through differentiation.

6 Innovation

We study innovation output at the portfolio company and financing round level, classified by community status. The first investment year in our analyses is 1985 and community status is inferred from syndication data from 1980 to 1984. We treat a financing round with at least one community-based VC as a community round.

6.1 Innovation Measures

We consider two widely-used measures of scale and quality of a portfolio company's innovation. One is the total number of patents applied for (that were finally granted) by a company in each year. The second is the average number of forward citation counts per patent, which measures the quality and value of patenting (Hall, Jaffe and Trajtenberg (2005)).

There are two key challenges with the use of patent citations. Citations suffer from truncation bias. A patent receives citations for many years after it is applied for and granted. As a result, patents granted towards the end of the sample period tend to have fewer citations. We use the estimated shape of the citation-lag distribution as a correction (Hall, Jaffe and

Trajtenberg (2001)). To address the variation of patenting and citation rates across industries and over time, we follow Seru (2014) and adjust the innovation measures for the average innovation in the same cohort to which the patent belongs. We use the application year and technology class available in the patent database to define these cohorts. Finally, we take logs and add one to both the patent count and citation variables.[20]

Table 8 describes basic properties of the innovation measures, the total patents and citations per patent adjusted for truncation bias, at the portfolio company $\times$ financing round level. Based on the full sample, Panel A shows that an average round in our sample has 0.36 patent applications and 3.68 average citations. On both measures, community rounds score higher than non-community rounds. In Panel B, we scale the two innovation measures by the average innovation in the same cohort. Compared with Panel A, the magnitudes are significantly smaller. However, the results are qualitatively similar. Community rounds are more innovative than non-community rounds.

Many firms in the VentureXpert sample are not active in patenting. At the portfolio company level, the median number of patents applied for is zero. Only 16% of the portfolio companies in our sample file a patent application. Thus, in Panel C, we consider only those financing rounds with at least one application for a patent. While excluding zeros predictably increases the number of patents and average citations, the differences between community and non-community rounds hold. These results indicate that the relation between community membership and patenting output is not driven by zero versus non-zero patenting. Rather, the positive relationship also obtains within samples of patent filers.

6.2 Multivariate Analysis

We turn to multivariate specifications next to test whether sourcing financing from a community VC is associated with greater innovation. We use the following specification with Total Patents and Average Citations as dependent variables:

$$y_{it} = \alpha + \beta \cdot Community_{it} + \delta \cdot \mathbf{X}_{it} + \mu_t + \mu_j + \epsilon_{it} \tag{2}$$

[20]Following Bernstein, Giroud and Townsend (2015), we address the truncation issue by running all our specifications with citations within the first three years of the application date. A slight difference is that the Bernstein, Giroud and Townsend (2015) count is as of the grant date rather than the application date. All our results with this alternative citation variable are qualitatively similar.

where i indexes portfolio companies, t indexes years (or financing round, depending on context), and j indexes portfolio company's industry. In Equation (2), y, the dependent variable of interest is total patents or average citations and *Community* is a dummy variable that indicates whether or not there is a community VC in the financing round. $\mathbf{X}$ is the vector of control variables. We report results for the full sample and also analyze heterogeneity to exploit samples in which community VCs are likely to have greater effects.

We briefly comment on the key controls here. We include a dummy variable for syndication, following the robust finding in the VC literature (e.g., Brander, Amit and Antweiler (2002)) that syndication predicts success. The variable also helps reinforce the distinction between syndication and community, the X variable of interest. We control for the stage of financing through the variable *Early Stage*, and, following Chen, Gompers, Kovner and Lerner (2010), we include controls for whether a VC is headquartered in the geographical agglomerates of California or Massachusetts. We control for VC experience and skill (Krishnan and Masulis (2011)). We include assets under management, centrality, *IPO Rate*, or the rate at which a VC takes its firms public, and *Experience*, the average age of the participating VCs as of the year before the financing round. We control for whether VCs are arms of financial institutions or corporate VC investors. We capture VC focus through the variables *Early Stage Focus* and *Industry Focus*, which are fractions of firms in the focus areas funded by the VC syndicate. All models include year and industry fixed effects.

Table 9 reports the results. In the base case in columns (1) and (2), we see that community rounds have a positive and significant coefficient. To allow for persistence in innovation, we control for past innovation, which is measured by the 3-year forward citations per patent issued in the 3 years prior to the financing round. We continue to find a positive and statistically significant coefficient for *Community*. Column (4) includes all other controls. We no longer find the *Community* dummy to be statistically significant although it is still positive. Column (5) uses average citation as the dependent variable, and we find a positive and significant effect of community VC on innovation, at the 10% level.

The control variable coefficients are interesting. We find mixed results for financing stage depending on whether the dependent variable is the scale of innovation or quality of innovation. Early stage firms patent less but when they do, they have more citations, perhaps reflecting the high risk and greater tail outcomes in early stage investments. Portfolio companies located in the geographical clusters of CA/MA exhibit positive and significant coefficients. Syndication has a positive effect on innovation. Greater VC networks, as reflected in centrality, and focus on the specific industry in which the portfolio company lies are associated with more innovative activities.

6.2.1 Heterogeneity by Financing Stage

Informational asymmetry is likely to be more pressing in early stage financings. Accordingly, in Table 10, we examine 22,923 early financing rounds in our sample. We also classify these financings based on whether the sample firm has prior history of patenting or not. Thus, we consider three specifications, one for the full sample and one for two subsamples based on whether there is prior history of patenting.

We estimate the multivariate specification (Eq. 2) and include all the controls from the full specifications (4) and (5) in Table 9. The variable of interest is the interaction between early stage and community round (i.e., $Early \times Comm$). A positive coefficient would suggest improved innovation in early stage ventures when there is a community VC. We find that community VC coefficient is significant in two instances. In the right panel, we see that community VC participation matters in all three specifications, although at only 10% for firms with prior histories of patenting. The community VC coefficient is significant in patenting volume for early stage ventures that do not have prior histories of patenting.

6.2.2 Heterogeneity By Community Types

We explore heterogeneity through the characteristics of communities that may drive the coefficients for innovation. As discussed in Table 4 and Table 5, a taxonomy of communities is the mean level of each characteristic and the standard deviation of a characteristic. The first represents the average VC in a community and the second captures the heterogeneity in a community. The attributes we consider are age, size, centrality, and the extent of portfolio concentration (HHI) based on industry, stage, and portfolio company's region. For each of these, we consider the mean for and the standard deviation within the community.

Because the focus is on community characteristics, the results in Table 11 are based on the sample of rounds with a community VC. Additionally, deals that are syndicated potentially include multiple community VCs. Thus, we include the average of the mean or standard deviation of the relevant community characteristic for all community VCs in the financing round. For brevity, we do not report coefficients for other controls.

We find that community VCs with high stage HHI are associated with more innovation. Put differently, communities that comprise VCs with more homogeneous early-stage focus have more significant coefficients. In addition, communities with greater breadth of industry experience (low industry HHI) and greater similarity in VC age (low Community Age SD) have more positive coefficients. Additionally, we find that high centrality among community VCs (high centrality, mean) and greater similarity in centrality among community VCs (low Centrality, SD) have positive coefficients on innovation quality as decided by citation. These two results are consistent with the view that more homogeneous communities are more effective in driving innovation.

6.2.3 Within Portfolio Analysis

We next consider the subset of portfolio companies that did not have a community VC in the first financing round but had a community VC in the second round. For this subsample, we include all subsequent financing rounds in which they had continuous community VC participation. If there was a round in which there was no community VC, we drop that round and all subsequent financing rounds for that portfolio company from our sample. This stringent sample, therefore, allows for a very pointed view on the role of community VC in innovation. The sample also mitigates concerns regarding the long-term effect on innovation of non-community VCs who may have participated in portfolio companies where previous rounds included community VCs. To mitigate concerns about unobservable portfolio company characteristics, we also conduct our analysis with portfolio fixed effects.

We present the results in Table 12. The specifications include the usual controls but sample size is smaller. The results are in the same direction as those in Table 10 but stronger. The coefficient for community VCs in early stage rounds is significant in portfolio companies with no prior patenting. The coefficient is 5.5% for patents and 17% for citations. Both coefficients are significant.

7 Performance of Portfolio Companies

Our final tests examine the relation between financing sourced from a community VC and exit. Successful exit in the venture capital context could be an IPO or an M&A transaction. An important issue in the VC literature is whether exits via M&A are indicators of success. Kaplan, Sensoy and Stromberg (2002) and Maats, Metrick, Hinkes, Yasuda and Vershovski (2008) use small subsets of venture capital data and find that IPO exits are accurate indicators of success. However, M&A exits have greater noise, divided between failed financings and success stories.

There is no received consensus on how best to deal with the noise in M&A exits in large scale VC studies such as ours.[21] The VC literature's response is to (a) be conservative using only IPO exits (Gompers et al (2012)); and (b) use multiple measures of exit. We follow these suggestions. We model exits via IPOs. We move the literature forward a notch by using a competing hazards model in which an M&A is a competing hazard for an IPO. This specification refines the IPO-only definition of successful exit by recognizing that IPOs are observed only if an M&A does not precede the IPO. For robustness, we also follow Hochberg, Ljungqvist and Lu (2007) and use the event of an M&A or an IPO.

We consider investments between 1985-2010 and track exits until December 31, 2012. In our sample, we have 2,545 unique IPOs and 5,106 M&A transactions. However, since our analysis is at the level of the financing round, Panel D of Table 3 shows that 23,476 (or 32%) of financing rounds exit through IPOs or M&As. IPOs account for 11%, or a third of these. In community rounds, 13% exit through IPOs and 25% exit through mergers, compared to 9% and 19% for non-community rounds, respectively. We find similar patterns when considering exits classified by the number of portfolio companies rather than the number of financing rounds. Finally, 69% of all rounds with a community VC proceed to a next round of financing compared to 57% of the rounds with no community VC.

We turn to multivariate models next. Table 13 reports the results. The key independent variable of interest is the *Community* dummy, which equals 1.0 if the round has at least one community VC and zero otherwise. Each of the specifications includes the set of controls

[21] Coding every M&A exit is a labor-intensive exercise that requires a search of news stories. The firms are private, not widely followed in the press, and the news stories do not let us sharply discriminate between failures and successes, especially in our sample period of 1980-2010. This exercise is likely worth a study in its own right.

we used in the baseline innovation analyses including an early stage dummy, geographical cluster controls, syndication, VC experience, skill, past rate of IPO exits, VC age, and VC type. For brevity, we do not report these coefficients.

Specification (1) in Table 13 is a Cox proportional hazards model in which success is exit through either M&A or an IPO. We report Cox estimates in the form of an exponentiated hazards ratio where a coefficient greater than 1.0 for a variable indicates that it speeds exit. The hazard ratio for community VC is 1.13 and is significant at the 1% level. Thus, community VC financing speeds exit. The probit result in specification (2) mimics the Cox result. Community VC is associated with a 9% increase in the likelihood of a successful exit.

Specifications (3) - (5) in Table 13 report estimates of competing hazards models. Specification (3) defines IPOs as success but recognizes that IPOs are observed only conditional on not having a prior M&A.[22] We find similar results - community VCs speed up exit. Finally, in specifications (4) and (5), we show results classified by the round of financing. As before, community effects are pronounced in round 1, where sourcing financing from a community VC speeds exit by 7%. The screening of firms is likely more intensive in the first round when all syndicate partners confront the portfolio firm for the first round. Trusted partners sourced from communities appear to matter more in these rounds than in subsequent ones.

8 Conclusion

Syndication is a pervasive feature of VCs financing. Over a period of time, the typical VC does syndicated deals often and with multiple partners. However, VCs are more likely to select some partners over others. The preference for a small set of partners results in the formation of several clusters of VCs, economic agglomerates that we term as VC communities. We characterize these new types of agglomerates, which are akin to soft conglomerates with informal and porous borders.

Our study makes precise the mathematical notion of VC communities and discusses their detection using syndication data. We discuss the complexity of the underlying computational problem because the number of clusters is not fixed and each may be of a different size. We use a flexible and fast computational method to detect communities using a large sample

[22]We get similar results if we consider IPOs alone without accounting for the selection effect due to a (non)merger when there is an IPO.

of syndications from 1980 to 2010. We then characterize their composition, their spatial locations relative to each other, and their economic activities, specifically the quantity and quality of innovation as reflected in patenting behavior and the maturation of firms they finance, as reflected in exits.

We identify several communities in each 5-year sub-period of our sample. About 20% of all VCs cluster into communities. We find that communities are both stable and tight-knit, suggesting that VCs prefer familiar syndicate partners. As Granovetter (1973) points out, partner preferences are potentially complex, reflecting the lower transaction costs, informational asymmetries, and contracting costs of dealing with similar-style partners and the converse incentive to reach outside comfort zones to generate new business and learning. We find that the preferences for similar style venture capitalists is the dominant force in the data. VCs within clusters tend to be similar in age, reach, and functional styles relative to simulated clusters that control for heterogeneity. The results support models such as Cestone, Lerner and White (2014) in which VCs syndicate with similar partners.

We find little evidence that VC communities occupy different style niches. Such a finding would be consistent with the view that clusters are formed to let VCs compete through differentiation in focus or style. Rather, communities appear to collect a range of VC skills. This behavior is conglomerate-like, indicating that VC communities compete across a broad range of financing opportunities. Alternatively, following Lucas (1978) or Maksimovic and Phillips (2002), generalized management skill can be important to all forms of VC financing. For instance, early stage firms may have good ideas within specific functional domains but may lack the organizational, management, or financial expertise to scale the ideas and translate them into successful businesses. The similarity of communities may reflect the demands for generalized skills for young firms.

The critical function of venture capitalists is to support innovation and to mature young firms from business plan and concept to a mature enterprise. We thus examine two key economic metrics associated with community VCs, viz., innovation and firm maturation. We find that community VC financing is associated with both quantity and quality of innovation, as measured by patenting and patent cites, with the effects particularly pronounced for early stage firms and firms without prior patenting. Using several metrics, we also find that community VC financing is also associated with greater likelihood of maturation of portfolio firms. We see this effect in models of exit via IPOs, exit via M&As or IPOs, or exit via IPOs with M&As as a competing hazard, and with progression to next stage financing.

Our study mainly contributes to the venture capital literature on syndication. The literature argues that syndication is beneficial as it permits risk-sharing, access to diverse resources, and second opinions on risky investments. Syndication can also pose a fresh set of problems for venture capitalists. Suspicions of ex-post hold up and free riding by partners can lead to insufficient effort and undo the benefits of syndication. Syndicating with familiar partners can mitigate these problems by reducing information asymmetry, building trust, and enhancing reciprocity between partners. Alternatively or additionally, familiar partners can enhance learning in models in which VCs learn by doing. The propensity to pick preferred syndicate partners can be interpreted as an outcome of these forces. Our study points out that syndication with the familiar results in preferences for some partners over others, which in turn implies the formation of economic agglomerates comprising VCs who tend to syndicate with each other. Our study can be viewed as a test of this implication.

Our approach towards analyzing the existence and nature of partner preferences has many applications outside VCs. In finance, syndications are even more pervasive in the commercial banking area and the investment banking field, where, for instance, syndicates form for underwriting public issues. The nature of partner preferences in these areas constitute a profitable avenue for future research. Yet another area of potential concerns the theory of the firm, specifically inter-firm alliances. Robinson (2008) highlights that simultaneous, non-overlapping inter-firm collaborations are quite common even between firms that compete in some product markets. An interesting question is whether these types of collaborations also exhibit preferred-partner clustering. Our study provides a coherent framework and a set of practical tools to analyze such questions. One can then better understand the *entire set* of partnerships of firms rather than the one-off ties that firms form.

Finally, communities appear to be interesting organizational forms that lie in between formal conglomerates and firms demarcated by legal organizational boundaries. Spot contracting between legally separate entities helps avoid the inflexibility and complexity of running large conglomerates. However, it also compromises the benefits of soft information flows and relationships from an integrated conglomerate. Communities can be regarded as organizational intermediates that provide some benefits of both forms of organizations, lying somewhere in between hard-boundary conglomerates that internalize all transactions and arms-length spot contracting with outside partners.

Appendix

A Detecting Communities and 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 **?**. 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

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{d_i \times d_j}{2m} \right] \cdot \delta_{ij} \tag{3}$$

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
Age	Number of years between a VC's last investment in year t and the VC firm's founding year
AUM	Total capital under management, in $ million, of all those VC funds that invested during a 5-year rolling window
AUM_Round	Natural log of one plus the average AUM, in $ million, of the participating VCs' funds that invested until the year prior to the financing round
Average Citations	Natural log of one plus the forward citations per patent applied for in the financing year corrected for truncation bias (based on Hall, Jaffe, and Trajtenberg (2001)) and adjusted for average citations of patents applied for in the same cohort (application year and technology class)
Centrality	VC's eigenvector centrality based on syndicated rounds during a 5-year rolling window
Community	Equals 1.0 if there is at least one community VC in the financing round and zero otherwise
Community Age, Mean(SD)	Mean (Standard deviation of) age of VCs within a community
Community Size, Mean(SD)	Mean (Standard deviation of) AUM managed by each VC within a community
Comm Centrality, Mean(SD)	Mean (Standard deviation of) centrality of VCs within a community
Co. Geog Cluster	Equals 1.0 if the portfolio company funded by the VC is in the state of California or Massachusetts and zero otherwise
Co. Region HHI	Herfindahl-Hirschman Index based on a VC's (or a community of VCs') share of total deals in each geographical region during a 5-year rolling window
Co. Region HHI, Mean(SD)	Company Region HHI of a community of VCs (standard deviation of individual VC's Company Region HHI within a community)
Company Region Rank i	Geographic region with the i^{th}-highest aggregate number of deals invested in by all VCs in a 5-year rolling window
Company Region Variation	Squared deviation of the proportion of a VC's (or a community of VCs') deals in a geographic region from the average of all VCs' (or all communities') proportions in the region, averaged across all VCs (or communities) and regions, during a 5-year rolling window
Corporate VC	Equals 1.0 if there is at least one venture capitalist in a financing round who is the corporate VC arm of a firm
Early Stage	Equals 1.0 if the round is an early stage financing and zero otherwise

B. Variable Definitions - Contd.

Variable	Description
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
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[23]
FI VC	Equals 1.0 if there is at least one financial institution VC in 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
Industry HHI	Herfindahl-Hirschman Index based on a VC's (or a community of VCs') share of total deals in each industry during a 5-year rolling window
Industry HHI, Mean(SD)	Industry HHI of a community of VCs (standard deviation of individual VC's Industry HHI within a community)
Industry Rank i	Industry with the i^{th}-highest aggregate number of deals invested in by all VCs in a 5-year rolling window
Industry Variation	Squared deviation of the proportion of a VC's (or a community of VCs) deals in an industry from the average of all VCs' (or all communities') proportions in the industry, averaged across all VCs (or communities) and industries, during a 5-year rolling window
IPO Rate	Natural log of one plus the average of each participating VC's ratio of IPOs to number of portfolio companies invested in the last three years prior to the financing round[24]
Ownership HHI	Herfindahl-Hirschman Index based on the proportion of VCs in a community from each ownership type (e.g., independent private equity, corporate VC, financial institution VC arm, others)
Past Innovation	Natural log of one plus the average 3-year forward citations per patent in a company's portfolio, averaged over company's patents applied for in the 3 years prior to the financing round
Prior Innovations	Equals 1.0 if the portfolio company applied for at least one patent in the 3 years prior to the financing round, and zero otherwise

[23] Our definition modifies Lindsey(2008)'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.

[24] Krishnan and Masulis (2011)

B. Variable Definitions - Contd.

Variable	Description
Stage HHI	Herfindahl-Hirschman Index based on a VC's (or a community of VCs') share of total deals in each stage during a 5-year rolling window
Stage HHI, Mean(SD)	Stage HHI of a community of VCs (standard deviation of individual VC's Stage HHI within a community)
Stage Rank i	Financing stage with the i^{th}-highest aggregate number of deals invested in by all VCs in a 5-year rolling window
Stage Variation	Deviation of the proportion of a VC's (or a community of VCs) deals in a stage from the average of all VCs' (or all communities') proportions in the stage, averaged across all VCs (or communities) and stages, during each 5-year rolling window
Syndicated	Equals 1.0 if the round is syndicated, zero otherwise
Total Patents	Natural log of one plus the number of patents a portfolio company applies for in the financing year, adjusted for total patents applied for in the same cohort (application year and technology class)
VC Geographical Cluster	Equals 1.0 if at least one participating VC in a financing round is in the state of CA or MA
VC MSA HHI	Herfindahl-Hirschman Index based on the proportion of VCs in a community from each MSA
VC Region HHI	Herfindahl-Hirschman Index based on the proportion of VCs in a community from each geographic region
VC State HHI	Herfindahl-Hirschman index based on the proportion of VCs in a community from each U.S. state

References

Baker, G., Gibbons, R., and Murphy, K. (2008). Strategic Alliances: Bridges between "Islands of Conscious Power", *Journal of the Japanese and International Economies* 22, 146-163.

Barabasi, A.L.; R. Albert (1999). "Emergence of scaling in random networks," *Science* 286 (5439), 509-512.

Barber, B., Lyon, J. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics, *Journal of Financial Economics* 43, 341-372.

Bengtsson, O., Hsu, D. (2010). How do venture capital partners match with startup founders? *Working Paper*, Wharton School, University of Pennsylvania.

Bernstein, S., Giroud, X., Townsend, R. (2015). The impact of venture capital monitoring, *Journal of Finance*, Accepted Author Manuscript. doi:10.1111/jofi.12370.

Bhagwat, V. (2013). Manager networks and investment syndication: Evidence from venture capital. *Working Paper*, University of Oregon. Available at SSRN: `http://ssrn.com/abstract=1962146` or `http://dx.doi.org/10.2139/ssrn.1962146`.

Bottazzi, L., Da Rin, M., Hellmann, T. (2011). The importance of trust for investment: Evidence from venture capital. *Working Paper*, NBER W16923.

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* 11, 423-452.

Brown, S., Goetzmann, W. (1997). Mutual fund styles. *Journal of Financial Economics* 43, 373-399.

Brown, S., Warner, J. (1985) Using daily stock returns: The case of event studies. *Journal of Financial Economics* 14, 3-31.

Burdick, D., Hernandez, M., Ho, H., Koutrika, G., Krishnamurthy, R., Popa, L., Stanoi, I.R., Vaithyanathan, S., Das, S.R. (2011). Extracting, linking and integrating data from public sources: A financial case study, *IEEE Data Engineering Bulletin* 34(3), 60-67.

Cestone, G., Lerner, J., White, L. (2006). The design of syndicates in venture capital, *Harvard Business School Working Paper*. Published as *Fundacion BBVA* White Paper.

Chen, H., Gompers, P., Kovner, A., Lerner, J. (2010). Buy local? The geography of successful venture capital expansion, *Journal of Urban Economics* 67(1), 90-102.

Chidambaran, N. K., Kedia, S., Prabhala, N.R. (2010). CEO-Director connections and fraud, *University of Maryland Working Paper.*

Chung, S., Singh, H., Lee, K. (2000). Complementarity, Status-Seeking, and Social Capital As Drivers of Alliance Formation, *Strategic Management Journal* 21(1), 1-22.

Cohen, L., Frazzini, A., Malloy, C. (2010). Sell-Side school ties, *Journal of Finance* 65, 1409-1437.

Cohen, L., Frazzini, A., Malloy, C. (2012). Hiring cheerleaders: Board Appointments of Independent Directors, *Management Science* 58(6), 1039-1058.

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

Currarini, S., Jackson, M., Pin, P. (2012). An Economic Model of Friendship: Homophily, Minorities, and Segregation, *Econometrica* 77, 1003-1045.

Du, Q. (2011). Birds of a feather or celebrating differences? The formation and impact of venture capital syndication, *University of British Columbia Working Paper*. Available at SSRN: `http://ssrn.com/abstract=1360545` or `http://dx.doi.org/10.2139/ssrn.1360545`.

Engelberg, J., Gao, P., Parsons, C. (2013). The value of a rolodex: CEO pay and personal networks, *Review of Financial Studies* 26(1), 79-114.

Fortunato, S. (2009). Community detection in graphs, *Physics Reports* 486(3-5), 75-174.

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.

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

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

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

Gompers, P., Mukharlyamov, V., Xuan, Y. (2012). The cost of friendship, forthcoming *The Journal of Financial Economics.*

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

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology* 78(6), 1360-1380.

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.

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

Gulati, R. (1995). Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances, *The Academy of Management Journal* 38(1), 85-112.

Hall, B., Jaffe, A., Trajtenberg, M. (2001). The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools, *NBER Working Paper 8498.*

Hall, B., Jaffe, A., Trajtenberg, M. (2005). Market value and patent citations, *RAND Journal of Economics* 36, 16-38.

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

Hegde, D., Tumlinson, J. (2014). Does social proximity enhance business relationships? Theory and evidence from ethnicity's role in US venture capital, *Management Science* 60(9), 2355–2380.

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. (2015). Resource accumulation through economic ties Evidence from venture capital. *Journal of Financial Economics* 118(2), 245–267.

Hsu, D. (2004). What Do Entrepreneurs Pay for Venture Capital Affiliation?, *Journal of Finance* 59, 1805-1844.

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

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., Hardymon, F., Leamon, A. (2007). Venture capitalist and private equity: A casebook, *John Wiley and Sons*.

Lucas, R., 1978. On the size distribution of business firms. *Bell Journal of Economics* 9, 508–523.

Maats, F., Metrick, A., Hinkes, B., Yasuda, A., Vershovski, S. (2008). On the Completeness and Interchangeability of Venture Capital Databases, *UC Davis Working paper.*

Maksimovic, V., Phillips, G. 2002. Do conglomerate firms allocate resources inefficiently across industries? *Journal of Finance* 57, 721–776.

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

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. (2010). Networks: An introduction, *Oxford University Press*, New York.

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

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

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

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. RePEc:ecl:ohidic:2010-21.

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

Scharfstein, D., Stein, J. (2000). The dark side of internal capital markets: divisional rent-seeking and inefficient investment, *Journal of Finance* 55, 2537-2564.

Seru, A. (2014). Firm boundaries matter: Evidence from conglomerates and R&D activity, *Journal of Financial Economics* 111(2), 381–405.

Sorensen, M. (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.

Stanfield, J. (2013). How does syndication influence leveraged buyout success?, *University of New South Wales, Working Paper*.

Tian, X. (2011). The causes and consequences of venture capital stage financing, *Journal of Financial Economics* 101(1), 132–159.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly* 42, 35-67.

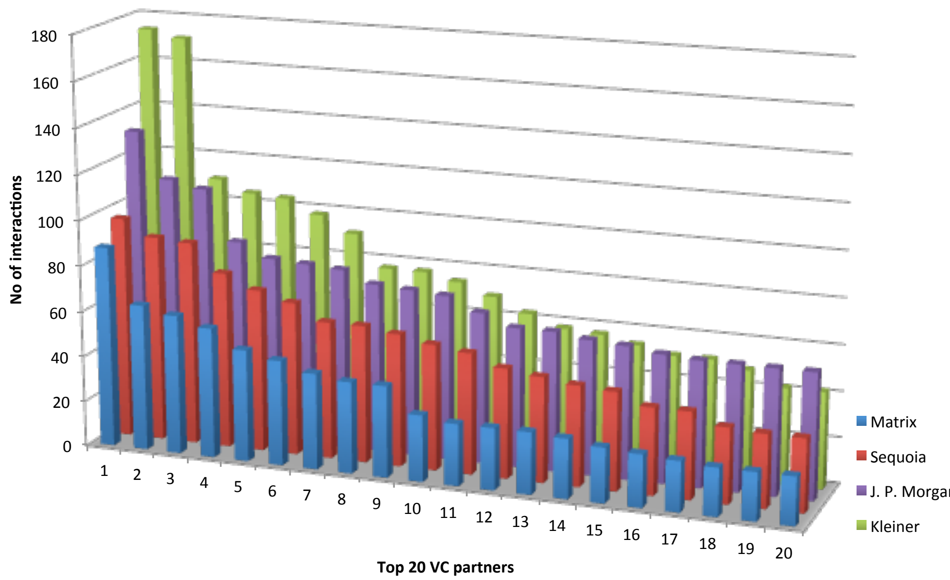

Figure 1: Distribution of the number of interactions of four top firms with their top 20 collaborators.

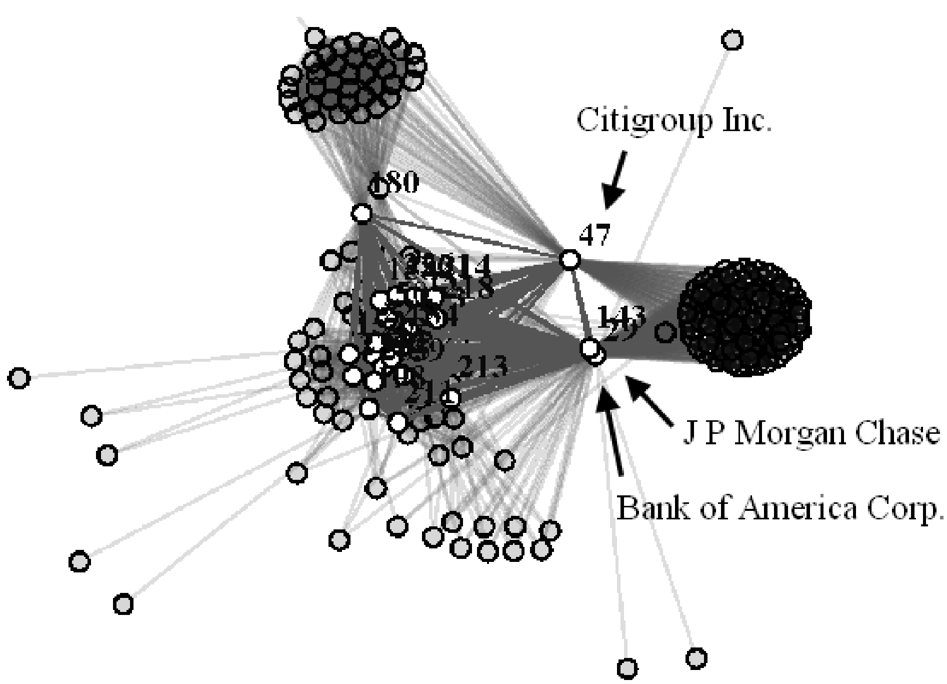

Figure 2: Communities and centrality in bank co-lending networks.

Table 1: Venture Capitalists in our sample. This table provides descriptive statistics of the 3,275 unique U.S.-based VCs in our database over the entire 31-year period, from 1980 to 2010. Data are from Venture Economics and exclude non-US investments, angel investors, and VC firms focusing on buyouts. We report the number of rounds of financing and the count of portfolio companies a VC invests in. Investment per round is the amount a VC invests in a round. % Deals Syndicated is the number of a VC's syndicated rounds as a percentage of all rounds that a VC invested in. % Early Stage Deals is the number of a VC's investment rounds classified by Venture Economics as early stage as of the round financing date, as a percentage of all Venture Economics deals for the VC between 1980 and 2010. AUM is the sum of the capital under management of a VC in all funds that invested during 1980-2010. Total investment is the sum of a VC's investments over this time period. Age is defined as the difference in the year of the VC's last investment in the period 1980 to 2010 and the VC firm's founding date. # VC firms per MSA is the total number of unique VCs headquartered a metropolitan statistical area (MSA). CA/MA VC is the fraction of all VCs that are headquartered in either California or Massachusetts.

Variables:	Mean	Median	# Observations
# Rounds	56.11	11.00	3275
# Companies	24.14	8.00	3275
Investment per round ($ mm)	4.25	2.74	3252
% Deals Syndicated	0.73	0.80	3275
% Early Stage Deals	0.30	0.27	3275
AUM ($ mm)	472.17	60.00	1859
Total Investment ($ mm)	285.26	32.73	3252
Age	11.50	9.00	3275
# VC firms per MSA	19.95	4.00	153
CA/MA VC	0.35	0.00	3275

Table 2: Stability of community status. The table provides data on the number of VCs who belong to community clusters in each 5-year window and the fraction of these that remain in a community after 1, 3, and 5 years from the initial window.

Window	# Community VCs	After 1 year	After 3 years	After 5 years
1980 -1984	87	1.00	0.98	0.93
1981 -1985	92	1.00	0.97	0.93
1982 -1986	101	1.00	0.97	0.92
1983 -1987	136	0.99	0.96	0.90
1984 -1988	139	0.98	0.93	0.81
1985 -1989	147	0.99	0.93	0.82
1986 -1990	112	0.99	0.94	0.88
1987 -1991	112	0.93	0.89	0.85
1988 -1992	83	0.98	0.96	0.89
1989 -1993	73	1.00	0.90	0.85
1990 -1994	71	0.99	0.94	0.93
1991 -1995	88	0.97	0.92	0.89
1992 -1996	101	0.96	0.92	0.94
1993 -1997	107	0.98	0.99	0.95
1994 -1998	129	1.00	0.98	0.95
1995 -1999	168	0.99	0.96	0.93
1996 -2000	248	0.99	0.98	0.86
1997 -2001	294	0.99	0.96	0.81
1998 -2002	297	0.99	0.88	0.76
1999 -2003	307	0.98	0.86	0.80
2000 -2004	257	0.96	0.88	0.84
2001 -2005	232	1.00	0.96	0.91
2002 -2006	186	0.99	0.98	
2003 -2007	198	0.99	0.98	
2004 -2008	194	0.99		
2005 -2009	202	1.00		
2006 -2010	207			

Table 3: Descriptive statistics for 73,414 rounds in 26,995 unique portfolio companies from 1985-2010. 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 but excludes non-US investments, angel investors and VC firms focusing on buyouts. Industry classifications are as per Venture Economics. Exit data are obtained by matching with Thomson Financial IPO and M&A databases. Data on follow-on funding is restricted to rounds financed until 2009.

Variable	Total	Community Round	Not Community Round
	Panel A: Counts By Round		
# Deals	73,414	32,547	40,867
—Round 1	23,253	7,232	16,021
—Round 2	15,771	6,773	8,998
—Round 3	11,237	5,626	5,611
Syndicated	35,088	22,812	12,276
Early stage	22,923	9,399	13,524
Geographical Cluster	35,414	21,074	14,340
Rounds with			
—Geographical Cluster VC	44,560	28,052	16,508
—Corporate VC	8,593	5,084	3,509
—FI VC	10,391	5,720	4,671
	Panel B: Percentage By Venture Economics Industry		
—Biotech	7.6	3.7	3.8
—Commu&Media	10.3	5.2	5.2
—Hardware	5.1	2.6	2.6
—Software	21.6	10.1	11.5
—Semiconductor, Electricals	7.4	3.9	3.6
—Consumer Products	5.0	1.3	3.7
—Industrial, Energy	4.9	1.4	3.6
—Internet	17.7	7.9	9.8
—Medical	13.1	6.5	6.6
—Others	7.2	1.8	5.4
	Panel C: Round Statistics		
Proceeds ($ million)	19 (8)	27 (15)	12 (4)
# VCs	2.09 (1)	2.83 (2)	1.50 (1)
—in syndicated rounds	3.34 (3)	3.69 (3)	2.69 (2)
—in early stage rounds	1.89 (1)	2.49 (2)	1.47 (1)
—in round 1	1.54 (1)	2.01 (2)	1.32 (1)
—in round 2	1.46 (1)	1.63 (2)	1.35 (1)
—in round 3	2.03 (2)	2.27 (2)	1.80 (2)
	PANEL D: Exit		
Rounds with			
—IPO exits	7,816	4,307	3,509
—M&A exits	15,660	8,049	7,611
—Follow-on funding	45,866	22,419	23,447

Table 4: Characteristics of Same-Community VCs. The table compares key community characteristics with those of simulated communities generated by choosing VCs who have invested at least once in the same industry, stage and region as the portfolio company. For each community (and simulated community), we generate the mean of the characteristic, and present the average value across 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 all VC funds that invested during a 5-year period. *Centrality* is based on each VC's eigenvector centrality determined for each 5-year rolling window. For the remaining attributes, we calculate the Herfindahl-Hirschman Index (HHI) as the sum of squared share in each subcategory of the attribute. *Industry HHI* is the Herfindahl index based on the % of a community VC's deals in each industry, while *Stage HHI* is the Herfindahl index based on the % of deals in each stage of investment. *Company Region HHI* is the Herfindahl index based on the % of deals in each geographic region. In unreported tests, we see similar results when we use HHI based on amount invested. The industry, stage and geographic region classifications are those provided by Venture Economics. 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	Simulated Community	p-value
Reach & Influence:			
Age	12.00	12.13	0.25
AUM	461.99	476.71	0.25
Centrality	0.15	0.14	0.05**
Style:			
Industry HHI	0.34	0.40	0.01***
Stage HHI	0.41	0.43	0.01***
Company Region HHI	0.40	0.44	0.01***

Table 5: Similarity of Within-Community VCs. The table presents variation in key attributes (in Panels A-B) and mean geographic location HHI (in Panel C) and ownership HHI (in Panel D) of VCs within communities, and compares these to those of simulated communities generated by choosing VCs who have invested at least once in the same industry, stage and region as the portfolio company. We calculate the Herfindahl-Hirschman index (HHI) as the sum of squared deviation of each subcategory of the attribute. *Age* uses the number of years between a VC's last investment in a 5-year rolling window and the founding year of the VC firm. *Assets under management (AUM)*, in $ million, uses the sum of all VC funds that invested during each 5-year rolling window. *Centrality* is based on each VC's eigenvector centrality determined for each 5-year rolling window. *Industry*, *Stage* and *Company Region* are based on % of a VC's deals in each of the 10 industries, each of the 5 stages, and each of the 14 U.S. geographic regions, respectively, as classified by Venture Economics. In Panels A and B, variations in Reach attributes and attribute HHI, respectively, are the standard deviation of each attribute of a community's VC, averaged across all communities. Variation in each attribute in Panel B measures the mean (across all communities) of the sum of squared deviation in each subcategory (e.g., Industry j) of each attribute (e.g., Industry) averaged across all subcategories and all within-community VCs. Panel C uses alternative geographic location variables, from the most granular (MSA) to the least granular (Region), and calculates the geographic HHI of a community's VCs, averaged across all communities. Panel D calculates the ownership HHI of a community's VCs, averaged across all communities. The last column shows the p-values testing the equality of the means of the community and simulated community characteristics. ***, **, and * denote 1%, 5% and 10% significance, respectively, from the test.

	Community	Simulated Community	p-value
	Panel A: Variation in Reach Attributes		
Age	7.56	7.87	0.01***
AUM	527.19	492.45	0.1*
Centrality	0.10	0.12	0.01***
	Panel B: Variation in Functional Styles		
Industry HHI	0.16	0.20	0.01***
Stage HHI	0.13	0.17	0.01***
Company Region HHI	0.18	0.21	0.01***
Industry Variation	0.11	0.13	0.01***
Stage Variation	0.08	0.11	0.01***
Company Region Variation	0.10	0.12	0.01***
	Panel C: Mean of Community Geographic HHI		
VC MSA HHI	0.28	0.29	0.25
VC State HHI	0.44	0.43	0.25
VC Region HHI	0.40	0.40	0.25
	Panel D: Mean of Community Ownership HHI		
VC Ownership HHI	0.60	0.59	0.25

Table 6: Functional Expertise Similarity of Within-Community VCs. We present the mean (across all communities) of the sum of squared deviation of VC's share of deal in some subcategories (based on number of deals in a 5-year rolling window in each of the top 4 industries, top 2 stages, and top 4 company regions, with the remainder share of investment comprising the last subcategory in each). We compare these values to those of simulated communities generated by choosing VCs who have invested at least once in the same industry, stage and region as the portfolio company. The last column shows the p-values testing the equality of the means of the community and simulated community characteristics. ***, **, and * denote 1%, 5% and 10% significance, respectively, from the test.

	Community	Simulated Community	p-value
Industry Rank:			
1	0.14	0.16	0.01 ***
2	0.12	0.15	0.01 ***
3	0.11	0.12	0.01 ***
4	0.09	0.11	0.01 ***
5 = Others	0.16	0.17	0.01 ***
Stage Rank:			
1	0.16	0.19	0.01 ***
2	0.15	0.17	0.01 ***
3 = Others	0.16	0.17	0.01 ***
Company Region Rank:			
1	0.18	0.21	0.01 ***
2	0.12	0.12	0.25
3	0.10	0.11	0.05 **
4	0.08	0.08	0.05 **
5 = Others	0.15	0.17	0.01 ***

Table 7: Similarity Across Communities. The table presents across community variation in (average) key VC attributes (in Panel A), in geographic location HHI (in Panel B) and in ownership HHI (in Panel C) of VCs within communities, and compares these to those of simulated communities generated by choosing VCs who have invested at least once in the same industry, stage and region as the portfolio company. We calculate the Herfindahl-Hirschman index (HHI) as the sum of squared deviation of each subcategory of the attribute. *Industry*, *Stage* and *Company Region* are based on % of a VC's deals in each of the 10 industries, each of the 5 stages, and each of the 14 U.S. geographic regions, respectively, as classified by Venture Economics. Panel A shows the mean values (across 5-year rolling windows) of standard deviation of the within-community HHI from the across-community average in each 5-year rolling window. Panel B uses alternative geographic location variables, from the most granular (MSA) to the least granular (Region), and calculates the standard deviation of geographic HHI of communities in each 5-year window, averaged across all such windows. Panel C calculates the standard deviation of ownership HHI of communities in each 5-year window, averaged across all such windows. The last column shows the p-values testing the equality of the means of the community and simulated community characteristics. ***, **, and * denote 1%, 5% and 10% significance, respectively, from the test.

	Community	Simulated Community	p-value
	Panel A: Variation in Functional Styles		
Industry HHI	0.08	0.09	0.25
Stage HHI	0.04	0.04	0.50
Company Region HHI	0.13	0.13	0.50
	Panel B: Variation of Community Geographic HHI		
VC MSA HHI	0.12	0.13	0.25
VC State HHI	0.20	0.19	0.25
VC Region HHI	0.17	0.17	0.50
	Panel C: Variation of Community Ownership HHI		
VC Ownership HHI	0.18	0.18	0.50

Table 8: Descriptive statistics for innovation outcomes from 1985-2006. The table provides mean values of innovation variables. The unit of observation is company $\times$ round. Total Patent is the number of patents applied for in the year of the financing round. Average Citations is the total number of citations per patent, corrected for truncation bias using Hall, Jaffe and Trajtenberg (2001) (as described in the text). Panel A considers all company $\times$ rounds. In Panel B, we adjust both innovation measures for the average innovation in the same cohort (based on application year and technology class) to which the patent belongs. Panel C uses the subsample where the total patent is positive. The last column tests for the equality of means of innovation outcomes for community and non-community rounds. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

Variable	Total	Community Round	Not Community Round	Mean equality
Panel A: All Observations				
Total Patents	0.36	0.46	0.29	***
Average Citations	3.68	4.99	2.66	***
Panel B: All Observations, Cohort-Adjusted Innovation Variables				
Total Patents	0.06	0.08	0.05	***
Average Citations	0.25	0.34	0.18	***
Panel C: Cohort-Adjusted Innovation Variables if Total Patent >0				
Total Patents	0.48	0.49	0.46	**
Average Citations	2.01	2.17	1.81	***

Table 9: Innovation and VC Communities. The table reports OLS estimates for innovation activity between 1985-2006. The dependent variable in specifications (1)-(4) is the logarithm of (one plus) the number of patents that portfolio company applied for in the financing year. In specification (5), the dependent variable is the logarithm of (one plus) the per patent citation count corrected for truncation bias. Both total patents and average citations are adjusted for average innovation in the same cohort (based on application year and technology class). The observations are at the company × round level. See Appendix B for a description of the independent variables. Standard errors are clustered at the portfolio company level. t-statistics are in parentheses. All specifications are overall significant at the 1% level. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	Total Patents				Avg Citations
	(1)	(2)	(3)	(4)	(5)
Community	0.022***	0.021***	0.014***	0.003	0.008*
	(12.31)	(12.41)	(9.40)	(1.31)	(1.80)
Past Innovation			0.085***	0.084***	0.196***
			(30.08)	(28.71)	(29.52)
Early Stage				-0.003**	0.024***
				(-2.12)	(5.94)
Company Geographical Cluster				0.012***	0.033***
				(6.36)	(7.79)
AUM_Round				0.001	-0.002
				(0.94)	(-1.48)
Corporate VC				0.004	0.013**
				(1.55)	(2.06)
FI VC				0.002	-0.003
				(1.10)	(-0.56)
Syndicated				0.006***	0.020***
				(4.04)	(5.77)
IPO Rate				0.003	0.006
				(0.66)	(0.59)
Centrality				0.020***	0.076***
				(2.78)	(4.40)
VC Geographical Cluster				0.001	0.003
				(0.77)	(0.69)
Experience				0.000	0.003
				(0.17)	(1.11)
Early Stage Focus				0.003	0.020
				(0.52)	(1.63)
Industry Focus				0.013***	0.022*
				(2.58)	(1.80)
# Observations	58,126	58,126	58,126	53,000	53,000
Year FE	No	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Adjusted R-2	0.005	0.087	0.206	0.209	0.176

Table 10: Innovation and VC Communities - Financing Stage. The table reports OLS estimates for innovation activity between 1985-2006. The table reports OLS estimates for innovation activity between 1985-2006. The dependent variable in specifications (1)-(3) is the logarithm of (one plus) the number of patents that portfolio company applied for in the financing year. In specifications (4)-(6), the dependent variable is the logarithm of (one plus) the per patent citation count corrected for truncation bias. Both total patents and average citations are adjusted for average innovation in the same cohort (based on application year and technology class). The key variable of interest is *Early* $\times$ *Comm*, were *Comm* is Community. Specifications (2) and (5) ((3) and (6)) are based on subsamples of observations where there are no patent applications (at least one patent application) in the 3 years prior to the financing round. The observations are at the company $\times$ round level. We include all control variables in Table 9 but do not list them for brevity. See Appendix B for a description of the independent variables. Standard errors are clustered at the portfolio company level. t-statistics are in parentheses. All specifications are overall significant at the 1% level. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	Total Patents			Avg Citations		
	All	Prior Innovations		All	Prior Innovations	
		NO	YES		NO	YES
	(1)	(2)	(3)	(4)	(5)	(6)
Community	0.002	-0.001	0.009	0.003	-0.012	0.030
	(1.03)	(-0.41)	(1.02)	(0.66)	(-1.39)	(1.64)
Early Stage	-0.004**	0.027***	-0.022**	0.016***	0.116***	0.026
	(-2.16)	(6.78)	(-2.20)	(3.48)	(7.70)	(1.14)
Early x Comm	0.001	0.016***	0.013	0.018**	0.053**	0.056*
	(0.42)	(2.77)	(1.00)	(2.40)	(2.52)	(1.67)
# Observations	53,000	16,337	10,245	53,000	16,337	10,245
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-2	0.209	0.081	0.140	0.176	0.083	0.127

Table 11: Innovation and VC Communities - Drivers. The table reports OLS estimates for innovation activity between 1985-2006. The dependent variable in specification (1) is the logarithm of (one plus) the number of patents that portfolio company applied for in the financing year. In specification (2), the dependent variable is the logarithm of (one plus) the per patent citation count corrected for truncation bias. Both total patents and average citations are adjusted for average innovation in the same cohort (based on application year and technology class). The observations are at the company $\times$ round level conditional on the round having at least one community VC. Besides the set of VC- and round-level controls used so far in our specifications, we consider community-level characteristics, averaged across all community VCs in a financing round. The community-level characteristics are mean and standard deviation of age, size, centrality, industry HHI, stage HHI and company region HHI. We include all control variables in Table 9 but do not list them for brevity. See Appendix B for a description of the independent variables. Standard errors are clustered at the portfolio company level. t-statistics are in parentheses. All specifications are overall significant at the 1% level. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	Total Patents	Avg Citations
	(1)	(2)
Community Age, Mean	0.000	0.002*
	(0.50)	(1.78)
Community Age, SD	-0.002***	-0.005**
	(-2.89)	(-2.45)
Community Size, Mean	0.000	-0.000
	(0.44)	(-1.11)
Community Size, SD	-0.000	-0.000
	(-0.64)	(-0.06)
Community Centrality, Mean	0.006	0.181**
	(0.20)	(2.06)
Community Centrality, SD	-0.027	-0.318**
	(-0.54)	(-2.40)
Industry HHI, Mean	-0.135***	-0.149**
	(-4.68)	(-2.13)
Industry HHI, SD	0.041**	0.095*
	(1.99)	(1.69)
Stage HHI, Mean	0.192***	0.447***
	(3.68)	(3.31)
Stage HHI, SD	0.018	0.096*
	(0.80)	(1.70)
Company Region HHI, Mean	0.021	0.005
	(1.16)	(0.11)
Company Region HHI, SD	-0.029	0.019
	(-1.14)	(0.31)
# Observations	25,033	25,033
Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Adjusted R-2	0.220	0.184

Table 12: Innovation and VC Communities in a Within-Portfolio Model - Financing Stage. The table reports OLS estimates for innovation activity between 1985-2006. The table reports OLS estimates for innovation activity between 1985-2006. The dependent variable in specifications (1)-(3) is the logarithm of (one plus) the number of patents that portfolio company applied for in the financing year. In specifications (4)-(6), the dependent variable is the logarithm of (one plus) the per patent citation count corrected for truncation bias. Both total patents and average citations are adjusted for total patents and average citation count, respectively, in the same cohort and technology class. The key variable of interest is *Early* $\times$ *Comm* where *Comm* is Community. Specifications (2) and (5) ((3) and (6)) are based on subsamples of observations where there are no patent applications (at least one patent application) in the 3 years prior to the financing round. The observations are at the company $\times$ round level. It includes those companies which did not have a community-based VC in the first round but have at least two consecutive financing rounds of community-based VCs immediately following the first round. We exclude all remaining financing rounds if there is a subsequent round without a community-based VC. This allows us to analyze the effect on innovation of a community-based VC joining a company which had never had a community-based VC. We include all control variables in Table 9 but do not list them for brevity. See Appendix B for a description of the independent variables. Standard errors are clustered at the portfolio company level. t-statistics are in parentheses. All specifications are overall significant at the 1% level. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	Total Patents			Avg Citations		
	All	Prior Innovations		All	Prior Innovations	
		NO	YES		NO	YES
	(1)	(2)	(3)	(4)	(5)	(6)
Community	-0.000	0.023***	0.004	0.010	0.094***	-0.016
	(-0.01)	(3.08)	(0.21)	(0.78)	(3.53)	(-0.38)
Early Stage	-0.017***	-0.055***	0.007	-0.039***	-0.190***	0.038
	(-3.61)	(-6.57)	(0.35)	(-3.02)	(-6.46)	(0.75)
Early x Comm	0.019***	0.055***	-0.006	0.059***	0.170***	0.004
	(2.74)	(4.96)	(-0.20)	(3.18)	(4.40)	(0.06)
# Observations	9,957	3,509	2,206	9,957	3,509	2,206
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Company FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-2	0.505	0.395	0.528	0.404	0.479	0.401

Table 13: 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 (IPO or merger) or April 10, 2013. Specification (2) reports the estimates of a probit model in which the dependent variable is 1.0 if there is an exit (IPO or merger) within 10 years of the investment round and 0 otherwise. Specifications (3)-(5) report estimates of a competing hazards model where the event of interest is exit only through an IPO (Specification (3)), IPO or follow on financing after round 1 (Specification (4)) or after round 2 (Specification (5)). A merger is the competing risk in the competing hazards models. The sample comprises VC deals obtained from Venture Economics but excludes non-US investments, angel investors and VC firms focusing on buyouts. All specifications include year and industry fixed effects, as well as control variables from Table 9 which are not reported for brevity. See Appendix B for a description of the independent variables. 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	Competing Hazards		
			IPO	Round 1	Round 2
	(1)	(2)	(3)	(4)	(5)
Community	1.130***	0.091***	1.184***	1.065*	0.961
	(6.13)	(5.94)	(5.06)	(1.95)	(-1.08)
# Observations	67,081	67,081	67,081	19,104	14,267
Controls	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Pseudo R-2	0.013	0.103			