

Emergent Properties of a New Financial Market: American Venture Capital Syndication, 1960–2005

Bruce Kogut, Pietro Urso

INSEAD, Boulevard de Constance, Fontainebleau 77305, France {bruce.kogut@insead.edu, pietro.urso@alumni.insead.edu}

Gordon Walker

Cox School of Business, Southern Methodist University, Dallas, Texas 75275, gwalker@mail.cox.smu.edu

The recorded transactions of venture capital investments permit a direct examination of the Braudel hypothesis that regional markets evolve dynamically and interdependently in reference to a global system. This hypothesis contradicts the popular belief that regional financial development is anchored in dense clusters. Using methods of complex graphs, we analyze 159,561 transactions over nearly 45 years to demonstrate the rapid emergence of a national network of syndications. A giant component emerges early in the history of the industry, which subsumes the regional and sectoral subgraphs. The results confirm the Braudel hypothesis over the role of regional clusters, rejects preferential attachment in favor of repeated ties among trusted partners, and emphasizes the importance of dynamics and complex weighted graphs for the analysis of social and economic behavior.

Key words: venture capital syndication; complex weighted graphs; giant components; financial market integration; network graphs

History: Accepted by Brian Uzzi and Luis Amaral, special issue editors; received April 5, 2005. This paper was with the authors 10½ months for 1 revision.

The venture capital (VC) market in the United States is a young industry, dating back to a few deals shortly after the close of World War II. Most evidence points to an early history, whereby investors poorly understood the risks attached to this new market (Gompers and Lerner 2001). During the 1960s and 1970s—for which better data exist—the deals still were relatively sparse. By the 1980s, an institutional infrastructure had been put into place that permitted a rapid growth in transactions, in volume and in value. By the start of 2000, over 103,000 transactions had been recorded, and a history of dozens of firms that grew from start-up to the industrial 500 had been written.

This infrastructure is what is called a *market*. It consists of a chain of investors and brokers who often do not know the start and the end of the chain, but who surely know their adjacent links. From this perspective, capital investment passes through social networks, in which the local character of investor and investee is preserved in the context of a global myriad of intersecting pathways. The key artifact that binds the links in this chain is the transaction, which in formal markets is recorded in contracts. The records of these contractual ties among investors who enter, transact, and exit over time permit a sophisticated analysis of the VC market as an emergent network.

Looking at the emergent network of contractual ties turns the question of financial market integration on its head. The question of integration assumes that local financial markets predate national ones; integration is the joining of existing local markets into larger national ones. We propose a contrary hypothesis, borrowed from the studies of Braudel (1981) on the emergence of markets in medieval Europe, that localities almost from the start are not only connected by bridges through commercial and social ties, but that their own evolution is deeply tied to the dynamics of a wider integrating market. In fact, it is not possible to understand local financial market developments in many important cases without understanding these interregional ties and connections.

The evolution of VC in the United States is an odd context in which to make this claim because past studies have found strong evidence for the local nature of VC lending. In their early article, Florida and Kenney (1988) noted the concentration of VC in a few regional centers. Castilla et al. (2000) trace the dense network linkages among VC firms in Silicon Valley, indicating strong local social ties. Sorenson and Stuart (2001) analyzed the proximity of the VC firm to the target firm and found that investments tended to be within a 500-mile radius. Critical to our analysis below, they

showed in their cross-sectional analysis that highly central VC funds bridged these local agglomerations.

By definition, these cross-sectional results do not capture the dynamics by which regions and national networks coevolved. As we show, national brokerage and regional development grew jointly and interdependently. The adaptation of the limited partnership agreement to VC investments in the late 1950s was an institutional innovation that suddenly permitted a national market to emerge in two decades. This innovation provided the missing complement for the effective governance of investments in unknown technologies owned by new firms. The explosion in entrepreneurial investments was geographically diverse and indicates a rapid integration of a national financial market in high-risk and new companies.

Our analysis proceeds as follows. After defining terminology, data, and theoretical concepts, we show first that VC deals quickly generated an integrated national market across regions. This central finding begs the question of what microbehaviors could plausibly generate a national market so early in its history. A benchmark hypothesis is that new and incumbent VC firms are drawn to highly connected firms. If this is true, we would expect to find power-law distributions for the number of deals by the VC firm. However, this distribution gives an unsatisfactory fit. The social structure of the network, while poorly explained by the degree distribution, is revealed through an analysis of repeated ties that are impressively distributed by a power law. In these repeated ties, we see a marked preference for VC firms to renew their ties with each other by investing jointly in a sequence of investment opportunities. By partitioning a measure of local density (i.e., clustering) by geography and sector, we show that repeated ties are national, and not local. Hence, the national VC market consists of strong ties in two senses: They are repeated, and they are “one step away.” The conservative preference for repeated coinvestments to remain geographically and sectorally local is countered by the opportunities to diversify into new sectors and regions that necessitate new partners. A simple logit model confirms the negative effect of diversification on repeated ties. We thus see a trade-off between trusted expertise and diversity.

The complexity of the sciences challenges us to characterize observations of the global structure of aggregate behavior (e.g., giant components, degree distributions) that are consistent with observations of the motives and strategies of microactors. From simple rules and strategies come surprisingly well-ordered structures. The statistical analysis of these structures is considerably indebted to Newman et al. (2001) for the analytical results for a bipartite graph; to Newman (2005), Mitzenmacher (2004), Gabaix

(1999), and Farmer and Geanakoplos (2004) for their explanation of power-law dynamics; to Guimera et al. (2005) for their treatment of the entrant and incumbent link probabilities in relation to percolation; and, finally, to Barrat et al. (2004a, b) for their analysis of weighted graphs. These works permit a rich analysis into the dynamics of VC market formation.

1. Financial Market Development and the Venture Capital Industry

1.1. Historical Background: The Braudel Hypothesis

That social foundations are required to support markets for capital investment is not a new observation and is hardly unique to VC. The history of the development of capital markets after the Great Restoration in England (Carruthers 1999), of credit in post-revolutionary France (Hoffman et al. 2000), of banking in 1800 New England (Lamoreaux 1994), and of stock market trading (Michie 1999, Baker 1984) are recapitulations of patterns that Braudel (1981) saw in the local nature of foreign commerce in late medieval Europe. These histories pose the following question: At what point did these local markets intersect, such that pricing and trade could be determined to be national or international—that is, could no longer be said to be local? The Braudel hypothesis is that local markets developed in response to an evolving global system.

The conventional treatment of financial integration has been to study the narrowing of price variations across local markets. Lance Davis (1965) launched a series of inquiries by studying the process of integration of capital markets in the United States, using implied interest data on both commercial and mortgage bank loans. He found slow convergence over the period of study (1865–1914), with persisting regional differences for long-term rates in regard to the southern states. Part of the reason for this convergence, he noted, was the geographic expansion of life insurance and brokerage companies. Buchinsky and Polak (1993) find that English credit and real estate markets reveal evidence of integration by the Napoleonic Wars, perhaps in consequence of the large government demand for loans; they do not, however, cite by what means such integration occurred. Absent public information on prices and transactions, the integration of prices that is suggested by these studies implies informal mechanisms of communication. However, these mechanisms are not specified.

Evidence on direct contacts and parties to transactions provides greater insight into the social foundations of markets. The benchmark expectation is that social ties are regional, hence, so are capital markets. Lamoreaux's history of bank lending in England reveals the lending of capital even among bank

directors; Padgett's and Ansell's (1993) detailed history of the Medici shows that commercial ties lie along local family and patronal lines. Even in financial markets, we find a strong bias for proximity to matter to portfolio holdings, as found by Braggion (2005) for late 19th century England, and by Franks et al. (2005) for their study of 20th century British finance. Goetzmann et al. (2004) found that the proximity bias in Swedish portfolios in the 1990s was related to urban specialization. This role of expertise is also found in our VC data, as discussed below.

However, there are also histories that indicate social ties often bridged spatially noncontiguous transactions, such as those among Maghribi traders (Greif 1993) or among notaries in France (Hoffman et al. 2004). Neal's (1990) superb study of the integration of the London and Amsterdam markets in the 1700s and 1800s notes in detail the movement of people from the continent to England, yet his analysis relies on price data to show convergence. In his study of the London stock market, Michie (1999, p. 125) writes that "by the early twentieth century, a complex web of contacts, clients, agents, and organizers linked the London broker to investors abroad." It is this web, which is captured by the data on VC transactions, that is amenable to analysis as a graph.

1.2. Venture Capital Syndication as a Graph: Definitions

The origin of the VC market is conventionally attributed to the founding of the closed-end fund ARD by George Doriot, a French entrepreneur on the faculty of the Harvard Business School, although there were important precedents (Hsu and Kenney 2005). This fund was imitated by several others in the 1950s—primarily in the northeast United States and in California—but disappointed small investors who were unaccustomed to substantial volatility. The institutional solution was to replace the closed-end fund by private funds that accepted investments only from financially qualified investors. Critical to the legal formation of syndications, the first VC limited partnership founding dates from 1958. The Draper, Gaither, and Anderson fund, founded in Palo Alto, included Rockefeller seed money, and thus was financed by East Coast money from the start (Avnimelech et al. 2004). In this period, the U.S. government promoted VC through small business investment corporations (SBICs), although this organizational form experienced substantial difficulty (Gompers and Lerner 2001). The decision in 1978 to permit pension funds to invest a restricted percentage of their assets in VC provided a substantial inflow of money and also legitimated the market for other investors.

The dominant business model of a VC company (or firm) is the attraction of capital from private or institutional investors for investment in VC funds under

the company's management that then invest in a portfolio of target companies. For a fund life of about five years to a maximum of roughly 10 years, the managing VC firm is likely to make deals in the same sector or geography because of the nature of the fund. In many cases, VC firms will join with each other in investing in a single start-up. The combination of investments by multiple VC firms in one start-up is called a *syndicated deal*, or *syndication*. Many VC firms involved in syndications invest with numerous VC partners across a range of start-ups, creating a chain of investments in which the syndications form the links. Syndications therefore define a network among the VCs, where the ties among them are their coinvestments.

The nodes in such a network or graph are VC firms. (We treat VC *funds* as "veils" and hence, we ignore them—as has been traditional in this literature.) However, because VC firms are indirectly linked to each other through joint investments in common targets, the graph structure is *bipartite*. By projecting the bipartite graph onto the VC firms, we can create a one-mode representation of the network, in which VCs are connected if they have invested in the same target company. In this case, information about the target companies can be used to characterize the VC nodes, but the targets themselves are not part of the network graph. The adjacency matrix is symmetric, consisting of nondirectional links between the nodes.

An important dimension of a VC transaction is its order in the sequence of financing of the target company. The first VC investment in a start-up is called the *first round*. It commonly consists of *general partners*, who bear more risk and earn more return, and *limited partners*. Investments can be made in subsequent rounds, at different dates. These later rounds can include the same investors or attract new ones. A VC firm can be a *leading* or *following* partner in the target company if it was, respectively, the first or the following investor. Hereafter, we use the term *follower* to designate VC firms investing in a given target company at a time that is subsequent to the seed investment made by another firm. We also distinguish between new and incumbent firms. VC firms that participate in a deal but had not previously invested in any previous deal are called *entrants*, while firms that have already made investments are called *incumbents*.

A few graph concepts are important for the following analysis. The *degree* is the sum of the links that a VC firm has built with other VC firms through syndicated deals. However, because it is sometimes useful to refer to the bipartite graph, we distinguish between the degree of the VC firm and the degree of the target. We use the term of *board size* of a target company to indicate the number of VC firms that have invested. The *syndication size* is the number of firms

participating in the same funding round (i.e., the same target at the same time).

In addition to degree, we also look at *weighted graphs*, in which the links between two VC firms are weighted by the sum of all the deals they have done together. These weights are the count of the repeated ties among two VC firms. Whereas weights refer to links, the *strength* of a VC is equal to the sum of all the weights, that is, to the total number of investments that the VC firm has made.

Path length is the geodesic between two nodes; the average path length for the graph is simply the average of these lengths. Nodes that are connected to each other constitute a subgraph, or *component*; a VC firm that does not syndicate is a component of size one. The *giant component* is the largest component in the network.

Finally, we also analyze *clustering*. The direct connections of a given node i define its *neighborhood*. If any two nodes j and h in this neighborhood are connected, a triangle is formed, because by definition they are connected to our node i . This triangle provides closure because the three nodes form a closed loop. The clustering coefficient is defined for each node and is simply the number of observed triangles divided by the number of all potential triangles in a neighborhood, then normalized to lie in the range of $(0, 1)$. The clustering measure for the graph is the average of these coefficients. We will define clustering for weighted graphs below.

The statistical distributions of graph-theoretic measures are not well understood. The tradition is to compare the empirical observations against those estimated for random graphs of known or arbitrary distributions. Similarly, a comparison can be made among empirical observations taken at different time periods. We use both approaches in the analysis below.

1.3. Making Deals: Trusted Expertise and Diversification

Coinvestments by VC firms in a target firm have two separate components. The first is the straightforward motivation to seek other investors, and the second is the search for a particular type of VC firm as a coinvestor. A conventional analysis of the first component involves the important feature of risk diversification, whereby the lead investor seeks to spread its capital among many targets by sharing their equity with other investors. The second component derives from the uncertainty over valuations that encourages firms to signal and build a reputation by syndicating with high-prestige partners.

A principal finding in this area (Lerner 1994) is that smaller VC firms tend to join syndicates in later rounds, while larger and more prestigious firms dominate the early rounds. Along these lines, Piskorski

(2004) distinguishes between power and prestige motives for ties, finding that VC firms are sensitive to the trade-offs between engaging in deals and building a reputation by venturing with prestigious partners. Hsu (2004) extends the importance of reputation to target firms by showing that their entrepreneurs forgo higher offers in deference to the value they place on investments by high-reputation VC firms. A common measure of prestige is *centrality*. In the analysis below, we use the degree count as a measure, as we wish to analyze the well-studied property of “preferential attachment,” by which a VC prefers to syndicate with other VCs that have already made a lot of deals.

An alternative view of signaling is that syndication is motivated by the need of expertise from other VCs who then sit on the boards of the target companies (Brander et al. 2002). This explanation implies that VCs seek partnerships to bring complementary bodies of expertise to new entrepreneurial ventures. This expertise may be geographical if one VC firm has more experience in a region, or expertise can be sectoral (industry specialization). As companies sample each other’s expertise, they develop trust and enter into new syndications, thus generating repeated ties. We should find that VC firms form ties with new partners only when the investment is in new regions or geographies. This corollary is tested below.

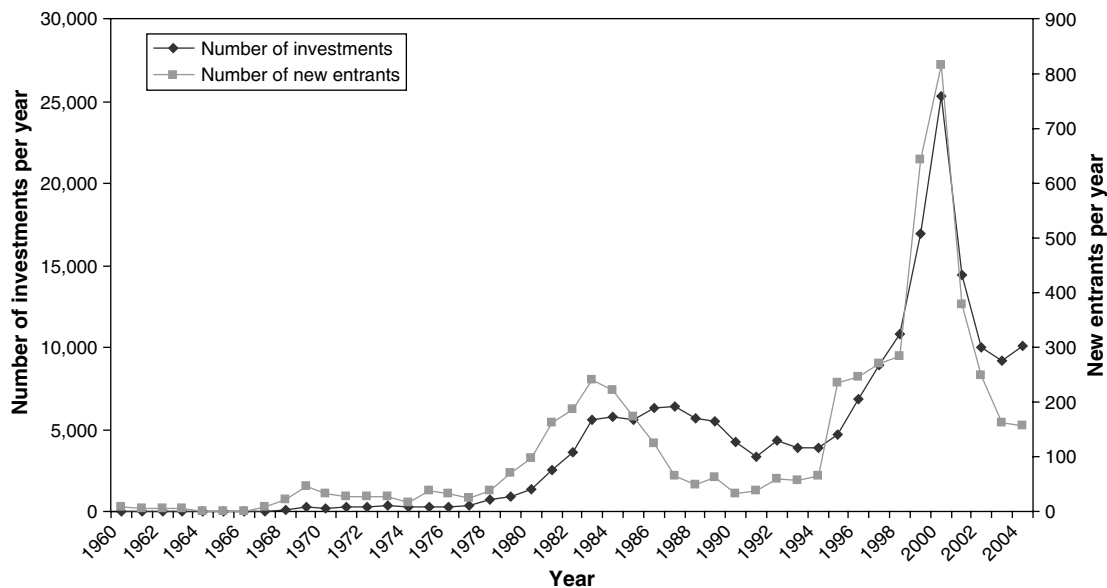
In our approach, VC firms inherit the geography and sector attributes of the target companies in which they have invested. The Sorenson and Stuart (2001) study assigned VC firm geography to the headquarters location (although for missing values; they assigned the location of the first investment). As we define diversification in terms of the target investment, we update a vector of locational attributes of VC firms by the geography of their investments. Similarly, we update the sectoral attributes of VC firms to reflect their investments in target industries. The paths of sectoral and geographical updating play a central role in understanding the local and dynamic properties of the VC network.

In addition to geographic and sectoral diversification, the second important element in network dynamics is the role played by new entrants. A surprising feature of this history is that despite high rates of entry in a finite sample of firms, statistical phenomena such as power-law distributions and percolation arise exceedingly quickly. These new entrants join the network at relatively high frequency, fragmenting the network initially, but quickly attaching themselves to established firms.

1.4. Deals Database

The Venture Economics database of the Securities Data Corporation (SDC) records VC transactions. We constructed a deals file containing the list of investments of VC firms in target companies since 1960;

Figure 1 Number of Investments and of Known Venture Capital Firms over Time



this year is only two years later than the creation of the first VC fund (Gompers and Lerner 2001). After removing errors and incomplete records, we were left with a data set of 159,561 VC deals made by 5,486 firms in 25,714 target companies located in 53 geographic areas and 10 industry sectors. The data set also includes about 1,000 deals in target companies where the VC firm names are not disclosed, although the targets' names are known. The SDC data have been criticized for spottiness prior to 1979 (Piskorski 2004).¹ For a more recent period, Kaplan et al. (2002) compare the SDC data against a small sample of hand-collected data; they find that SDC excludes about 15% of the rounds. This error would thus be rather small for the degree count (given the high number of repeated ties discussed later), but would understate the arrival time of the giant component and the number of repeated ties; on balance, these errors provide a conservative bias given our hypothesis that a national component accompanies local development.

Figure 1 provides the growth of deals over time and the entry of new VC firms (nodes). The Internet bubble clearly marks the final years of the data set. As 1979 is the year following a Department of Labor decision to allow for employee pension savings—ERISA—to be invested in VC funds, the growth of the network has a clear inflection at this juncture. Moreover, as we will show, these early years already show a pattern that is replicated in subsequent periods.

¹ We are grateful to a reviewer who noted that the SDC data start in 1960, as they were derived from the mandatory evaluations of the government-subsidized SBICs.

2. Are the Early Graph Dynamics Local or National?

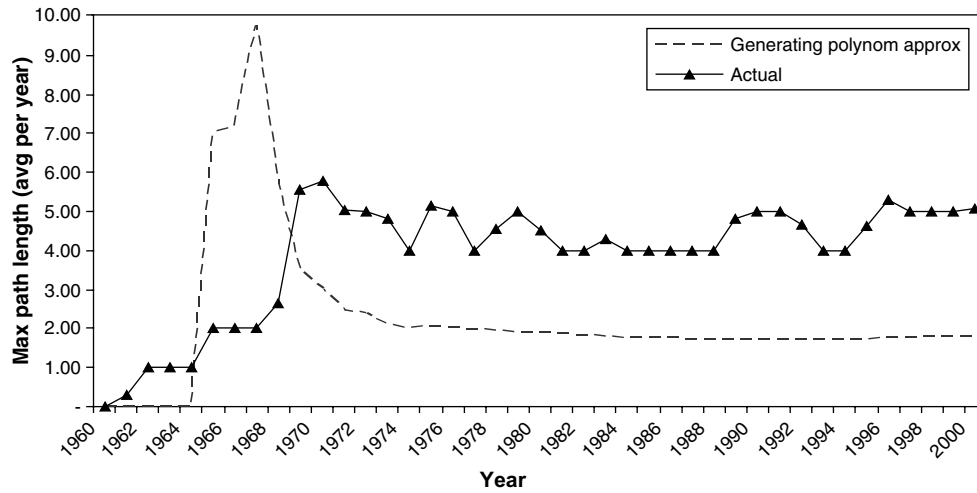
The Braudel hypothesis is that local markets develop in reference to the dynamics of a global market. To address this hypothesis in the context of the United States VC syndication network, we take snapshots of the network at discrete cross-sections and then trace the evolution of key statistics for understanding the global connectivity of the graph. We begin with the statistics on path length and then turn to the emergence of a giant component.

2.1. Length and Connected Graph Components

Path length between a given node and all other nodes provides an assessment of the *global* connectivity of a node or, when summed over all nodes and divided by the number of nodes, of the entire graph. This statistic has been used extensively in recent years due to the path-breaking investigations of Watts and Strogatz (1998) on the properties of small worlds. Graphs with short path lengths are highly connected. In the following plots, we assume that two isolated nodes have a null distance. This assumption allows us to treat unconnected graphs.

To provide a sense of magnitude, we compare our empirical estimation against the expected path length drawn from a graph of an arbitrary distribution. Using a generating function formula proposed by Newman et al. (2001), Figure 2 compares this approximated value for the path length with the actual value for the network realized (cumulative deals from 1960 to 2005). This plot indicates a substantial difference between the actual network structure and the results from the Newman et al. formula. The empirical path

Figure 2 Actual Maximum Path Length Against the Newman et al. (2001) Approximation



length is substantially larger than the formula, suggesting that simplifying assumptions do not hold for our empirical data set.

In any study on the emergent properties of a network, an entity of primary importance is the formation of the giant component. A giant component is the largest subgraph or component in the network. A visual inspection of Figure 3 suggests that the network underwent a phase transition in early 1970s in which many disconnected nodes and components combined to form one giant component. We see that the stabilization of path length is more or less coincident with the emergence of this giant component. Figure 3 indicates that the national giant component consists of all regions and sectors also in our database at this time. The rapid emergence of a giant component despite longer path lengths than predicted by

a random model point to a role by actors to bridge the regional and sectoral components in an otherwise disconnected graph. We study this role next.

2.2. The Relationship Between the Local and National Components

The results discussed above point to the obvious conclusion that the regions or sectors can be more fragmented than the national graph. The number of components in the national graph itself represents a lower bound on the sum of the components from the disaggregated data; by construction, the national giant component gives the upper bound on the maximum size of the local components. These boundary relationships provide a potentially important insight into the communication among the regions and the national network.

Figure 3 Emergence of the Largest Network Component—the “Giant Component”

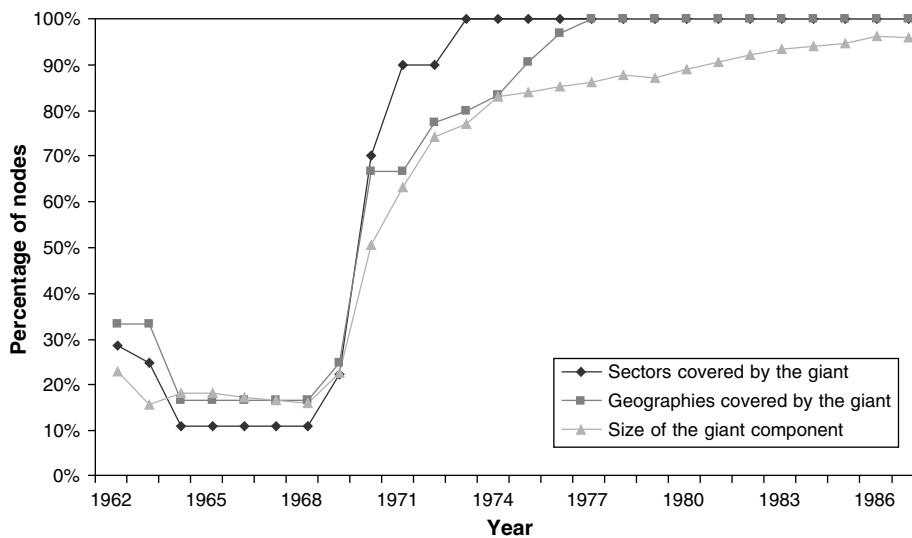


Table 1 Statistics of the Components Across Geographies and Sectors from 1980 to 2005

Year	Geographies			Sectors		
	Number of components	Size of giant component	Number of nodes	Number of components	Size of giant component	Number of nodes
1980						
Min.	1	1	1	2	1	2
Max.	10	149	158	9	168	169
Average	2.6	19.6	21.3	4.6	75.9	79.6
1990						
Min.	1	2	3	2	93	94
Max.	17	1,069	1,085	16	878	886
Average	2.5	130.6	132.2	8.4	584.8	592.4
2000						
Min.	1	2	3	7	812	818
Max.	38	2,131	2,169	35	1,537	1,566
Average	4.6	250.1	253.8	17.7	1,083.2	1,100.2
2005						
Min.	1	2	3	8	1,012	1,035
Max.	50	3,201	3,252	39	2,599	2,635
Average	5.7	379.6	384.6	22.8	1,627.6	1,605.0

Note. These values refer to the whole deals data set, excluding cases where geography or sector is unknown.

We take the whole network resulting from all deals and look at the VCs that have invested in a particular geography or sector. The deals made in other areas effectively increase the connectivity among firms, reducing the number of components and increasing the size of the giant. Table 1 shows the census of the distribution of components by region and sector. For instance, in year 2005, the geographic units varied from one to 50 components, with an average of 5.7. The sectors have a more compact distribution, varying from eight to 39 components, with an average value of 22.8. The giant component in the sectors is on average bigger than in the geographies (1,627 versus 379) even if the biggest components are in a single geography (e.g., the giant component of 3,200 nodes in California).

Another way to analyze the relationship of the national and local components is to observe the change in the number of components and the size of the giant change at the national level. Table 2 shows that, in year 2005, the giant component size increases from a maximum value of 3,201 nodes for the regions to 5,192 for the national giant component. Symmetrically, from the industrial sector perspective, we observe that the consolidation effect by national integration is even

Table 2 Statistics of the Components for the Entire Graph

Year	Number of components	Size of giant component	Total nodes number
1980	45	384	431
1990	62	1,744	1,811
2000	185	3,533	3,726
2005	278	5,192	5,486

more manifest because in 2005, the largest sector component consists of only 2,599 nodes. Because the giant components by the largest regions or by the largest sectors are not close to the national size, there is obviously a syndication network that not only spans region and sector, but also bridges components *within* sectors and regions. We can thus conclude that the structural dynamics in regions and sectors are part of a larger national dynamic, consistent with the Braudel hypothesis.

3. Analysis of Venture Capital as Nodes

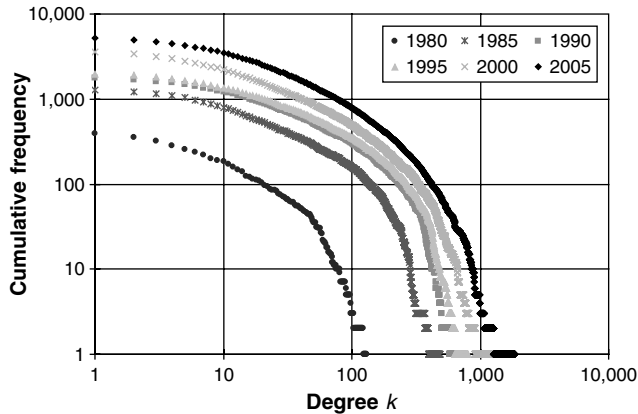
We can repeat our analysis of the network at the level of nodes—that is, the VC firms—by asking if the social mechanisms that generate a connected national market are themselves local or national. A measure of particular theoretical interest is the *degree* because reputation and signaling should lead to a pattern of “the rich get richer” in syndications, thus generating a highly skewed degree distribution characterized by a power law. We thus begin by fitting the distribution of degrees to a power law, and then see if the social dynamics are local or national.

Power-law distributions can arise for many reasons (even by a random process). They are consistent with microbehaviors that are of interest here, such as preferential attachment (nodes of high degree attract proportionally more new links) that is suggestive of the Lerner (1994) hypothesis (smaller firms join larger ones in subsequent syndication rounds). The failure to find a well-behaved power-law distribution contradicts the supposition of particular micromotives—such as preferential attachment—and thus their estimations can provide very powerful implications.

In a random graph, the theoretical degree distribution follows a Bernoulli law, approximated by a discrete Poisson distribution or, asymptotically, a normal distribution. As many have shown, the empirical distributions of graphs usually deviate from the theoretical. Barabási and Albert (1999) have found that many empirical distributions follow a power law. More importantly, power-law distributions are suggestive of a potential self-organizing criticality of a network (Newman 2005, Mitzenmacher 2004, Albert and Barabási 2002). They can be a product of competing exponentials: (1) the growth of new nodes in a network, and (2) the probability of these nodes to link (or syndicate in our data). If this probability to attach is sufficiently high relative to growth, then the network will experience a phase transition in which a giant component will emerge. At this point, the degree distribution will be power-law distributed.

As the node’s degree (k) is spotty for large degree counts, a direct fit of the degree distribution is poor

Figure 4 Cumulative Node Degree Distribution for Whole Venture Capital Network in Six Representative Years



and highly skewed. Therefore, cumulative frequency curves are commonly used in the literature. We follow Newman (2005) in plotting a cumulative function instead of the density for a sparse set of degree values. This method does not require binning of the samples, and the cumulative distribution is defined for all values of k by the expression

$$F(k) = \sum_{i>k}^{\max\{k\}} f(i).$$

The cumulative frequency is a Pareto distribution, as the frequency declines relative to incremental increases in k . In Figure 4, we show that the cumulative distribution of degree is not a straight line in a log-log scale. Thus, neither preferential attachment, as implied by a social mechanism of signaling, nor any other micromechanism is sufficiently strong in these data to generate a strong power-law distribution in its cumulative distribution form.

3.1. Fitting by a Power Law

By now, many studies have estimated the slope of the fit of a power law to the degree distribution, and its numerical value has a substantive and theoretical interest. Given that we have more than 150,000 deals and 5,000 nodes, the degree values could be very skewed. However, by inference, VC firms are limited in their capacity to handle a very large number of ties. In fact, Newman (2001) and Amaral et al. (2000) have all found power-law distribution of social networks that were better fitted by estimating a cut-off value. We therefore start with a power law, and then consider a fit with an exponential decay.

We use the distribution $f(k) = \beta k^\alpha$, where k is the network node degree, α is the parameter to be estimated, and β is a scale factor. Figure 5 shows the yearly evolution of the exponent value. Note that power law regressions of disconnected graphs are biased by the presence of isolated nodes that have made deals without syndication (these nodes have

$k = 0$), and some missing values in the discrete set of degree values. We do not consider the isolated nodes, as is the convention in the literature. These figures reveal a startling structure: By 1980, even though fewer than 3,800 out of 159,000 deals had been consummated, the alpha and standard errors approach a relatively steady state. (Not surprisingly, the intercept increases dramatically with the total number of deals.)

3.2. Deviation from Simple Power-Law Distributions

A standard power law with only two parameters is convenient to estimate and leads to good accuracy. However, the values of the exponent α differ from the results obtained, for example, by Newman (2005) or Barabási and Albert (1999) for other data sets.² Besides the problems of having a discrete set of degree values and the difficulties of using bins to plot frequencies, estimates for other common networks generally lead to $|\alpha| > 1$. Low values of α are also theoretically problematic, because such distributions do not converge in the limit.

It is also important to consider other distributions that may fit the data better. A distribution close to a power law is a log normal; we tested this distribution and found that it was an inferior fit. The rejection of the log normal is not surprising because the visual inspection of the data in the early figure reveals an extreme skewness.

Alternatively, a corrected power-law distribution might provide stronger evidence for preferential attachment. In Table 3, we show the estimates using a corrective exponential factor that multiplies the power law. This allows a degree distribution that converges and can be associated with a probability density even if $|\alpha| < 1$. The best fit of the cumulative distribution is obtained by a least square error estimation against the function³

$$F(k) = e^{\lambda k} \beta k^\alpha.$$

Newman found that this expression can be normalized for any α by a constant $C = [\beta Li_{-\alpha}(e^\lambda)]^{-1}$, where we use the polylogarithm function

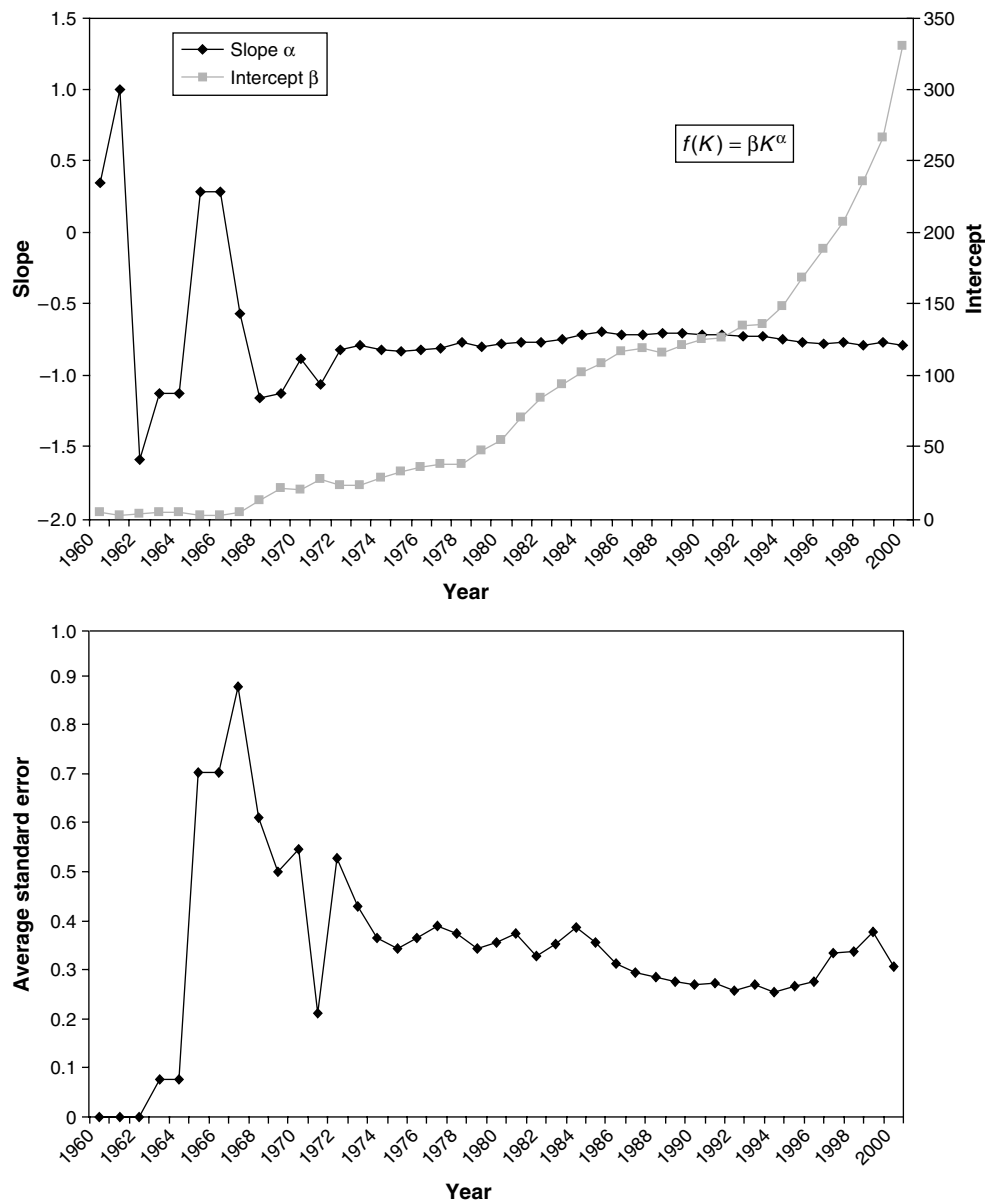
$$Li_\alpha(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^\alpha} \quad \forall |z| < 1.$$

However, the exponential correction does not improve the fit for high degree values. The tail of the distribution is affected by measurement noise that is difficult to model with simple functions. Other models also performed poorly.

² However, many important phenomena are characterized by power-law distributions with low exponent values, such as scaling in biological systems (see West et al. 2000).

³ This is equivalent to fitting the density distribution by $f(k) = \beta e^{\lambda k} k^{\alpha-1} (\lambda k + \alpha)$.

Figure 5 Power-Law Degree Distribution Estimated Coefficients and Standard Error



In summary, the power-law fit for the cumulative distribution is poor, even if it provides the best fit of estimated models. Linear models exaggerate the tail of the distribution, whereas common lognormal fitting leads to a shorter-than-actual tail extension. Because the standard approaches are ineffective for the extreme values of the distribution of the high degree values in our case, more complex models are generally accepted.⁴ The network of VC firms is also characterized by a few nodes that have a high degree

of connection and an extremely high number of nodes that have only a few ties. The low exponent value to this distribution is not consistent with a theory of preferential attachment, which would be (naively) expected for the formation of links where prestige and reputation should influence investment ties.

3.3. Network Cohorts

The above estimations used the cumulative data set from 1960 to 2005. It is possible that we are mixing distributions, and we can recover the power law if we look at cohorts. Cohort analysis standardizes age and period effects. In the absence of technological and competitive interventions, the distributions of cohorts should be ordered by age. Figure 6 shows a reversal: By 2005, the frequency of nodes belonging

⁴ Laherrère (1996), for instance, has shown that, for data drawn from urban areas extension, world population, spoken languages, galaxies intensity, and oil fields, we often observe a dominant value, or a largest object is much larger than the next in line; these are better fitted by a fractal specification.

Table 3 Estimated Parameters for a Power-Law Distribution Fitting the Cumulative Frequency Distribution, with an Exponential Cut-Off Factor

	1980	1985	1990	1995	2000	2005
Slope α	-0.22	-0.25	-0.23	-0.24	-0.29	-0.27
Intercept β	425.6	1,485.1	2,070	2,320.9	4,223.2	6,276.9
Lambda λ	-3.5E-02	-1.1E-02	-8.3E-03	-7.5E-03	-7.4E-03	-7.6E-03
$Li(-\alpha, \exp(\lambda))$	15.46	35.27	47.24	49.16	40.89	43.30
C	1.5E-04	1.9E-05	1.0E-05	8.7E-06	5.8E-06	3.7E-06

to the 1995 cohort, and with degree value between 100 and 300, is higher than in the 1990 cohort. New entrants in 1995 were able to invest in more targets and develop their network much faster than were 1990 entries. Clearly, this difference implies that VC firms are specialized in their sectoral knowledge and that the 1995 cohort entered with skills appropriate to Internet-related opportunities. The year 2000 entries, which consist of 10 times more firms than the 1995 cohort, show a lesser connectivity in the same range of degree value. This reversal is a strong indication that sectoral and technological expertise can change the dynamics in the network.

3.4. Exit of Nodes

Our final effort to recover a power-law distribution of degrees is to take into account the decay of old links among nodes. We assume that a firm has exited the network when all ties have expired. We decay ties by removing them after observing a prolonged period of inactivity. The degree α distribution including decay is fundamentally the same. (The graph is available upon request from the authors.) Exits do not change our fundamental results regarding the power-law distribution and its inconsistency with preferential attachment. Surprisingly, the degree distribution is similar even for exits of VC firms after only one year of inactivity. We obtain this result for two reasons: the high level of new entries that replace “broken” ties and the strength of edges between nodes given by repeated

syndicated deals over time. The latter reason points to the importance of repeated ties among VC partners, which is a microbehavior distinct from preferential attachment favoring large degree nodes.

3.5. Analysis of Regions and Sectors

Power-law distributions have the interesting property that they can also result from multiplicative or additive mixtures (Farmer and Geanakoplos 2004). It is possible, then, that the power-law distribution that we observe nationally is in fact the mixture of regional or sectoral power laws. This hierarchy suggests a “fractal-like” pattern in which the structure of the whole is repeated in the most popular regions and sectors, where we have a substantial number of samples. Such a pattern would suggest two possible explanations: (1) There is a similar microdynamic in each region or sector, or (2) there is a “communication” among regions that results in a common dynamic. Or, alternatively, we may find that preferential attachment is operative at the regional and sectoral level, thus contradicting the conclusions from connectivity that the national market emerged early. We address these points by analyzing the estimations from the principal regions and sectors.

Table 4 summarizes the above findings by comparing the estimates across the two dominant regions and sectors at the beginning of 1985, 2000, and 2005. (Because power-law estimates are sensitive to extreme values, they require a lot of data to attain sensible estimates; the estimates reported below are similar to those of smaller sectors and geographies.) We use the power-law distribution with an exponential limit to facilitate comparison with the earlier results. The estimates are remarkably close to each other, except for Massachusetts and Internet companies in 1985; the disparity for the Internet sector is, of course, not surprising given the very low frequency of deals at that time. By 2000, this disparity disappears.

In summary, the results for the degree distribution show slope coefficients at the sectoral and regional levels close to those for the national level, suggesting similar processes, and the global connectivity analysis indicates that these processes were surely communicative. However, the low values indicate that these

Figure 6 Cumulative Nodes Degree Distribution in 2005 for Four Different Cohorts

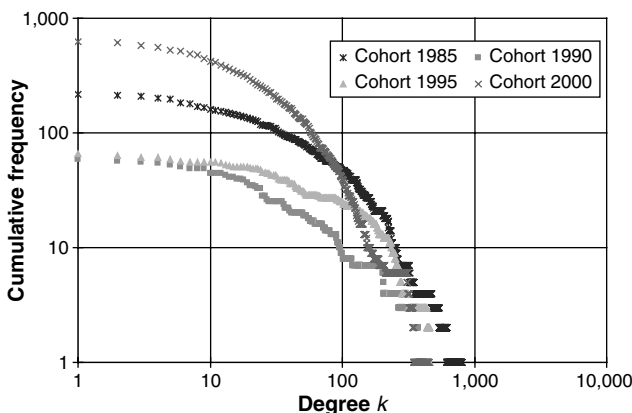


Table 4 Power-Law Estimates for Dominant Regions and Sectors

Year	Computer services			Internet-specific			California			Massachusetts		
	1985	2000	2005	1985	2000	2005	1985	2000	2005	1985	2000	2005
Nodes (%)	35.6	46.1	48.0	2.1	42.9	43.0	32.4	57.1	59.1	2.1	42.9	43.0
Slope α	-0.28	-0.28	-0.29	0.04	-0.31	-0.30	-0.27	-0.24	-0.29	0.04	-0.31	-0.30
Intercept β	537.6	2,094.1	2,634.5	33.1	1,947.5	2,359.0	1,777.1	861.5	2,695.3	33.1	1,947.5	2,359.0
Lambda λ	-1.2E-2	-7.0E-3	-8.2E-3	-1.9E-1	-3.1E-2	-2.1E-3	-5.1E-3	-1.6E-2	-1.1E-2	-1.9E-1	-3.1E-2	-2.1E-3

social influences are fairly weak. The cohort analysis points to an important role played by technological change and technical expertise, which may also be correlated with geographical changes as well. We now examine these waves of technological breaks in the industry and show how they generate a trade-off between a social mechanism called *trusted expertise* that leads to repeated ties, and the economic rationale to enter new markets that leads to partner diversification.

4. Exploring Industry Dynamics: Entrants, Incumbents, and Diversification

The analysis discussed above presents the puzzle that a striking national coherence emerged early and has persisted despite massive entry, and allowing even for exit. In this regard, the evolution of the VC industry looks distinctly different from the pattern seen in many other industries that are marked by a period of shake-out and high levels of concentration. For example, Klepper (2002) found that the number of firms that entered the automobile industry grew steadily from 1895 to 1907, peaking at 82 in 1907. Less than 30 years after the industry founding date, only 15 firms entered from 1923 through 1966. The VC industry experienced, as we know from above, massive and growing entry throughout the 40-year history observed in our data set.

Surely, part of the explanation is the far lower capital equipment expenditures required for VC, but the importance of social capital and signaling is far from negligible (Sorenson and Stuart 2001, Piskorski 2004). Yet, the analysis of the degree distribution contradicted a type of rich-get-richer (or preferential

attachment) motivation. What, then, are the microbehaviors that we observe in the data that are consistent with the emergent properties that we observe? We propose that these motivations can be understood by analyzing the dynamics of new entrants and incumbents regarding the choice to build new links with new targets, or to link to incumbents.

4.1. New Entrants vs. Incumbent Behavior

In the spirit of clearly specifying the evolutionary details of the VC market, we focus on the basic facts of entry and of link formation. In one of the earliest studies on social network dynamics, Walker et al. (1997) showed that new entrants evince different logics in forming ties than incumbents, with incumbents tending to reinforce existing ties. This distinction between incumbents and entrants is also found to have behavioral consequences in Guimera et al. (2005) and in Powell et al. (2004). In an important article, Lerner (1994) proposed a different dynamic whereby smaller VC firms tend to join late and to join bigger syndicates, that is, they link to incumbents.

Incumbents tend to have an increasing number of deals as the network grows. One might assume that this trend is true by construction because there are monotonically more incumbents over time. However, in the last decade there were an increasing number of deals led by new entrants, and the share of incumbent deals fell. By 2002, incumbent shares returned to their early-1990 levels. Clearly, technological breaks led to a change in deal making, as the expertise in Internet technologies was initially held by the new entrants.

4.2. Diversification of Incumbents

We showed earlier that a national giant component tied together the components within sectors and

Table 5 Diversification Odds for Incumbents Across Sectors and Geographies (All Deals from 1960 to 2005)

Sector	Geography by number of investments			Geography (%) by column			Geography (%) by row		
	Inbound	Outbound	Total	Inbound	Outbound	Total (%)	Inbound	Outbound	Total (%)
Inbound	129,923	10,142	140,065	95.4	56.7	90.9	92.8	7.2	100
Outbound	6,260	7,750	14,010	4.6	43.3	9.1	44.7	55.3	100
Total	136,183	17,892	154,075	100	100	100	88.4	11.6	100

Note. γ_2 : significant at $p < 0.01$.

Table 6 Diversification Across Geographies and Sectors: Odds for Incumbents Across Sectors and Geographies (All Deals from 1960 to 2005)

Follower	Incumbent by number of investments			Incumbent (%) by column			Incumbent (%) by row		
	True	False	Total	True	False	Total (%)	True	False	Total (%)
True	121,904	3,757	125,661	79.1	68.5	78.8	97.0	3.0	100
False	32,171	1,729	33,900	20.9	31.5	21.2	94.9	5.1	100
Total	154,075	5,486	159,561	100	100	100	96.6	3.4	100

Note. χ^2 : significant at $p < 0.01$.

geographies. This global connectivity reflects, tautologically, the decision to diversify syndications geographically and sectorally. These two dimensions are differently correlated; for instance, Table 5 shows that 55% of sector-diversifying deals also occur in new geographies, while only 43% of geography-diversifying deals involve new sectors. The over-time data indicate that diversification was most pronounced in the early period of the industry, stabilized, and then increased in the 1990s. During this latter period, firms sought to diversify more through new geographies than through new sectors.

An interesting comparison between diversification in sectors and geographies is given by the Herfindahl-Hirschman index of concentration. Using the sum of squared market shares to indicate the degree of concentration of deals across geographies and sectors, the calculations indicate that particular geographies were far more favored (i.e., California and Massachusetts) than other regions. By the end of the 1990s, the large investments in computer services led to a high level of concentration level averaged across sectors. The entry in the 1990s due to the Internet boom reduced the concentration levels before they returned to their historical values.

4.3. Followers or Start-Up Deals

As noted earlier, Lerner (1994) found that for biotechnology investments smaller VC firms are drawn toward investments that have already established investors in later rounds; this finding has been routinely corroborated (Piskorski 2004, Sorenson and Stuart 2001). We can investigate this claim more directly by analyzing the propensity of new and incumbent investors to invest in larger syndications. It should be recalled that investments in VC can be made at several stages. In our data, Stage 1 is the recording of the first investment by a VC firm into a new target company. If another VC firm invests at the same time, a syndication event occurs at the same time of the first investment. It is common that subsequent to the first stage, there are further stages of investment in which new investors—the followers—take a stake in the company.

Table 6 shows that the majority of deals are late rounds, and they are the preferred method by both

incumbents and new entrants. Thus, we see a fuzzier tendency than that described by Lerner (1994) because both entrants and incumbents prefer an established syndicate rather than investing alone; a chi-square test cannot reject the null of no difference between the behavior of incumbents and new entrants.⁵

It is fair to conclude, then, that entrants show a dual propensity: Their entry tends to be relatively more in start-ups and stand-alone companies, and yet they, like incumbents, expand through investments in targets that already have a large number of investors. The first investment buys the ticket to the game; the subsequent investments tend to show the prestige deference noted by Lerner (1994), Sorenson and Stuart (2001), and Piskorski (2004). Thus, ties should be repeated often. The dynamics of repeated ties, then, should be more revealing of the relationship between the structure of the network and the individual motivations of VC firms to syndicate.

5. Repeated Ties and Social Rules

The statistical description above indicates that there is a marked tendency for incumbents to form links to other incumbents. This tendency has been noted for other industries, especially by Guimera et al. (2005) for Broadway musicals. They emphasize the importance of diversity for creativity, while noting that the overall tendency is for incumbents to align themselves with incumbents. In their data, 77% of all ties are made by incumbents, while incumbents tied up with incumbents 16% of the time. In our data, 96.6% of all ties are by incumbents, and 76.4% are ties among incumbents. Clearly, the VC market is characterized very strongly by a conservative dynamic in which incumbents dominate, and favor dealing with each other.

Barrat et al. (2004a, b) described the structure of *weighted graphs* and pointed out that the weights distribution can be very different from degree distribution. A weighted graph associates a measure with every edge in the graph. For example, air traffic

⁵ Lerner (1994) distinguished between small and large VC firms; he did not compare entrants and incumbents.

among airports constitutes a weighted graph in which airports are nodes and the number of flights between two cities gives the weight to the edge that links the cities. The weights of links, and resulting node strengths, can have a major influence in the dynamics of the network formation and evolution.

We failed to capture the social dynamics earlier by fitting the nodes degree distribution because degree is a measure independent of repeated ties that characterize a weighted graph. We define w_{ij} as the number of deals (or contracts) made by two VC firms, i and j . The degree distribution ignores these weights, which in our data reflect the most important social fact in the transactions: VC companies form enduring partnerships.

The sum of the weights for any node (VC firm) is a measure of strength. More formally, the total number of repeated ties of a node can be written as

$$s_i = \sum_j \sum_t w_{ij}^t,$$

where $w_{ij}^t = 1$ if firm i and j invested in company t .

Strength of a node is likely to be correlated with its degree because large and wealthy VC firms are likely (although not necessarily) to have more partners and more deals. In fact, the log-log graph of strengths and the statistical fit to a power-law distribution are remarkably similar to our findings for the degree distribution.

More interesting is the distribution of the weights themselves. We estimate the fit by a power law to the links weight distribution according to the equation

$$f(w) = \beta w^\alpha, \quad w \in \{w_{ij}\} \forall (i, j); \quad w_{ij} = \sum_t w_{ij}^t.$$

Figure 7 clearly shows a linear distribution. The number w of repeated ties generated by deals between firms in the same or a different target company is distributed according to a classic power law, with an

Figure 7 Histogram (in Log-Log Scale) of the Number of Repeated Ties (i.e., Link Weights) in 2005

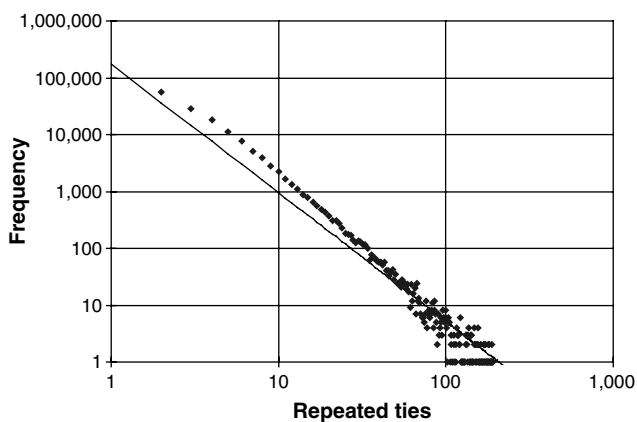


Table 7 Slope and Intercept for an Estimated Power Law on the Cumulative Repeated Ties Distribution in 2005

	California	Massachusetts	Internet-specific	Computer services
Slope α	-1.64E+00	-1.62E+00	-2.55E+00	-2.94E+00
Intercept β	7.25E+05	1.25E+05	1.10E+06	6.42E+06

Note. Samples with fewer than 10 repeated ties are excluded from the regression.

exponent of -2.26 . This fit indicates that many relationships are often reiterated, signaling a preference in doing repeated deals with a few firms. On the other hand, a power law does not fit the nodes' degree and strength distribution, which appear skewed in the tail for low-frequency values. The social mechanism of repeated ties is not the same as the preferred attachment by which anonymous firms are attracted to the high-degree firms. On the contrary, this skewness in repeated ties points to a social mechanism by which firms build relationships on the basis of trusted expertise.

Consider, for example, New Enterprise Associates (NEA), a VC firm founded in 1978 on both coasts, and specializing in medical and information technologies. With \$6 billion under management, NEA has many partners, but a few form a core group. Our data indicate that it has 1,626 repeated ties with its top five partners. Similar patterns can be found for such well-known firms as Kleiner Perkins. Because the distribution is very skewed, a few firms dominate the extreme tails.

We again compare the estimates for the national distribution against those for the main regions and sectors. Scatter plots of the repeated-ties distribution in the most common geographies and industries are characterized by a curved distribution closer to an exponential corrected model.⁶ In each geography or sector, the distribution is not power law, while it is at the national level. Therefore, geography and sector diversification leads to a power-law distribution at the national level. These repeated ties across the national graph create very strong ties and clearly indicate a deep global structure to the venture financial market.

As calculated in Table 7, we can observe an excellent fit to the power law after removing the samples for fewer than 10 repeated ties. The exponents for the sectors are substantially larger (in absolute value) than those of the regions, indicating a more pronounced tendency for repeated deals to be driven by sector expertise. (See the logit estimates below that confirm this inference.) The regional slope coefficients are also substantially smaller than those for the entire graph.

⁶ These plots are available upon request from the authors.

The above analysis describes a growing network of two competing forces: new entry fragments the network, particularly at the regional level, but the diversifying behavior of the VC firms drives global connectivity. Contrary to the conventional hypothesis of regional dominance, trusted relationships emerge at the national level. It is then a tension between diversity (diversification) on the one hand and expertise on the other that drives link formation in the VC network. One can speculate that geographical diversification is often channeled along conduits of sectoral expertise. Repeated ties are favored by incumbents who are making investments in familiar sectors and regions, and hence can rely upon their traditional partners. It is diversification that drives the search for new and diverse partners and, consequently, the formation of new links, but eventually a conservative force dominates by which exploration leads to learning to repeat deals with known firms over those untried.

5.1. When Network Clusters Are Not Geographic Clusters

For a final application of graph-theoretic concepts to analyzing local and national dynamics, we turn to an analysis of clustering. Clustering is a measure of closure: To what extent are the friends of friends also friends? Clustering is agnostic as to whether a VC firm is connected to firms in the same geography or sector. Clustering is not, then, the clusters used in regional geography to denote high density within a region. By the above analysis, we should in fact expect national clustering to dominate local clusters.

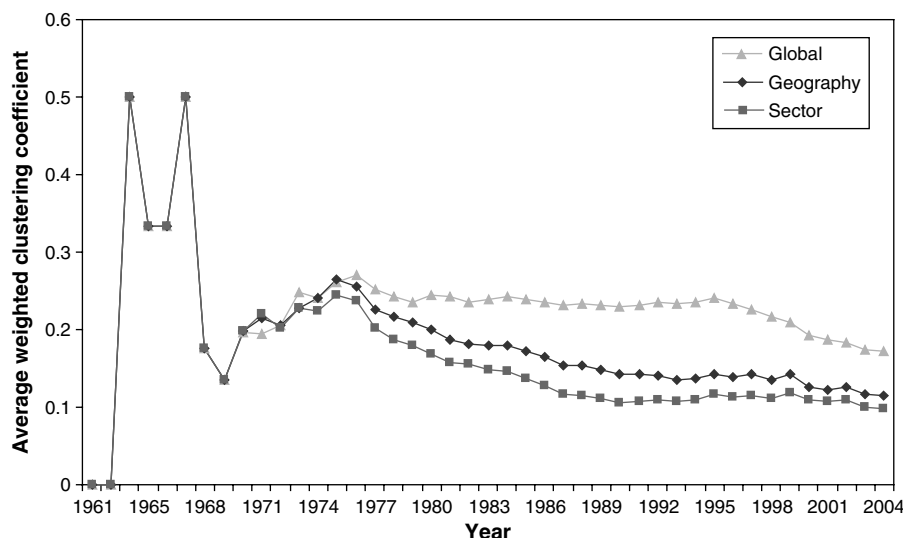
Because we have shown that repeated ties capture the social mechanism of trusted expertise in VC syndication, we do not use the standard clustering measures found in Watts and Strogatz (1998) or in Newman et al. (2001). Instead, we analyze the weighted clustering coefficient as defined by Barrat et al. (2004a, b). This measure counts the number of triangles, but weights them by an average of the focal firm's number of repeated ties with its two counterparts; this weighted count is then normalized to assure that the clustering coefficient value lies in the range of (0, 1).

The plot in Figure 8 indicates that the clustering coefficient is about constant during the period under analysis and has an average value of 0.25. This signals a higher propensity to invest for firms that are at the center of a group of other firms that have a lower connection degree than the central node.

Because VC firms are likely to diversify their investments across geographies and sectors, we are also interested in looking at the composition of the cluster around a given node. Indeed, the ties between a firm and its closest neighborhood do not necessarily represent the funding of companies in the same geographic area or industry sector. When we do not consider this localization aspect of the ties, we have a standard, and from this prospective national, definition of clustering coefficient.

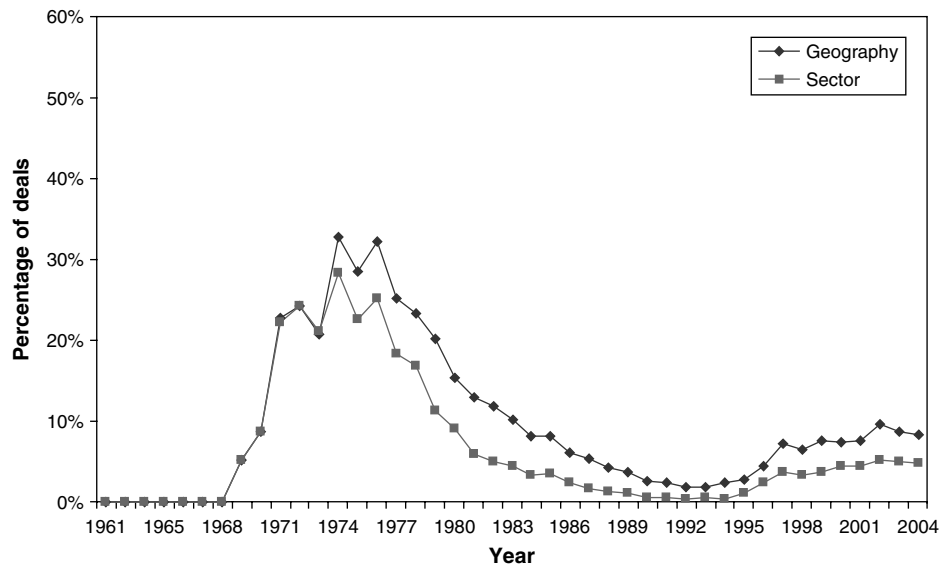
We have analyzed, for all firms, the weighted clustering coefficient calculated considering only the deals within the geography or the sector of the latest investment. Figure 8 shows that the average value of the geographic- and sectoral-clustering coefficient

Figure 8 Average Weighted Clustering Coefficient of Firms Investing in a Given Year



Notes. The global value corresponds to a clustering coefficient calculated for all node ties, while the other local coefficients account either only for ties in the geography or the sector of the deal.

Figure 9 Percentage of Deals for Which the Local Weighted Clustering Coefficient Is Higher than the Global One



is lower than the *national* weighted clustering coefficient. At the VC firm level, there are, of course, exceptions: Sometimes firms are very well connected in the geography or sector of the deal, leading to a clustering level higher than the global one. In particular, Figure 9 shows how this higher connectivity by local deals involves almost 40% of the deals in the 1970s and disappears early in the 1990s. Therefore, from a clustering perspective, we can say that the national structure took over the local subnetworks during the 1980s.

5.1.1. A Simple Test. We can test this inference more rigorously by estimating the effects of the network statistics on the choice to form a repeated tie or a new link. Link formation has been analyzed by other researchers, such as Gulati and Gargiulo (1998) and Powell et al. (2004), and most relevantly by Sorenson and Stuart (2001), Hsu (2004), and Piskorski (2004) in the case of VC. In our analysis, we wish to keep the analysis to a few essentials to estimate the effects that the above analysis has indicated are important: diversification, the degree count of the target investor, the degree of the VC firm, and new entrants. Because our interest is not in analyzing why firms invest in particular sectors and regions, we do not include economic data on their size. The analysis here focuses on whether the choice of a repeated tie (coded 0) or new tie (1) is related to diversification (whatever the true economic opportunity), favoring past partners, or new entrants.

The data consist of 159,561 transactions from 1960 to 2005. The dependent variable is whether a link represents a new edge between two nodes (VC firms) or is a repeated tie. We code the dependent variable as zero when the link is repeated and as one

when it is new.⁷ We remove 1,321 entrants because by definition they cannot form a repeated tie; we thus also lose 5,486 transactions. The diversification variables (geographical, sectoral, and their interaction) are binary (0, 1), indicating whether the link has been created by a diversifying investment—that is, when a firm invests in a company that is in a new geography or new sector as compared to its current investment portfolio. The degree count of the target is a refined measure of the investment round and also of the attractiveness of the target. The degree of the VC firm measures its stock of deals, which is a reasonable measure of its size.

In Table 8, we report the results of panel logit regressions estimating the probability of having a repeated tie formed depending on the sector and geography diversification and on whether the node is a new entrant.⁸ All coefficients are significant. We use the Huber White correction for heteroscedasticity and a panel design to account for ties formed by the same firm.

The first run establishes the baseline by estimating the effects of diversification on repeated ties. By themselves, geography and sector have positive coefficients. When their interaction is added to the analysis, the coefficients of these two variables remain positive, but the interaction is strongly negative (see the first column). The slightly larger coefficient for sector

⁷ A new tie is recorded under two cases: The first is when a VC firm invests in a new target, and the second is when a VC firm invests in a target whose previous investors include a new VC partner (that is, it does not have any previous investments with the new investing VC firm).

⁸ The results are not sensitive to reasonable alternative specifications, such as clustered logit or complementary log-log (for rare events).

Table 8 Effects of Diversification and Degree on Repeating or Making a New Tie Logit Panel Regression with Robust Errors (Coefficients and Standard Errors)

	Diversification	VC degree	Target degree	Year effects
New geography	-1.577 -0.023	-1.504 -0.023	-1.559 -0.023	-1.57 -0.024
New sector	-1.723 -0.029	-1.613 -0.029	-1.663 -0.029	-1.669 -0.029
Interaction	1.502 -0.043	1.449 -0.043	1.483 -0.043	1.499 -0.044
Firm degree		0.002 0	0.002 0	0.002 0
Target degree			-0.046 -0.002	-0.056 -0.002
Constant	1.398 -0.015	1.24 -0.018	1.487 -0.02	23.439 -26,313.25

Note. Observations: 154,075; Number of firms 4,165; New tie: 0; Repeated tie: 1.

confirms the results from the weighted-link analysis, where we found that repeated ties were stronger in bridging the graph. In all, the search for new opportunities encourages the formation of new links. The second and third regressions add in the degree of the investing VC firms and the degree of the target; target degree is equivalent to what we have called *syndicate size*. Firms with many syndicate partners form repeated ties compared to new edges; this pattern is a type of inbreeding—where friends tend to stay with friends. Also, new edges are formed when the target already has a high number of investors. This result suggests that indeed new investors in a syndicate are attracted to the size of the board and to an unobserved variable regarding target quality. When year effects are added (Model 4), the results are the same, even though the likelihood function may not be globally maximized for the year and constant effects.

We also ran a panel logit regression on whether new links were formed in a syndicate; for entrants, this coding means whether the entrant invested in an isolated target (no previous investors) or in a target with previous investors. Using the same variables given in Model 3 of Table 8, we included a dummy for whether the firm was an entrant. The positive coefficient is not consistent with the Lerner (1994) hypothesis regarding smaller VC firms; at least for new entrants, they are relatively likely to favor new ties. We would then amend the Lerner hypothesis in this way: On the condition that a new VC firm joins a syndicate, its investments are in relatively later rounds of syndication. (Note that a VC firm that invests for the first time in a new target company, which has not received other investments, does not immediately build new ties with incumbent firms.)

These results very clearly confirm the patterns we described above. Diversification promotes links to

new partners; the tendency toward exploration by new partnering is offset by a preference for investing with past partners. It is then easy to see why a giant component arises nationally, for it is by acts of sectoral and regional diversification that new edges are formed. At the same time, incumbents conservatively prefer their trusted partners. Thus, there is a tension between diversity and trusted expertise that is consistent with our power-law results. High-degree firms favor past partners, and hence the distribution is not sufficiently skewed to generate a good fit to a power-law distribution, but some VC firms have many repeated ties that are distributed by a power law to textbook precision.

6. Conclusions

We analyzed syndication contracts to explore the geographic and sectoral diversification of VC firms from the time of their founding. In contradiction to studies that conclude that regional development precedes national, our results point to the importance of understanding the dynamics of the global graph to understand regional development. The analysis of the microbehavior of VC firms shows a high degree of diversification for firms that played a historic role in creating a national market. Given the past evidence on signaling and prestige, we were surprised *not* to find high-degree firms to be relatively more inclined toward making new edges. However, at the same time, the VC firms preferred the conservative rule to invest with known partners. It was only the force of diversification that broke the lock of proven expertise of trusted partners.

In addition, the statistical analysis indicates that new edges are formed in reference to diversification, sectoral and geographic, and to targets that have already received endorsements by established firms. Degrees, as a measure of size, do not explain the formation of new ties well, for incumbents tend to tie to prior incumbent partners—a finding made by Walker et al. (1997) for biotechnology and Sorenson and Stuart (2001) for VC. It is not surprising, then, that the ranking of the largest investors by dominant region and sector is not the same for the national graph. (The tables are available upon request from the authors.)

The statistical results of this study differ in an important way from many other studies of social phenomena. The degree distribution among connected nodes deviates from the power law, even if this distribution is the best-fitting curve for the degree count. The distribution of the repeated ties reveals a power law indicating the tendency of a few firms to do many deals with each other. Because neither the degree distribution, nor the node strength distribution is a

power law, there must be, by implication, a very strong tendency of firms to repeat deals with each other. These results suggest that the microrules governing the formation of syndication differ from a mechanism of preferential attachment (Barabási and Albert 1999). The statistical results point to a few simple rules (or strategies) that trade off between the reliance on proven expertise and the search for new opportunities through diversification, often achieved by coinvesting with new partners.

The microrules by which small firms—or in our data, entrants—connect to larger entities rapidly produced a connected network, displaying a giant component and power-law distributions in the major geographies and sectors. Fleming et al. (2004) estimated that the emergence of a giant component in a graph of coauthors to a patent occurred only in the early 1990s for Silicon Valley, and several years later for the Boston Route 128 area. For VC, this result of percolation was nationally and regionally observed by the early 1980s. It is premature to conclude causality—that connectedness in VC caused a high density of patenting. We note, though, that such an implication is consistent with Kortum and Lerner's (2000) study of VC and innovation (as measured by patent counts): namely, those VC investments promote innovation. It is an interesting speculation that such investment networks may cross critical thresholds of interconnectivity to provide an externality in the form of an infrastructure of trust and cumulative local knowledge and expertise.

The American VC market entailed national brokerage from its start, with dense localities evolving in reference to its own internal economics, but also to these national ties. It is quite possible that financial centers favor those regions that are already rich in social ties. Braudel (1981) would dissent, however, arguing that international trade and finance influenced medieval material life at all levels and bridged local markets. It is this local and national connectivity that we find to be an early and remarkable characteristic of the evolution of the market for VC in the United States. No doubt, the primary difference in today's market is that the intensity of these links evolves so much faster than before. Local agglomerations are not independent of global connectivity. It is an important observation in a current world in which policies for financial and industrial regional development are often fostered in opposition to, rather than in conjunction with, the dynamics of the global market.

Acknowledgments

The authors thank INSEAD for providing the financial support for this project, as well as the project on complexity funded by the European Union, organized by Felix Tsochas-Reed of the Said Business School, Oxford University. Mariano Belinky provided much appreciated assistance and

advice. Elements of this paper have been presented at New York University, Max Planck Institute at Jena, Harvard Business School, and WHU at Vallendar. The authors thank Fabio Braggion, Jean Laurent Rosenthal, Brian Uzzi, and the reviewers for their comments.

References

- Albert, R., A.-L. Barabási. 2002. The statistical mechanics of complex networks. *Rev. Modern Phys.* **74** 47–97.
- Amaral, L., A. Scala, M. Barthelemy, H. E. Stanley. 2000. Classes of small world networks. *Proc. Natl. Acad. Sci. USA* **97** 11149–11152.
- Avnimelech, G., M. Kenney, M. Teubal. 2004. Building venture capital industries: Understanding the U.S. and Israeli experiences. BRIE Working Paper 160, University of California, Berkeley, CA.
- Baker, W. 1984. The social structure of a national securities market. *Amer. J. Sociol.* **89** 775–811.
- Barabási, A.-L., R. Albert. 1999. The emergence of scaling in random networks. *Science* **286** 509–512.
- Barrat, A., M. Barthelemy, A. Vespignani. 2004a. Modeling the evolution of weighted networks. *Physical Rev. E* **70** 066149-1–066149-12.
- Barrat, A., M. Barthelemy, R. Pastor-Satorras, A. Vespignani. 2004b. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **101** 3747–3752.
- Braggion, F. 2005. Credit market constraints and financial networks in late Victorian Britain. Mimeo, Department of Finance, Tilburg University, Tilburg, The Netherlands.
- Brander, J., R. Amit, W. Antweiler. 2002. Venture-capital syndication: Improved venture selection versus the value-added hypothesis. *J. Econom. Management Strategy* **11** 423–452.
- Braudel, F. 1981. *The Wheel of Commerce. Civilization and Capitalism: 15th–18th Century*, Vol. II. Harper & Row Publishers, New York.
- Buchinsky, M., B. Polak. 1993. The emergence of a national capital market in England, 1710–1880. *J. Econom. Hist.* **53** 1–24.
- Carruthers, B. 1999. *City of Capital: Politics and Markets in the English Financial Revolution*. Princeton University Press, Princeton, NJ.
- Castilla, E., H. Hwang, M. Granovetter. 2000. Social networks in Silicon Valley. C. M. Lee, W. F. Miller, H. Rowen, M. Hancock, eds. *The Silicon Valley Edge: A Habitat for Innovation and Entrepreneurship*. Stanford University Press, Stanford, CA, 218–247.
- Davis, L. 1965. The investment market, 1870–1914: The evolution of a national market. *J. Econom. Hist.* **25** 355–399.
- Farmer, D., J. Geanakoplos. 2004. Power laws in finance and their implications for economic theory. Mimeo, Santa Fe Institute, Santa Fe, MN.
- Fleming, L., L. Colfer, A. Marin, J. McPhie. 2004. Why the valley went first: Agglomeration and emergence in regional inventor networks. Working paper, Harvard Business School, Cambridge, MA.
- Florida, R., M. Kenney. 1988. Venture capital and high technology entrepreneurship. *J. Bus. Venturing* **3** 301–319.
- Franks, J., C. Mayer, S. Rossi. 2005. Ownership: Evolution and regulation. Working paper, London Business School, London, UK.
- Gabaix, X. 1999. Zipf's law for cities: An explanation. *Quart. J. Econom.* **114** 739–767.
- Goetzmann, W., M. Massa, A. Simonov. 2004. Portfolio diversification, proximity investment, and city agglomeration. CEPR Discussion Paper 4786, Centre for Economic Policy Research, London, UK.
- Gompers, P., J. Lerner. 2001. The venture capital revolution. *J. Econom. Perspect.* **20** 145–168.

- Greif, A. 1993. Contract enforceability and economic institutions in early trade: The maghribi traders' coalition. *Amer. Econom. Rev.* **83** 525–548.
- Guimera, R., B. Uzzi, J. Spiro, L. Amaral. 2005. Team assembly mechanisms determine collaboration network structure and team performance. *Science* **308**(5722) 697–702.
- Gulati, R., M. Gargiulo. 1998. Where do interorganizational networks come from? *Amer. J. Sociol.* **104**(5) 1439–1493.
- Hoffman, P., G. Postel-Vinay, J.-L. Rosenthal. 2000. *Priceless Markets: The Political Economy of Credit in Paris, 1662–1869*. University of Chicago Press, Chicago, IL.
- Hoffman, P., G. Postel-Vinay, J.-L. Rosenthal. 2004. Révolution et évolution: Les marchés de crédit notarié en France 1780–1840. *Annales ESC* **59** 387–424.
- Hsu, D. 2004. What do entrepreneurs pay for venture capital affiliation? *J. Finance* **59** 1805–1844.
- Hsu, D., M. Kenney. 2005. Organizing venture capital: The rise and demise of the American research and development corporation, 1946–1973. *Indust. Corporate Change* **14** 579–616.
- Kaplan, S., B. Sensoy, P. Strömberg. 2002. How well do venture capital databases reflect actual investments? Working paper, University of Chicago, Chicago, IL.
- Klepper, S. 2002. The capabilities of new firms and the evolution of the US automobile industry. *Indust. Corporate Change* **11** 645–666.
- Kortum, S., J. Lerner. 2000. Assessing the contribution of venture capital to innovation. *RAND J. Econom.* **31** 674–692.
- Laherrère, J. 1996. Parabolic fractal distributions in nature. *C.R. Acad. Sci., Ser. Ila* **322**(7) 535–541.
- Lamoreaux, N. 1994. *Insider Lending: Banks, Personal Connections, and Economic Development in Industrial New England*. Cambridge University Press, New York.
- Lerner, J. 1994. The syndication of venture capital investments. *Financial Management* **23** 16–27.
- Michie, R. 1999. *The London Stock Exchange: A History*. Oxford University Press, Oxford, UK.
- Mitzenmacher, M. 2004. A brief history of generative models for power law and lognormal distributions. *Internet Math.* **1** 226–251.
- Neal, L. 1990. *The Rise of Financial Capitalism: International Capital Markets in the Age of Reason*. Cambridge University Press, Cambridge, UK.
- Newman, M. 2001. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* **98** 404–409.
- Newman, M. 2005. Power laws, Pareto distributions, and Zipf's law. *Contemporary Phys.* **46** 323–351.
- Newman, M., S. Strogatz, D. Watts. 2001. Random graphs with arbitrary degree distributions and their applications. *Physical Rev. E* **64** 026118-1–026118-17.
- Padgett, J., C. Ansell. 1993. Robust action and the rise of the Medici, 1400–1434. *Amer. J. Sociol.* **98** 1259–1320.
- Piskorski, M. J. 2004. Networks of power and status: Reciprocity in venture capital syndicates. Working paper, Harvard Business School, Cambridge, MA.
- Powell, W. W., D. R. White, K. W. Koput, J. Owen-Smith. 2004. Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *Amer. J. Sociol.* **110** 1132–1205.
- Sorenson, O., T. Stuart. 2001. Syndication networks and the spatial distribution of venture capital investments. *Amer. J. Sociol.* **106** 1546–1588.
- Walker, G., B. Kogut, W. Shan. 1997. Social capital, structural holes and the formation of an industry network. *Organ. Sci.* **8** 109–125.
- Watts, D., S. Strogatz. 1998. Collective dynamics of "small world" dynamics. *Nature* **393** 440–442.
- West, G., J. Brown, B. Enquist. 2000. The origin of universal scaling laws in biology. J. H. Brown, G. B. West, eds. *Scaling in Biology*. Oxford University Press, Oxford, UK, 87–112.