Venture Capital: A Tale of Three Networks

Ke Shi*

Caltech

May 5, 2025

Preliminary. Comments welcome.

Abstract

This paper examines how informal and formal networks shape performance in the venture capital (VC) industry. Using data on all U.S.-based VC investments from 1990 to 2009, supplemented with partner-level educational and employment histories from LinkedIn, I develop a structural framework that connects three types of networks: coinvestment ties, historical affiliations, and latent social connections. In the baseline model, VC performance is a function of peer performance, capturing network spillovers through a micro-founded production function. To address endogeneity in network formation, I extend the model using a two-step instrumental variables strategy that leverages variation in past professional and alumni ties. Finally, I introduce an endogenous network formation model in which VCs strategically choose connections based on expected peer quality, allowing for the recovery of latent social networks from equilibrium outcomes. Across specifications, better-connected VCs exhibit significantly higher exit rates. Estimates from the endogenous model suggest that a 1% increase in social connectedness raises a VC's exit rate by 0.2 percentage points, while a 1% improvement in peer performance leads to a 0.74 percentage point increase in connection intensity. These findings highlight the economic value of informal relationships and offer new empirical tools for measuring network effects in private capital markets.

---

*California Institute of Technology. Email: keshi@caltech.edu.

1

1 Introduction

The venture capital (VC) industry plays a central role in financing innovation and entrepreneurship. Yet access to deals, capital, and follow-on support is often governed not just by investment strategy or firm performance, but by personal relationships built through shared education, early career ties, or social circles (Da Rin, Hellmann, and Puri 2013). While it is well-documented that VC firms frequently collaborate extensively through coinvestment (syndication), the informal ties that underlie these partnerships are far less visible and poorly understood. These social connections may facilitate trust, information sharing, and access to deals, but they may also entrench a small group of insiders ("the old boy network") and restrict entry into an already concentrated industry.[1] Understanding how these relationships shape investment outcomes is critical to evaluating both the efficiency and equity of the VC ecosystem (Hochberg, Ljungqvist, and Lu 2010; Ewens 2023).

This paper studies the causal impact of VC networks on fund performance, with a particular focus on the informal, personal connections that traditional data sources overlook. While prior research has found that better-connected VCs perform better (Hochberg, Ljungqvist, and Lu 2007; Tian 2011), it remains difficult to determine whether networks drive performance or simply reflect it (Da Rin, Hellmann, and Puri 2013). Coinvestment-based measures capture only formal, observed relationships and cannot account for latent social ties. Moreover, existing empirical approaches often rely on coarse measures of network centrality without a clear economic interpretation. To address these challenges, I develop a structural network framework that connects three networks and estimate how VC performance improves through these connections: (1) formal *coinvestment* ties formed through joint startup funding; (2) *historical* (alumni and professional) connections based on shared education and prior employment; and (3) informal *social* networks that emerge from these past affiliations and shape ongoing collaboration and information flow.[2] By

---

[1]See, for example, `https://www.forbes.com/sites/oliversmith/2019/02/03/new-industry-report-exposes-british-vc-industry-as-an-old-boys-club`.

[2]The three networks refer to different layers of relationships in the VC industry. (1) The coinvestment network consists of formal ties established when VC firms jointly invest in startups, a common practice that connects nearly all major VCs in the U.S. market (Lerner 1994; Brander, Amit, and Antweiler 2002; Lerner, Shane, and Tsai 2003; Hochberg, Ljungqvist, and Lu 2007). (2) The historical network captures long-standing connections between VC partners formed through shared educational and professional backgrounds, such as attending the same universities or working at the same firms before entering venture capital (Rider 2012; Shue 2013; Huang 2022). See also `https://news.crunchbase.com/data/venture-capitalists-go-college/`. (3) The social network reflects informal and personal relationships among VC partners built upon

leveraging quasi-exogenous variation in VC partners' past affiliations and endogenizing network formation, I identify the causal effects of social connections on investment success and recover the structure of informal networks that shape outcomes in venture capital.

At the heart of this paper is a simple intuition: in a networked environment, a VC's productivity depends not only on its own effort and capabilities, but also on the productivity of its connected peers. This interdependence arises because venture capital is an information-intensive business where relationships facilitate the flow of soft information, expertise, and reputational signals. Networks allow VCs to reduce uncertainty and improve decision-making at two critical stages of the investment process: screening and value creation. During screening, VCs benefit from shared signals, referrals, and joint due diligence with trusted peers, which improves selection quality and mitigates adverse selection risk. In the post-investment phase, networks expand the resources available to portfolio companies such as strategic advice, hiring support, and operational contacts and increase the likelihood of securing follow-on funding. Moreover, in the two-sided matching process between startups and investors, networks serve as a signal of reputation and credibility. Well-connected VCs are more attractive to high-quality entrepreneurs, not only because of their resources but also because their connections reflect market validation (Sørensen 2007; Nahata 2008). The structural model formalizes this mechanism by allowing a VC's performance to depend on the expected performance of its network neighbors, capturing how information and influence propagate through the network to shape investment outcomes.

The structural network model presented in Section 3 builds on the framework developed by Battaglini, Sciabolazza, and Patacchini (2020) and Battaglini, Patacchini, and Rainone (2021), and is implemented in two stages. The first stage, described in Section 3.1, introduces a baseline model in which VCs are endowed with a fixed set of connections, and performance arises through information diffusion across the network. At its core is a simple production function: a VC's performance depends on both its own effort and the performance of its connected peers. This specification captures the idea that connected VCs share information and thereby improve each other's outcomes. A key innovation relative to previous VC network literature is the incorporation of a micro-founded mechanism linking networks to performance. Since effort is chosen in anticipation of peer outcomes, and performance is itself shaped by the network, the model yields a system of interdepen-

---

these historical ties that facilitate mutual support and information exchange, even in the absence of direct coinvestment activity.

dent equations in which all VCs' performances are jointly determined. Under standard regularity conditions, the system admits a unique equilibrium. This formulation allows for a direct quantification of social spillovers and provides a structural interpretation of performance as an equilibrium-based measure of network centrality.

To address the endogeneity of network formation, I extend the model to allow for endogenous link choice among VCs. In this formulation, detailed in Section 3.2, VCs select their social connections in a first stage based on rational expectations about the equilibrium performance of their peers. The model is structured as a two-period game: in period one, agents choose links in anticipation of future benefits; in period two, they select effort levels given the realized network. Connection costs depend on observed compatibility, proxied by shared professional and educational history as well as characteristic similarity. Agents internalize both the informational benefits and formation costs of each connection, resulting in an equilibrium network shaped by strategic behavior.

This structure allows the model to jointly identify the magnitude of peer spillovers and the elasticity of link formation. In doing so, it recovers latent social ties, informal relationships not directly observed in coinvestment data, by leveraging variation in performance, historical affiliations, and cross-sectional differences in characteristics. The intuition is straightforward: if two VCs are highly similar, share extensive past connections, and both perform well, a strong underlying social link is likely; if performance diverges despite those similarities, the connection is likely weaker. Crucially, this approach does not rely on observed coinvestment data, allowing for a conceptual and empirical distinction between formal investment ties and informal social networks.

The results in Section 5 follow the structure of the modeling framework. In the baseline model, performance is systematically related to the performance of a VC's connected peers. A 10 percentage point increase in a coinvestor's exit rate is associated with a 0.1 percentage point increase in the VC's own exit rate. The magnitude of this estimate is comparable to the reduced-form regression relating performance to centrality measures. This peer effect is robust across alternative specifications that use professional and alumni networks in place of coinvestment ties, suggesting that informal connections can carry similar informational value and play a comparable role in driving performance.

The baseline model is then extended using a two-step IV approach. This method leverages professional and alumni networks, constructed from LinkedIn profiles of VC partners, as sources of exogenous variation. These historical affiliations serve as proxies for prior relationships that are unlikely to be influenced by current fund performance. In

4

the first step, coinvestment links are explained using shared educational and professional backgrounds as well as characteristic similarity, under the assumption that VCs tend to partner with peers who are similar to themselves. The residuals from this link formation model capture unobserved factors affecting connection decisions. In the second step, these residuals are used as controls in the main performance equation to address potential selection and simultaneity. Results from this specification show that network spillovers remain positive and statistically significant, with magnitudes comparable to those in the baseline model. Moreover, the insignificance of the residual term suggests that professional and alumni networks account for much of the unobserved heterogeneity in network formation. These findings reinforce the view that long-standing social ties play a meaningful role in structuring VC networks and shaping fund outcomes.

Building on the baseline and IV models, the final specification allows for endogenous network formation, where VCs strategically select their connections in anticipation of performance gains. This approach jointly estimates both the impact of social ties on outcomes and the responsiveness of network formation to peer quality. The results show that a 1% increase in a VC's social connectedness, whether through forming new links or strengthening existing ones, is associated with a 0.2 percentage point increase in its own exit rate. At the same time, a 1% increase in a peer's performance leads to a 0.74 percentage point increase in connection intensity, indicating that VCs actively reconfigure their networks in response to the quality of their peers. Unlike the previous specifications, which rely on observed coinvestment or historical affiliations, this model recovers the latent social network directly from performance outcomes, past ties, and characteristic similarity. The recovered network shares many features with the coinvestment network but also reveals distinct patterns of informal connectivity. These differences suggest that personal and unobserved relationships play an important and independent role in shaping performance in venture capital.

The remainder of this article is organized as follows. Section 2 describes the data and establishes reduced-form evidence consistent with Hochberg, Ljungqvist, and Lu (2007). Section 3 presents the structural network model following Battaglini, Sciabolazza, and Patacchini (2020) and Battaglini, Patacchini, and Rainone (2021) and Section 4 presents the details of the estimation. Section 5 presents the estimation results of the structural models. Section 6 concludes.

1.1 Related literature

This paper contributes to three strands of literature. First, it advances research on the determinants of venture capital performance. Unlike traditional asset classes, VC investing requires intensive screening, monitoring, and value-added engagement under high uncertainty. A growing literature highlights the importance of partner characteristics, fund size, stage specialization, and experience in shaping returns (Kaplan and Schoar 2005; Cochrane 2005). Another distinctive feature of the VC industry is syndication. These coinvestment partnerships serve not only financial purposes but also function as channels for information exchange and strategic alignment. Hochberg, Ljungqvist, and Lu (2007) show that centrality in the coinvestment network is positively associated with fund success, suggesting that better-connected VCs benefit from enhanced deal flow and information. Similarly, Sørensen (2007) models VC-startup matching as a two-sided process, emphasizing how relationships shape selection. However, most of this literature relies on reduced-form methods and observable investment ties, which limit causal interpretation and overlook the role of latent social connections.

This paper provides the first structural estimation of VC networks, capturing how performance is endogenously shaped by the productivity of peers. The model formalizes the information-based mechanisms emphasized in the literature, both screening and value-adding, and extends earlier insights on syndication and matching (Sørensen 2007; Hochberg, Ljungqvist, and Lu 2007; Sorenson and Stuart 2001; Sorenson and Stuart 2008; Das, Jo, and Kim 2011). Crucially, the structural framework identifies both how networks influence performance and how performance, in turn, affects network formation. This dual direction of influence is often missing from prior work, which typically focuses on either the value of network position (Sorenson and Stuart 2001; Sorenson and Stuart 2008) or the determinants of link formation (Lerner 1994; Du 2016; Bubna, Das, and Prabhala 2020). By modeling both sides simultaneously, this paper offers a more comprehensive view of how networks shape outcomes in venture capital.

Second, this paper is related to a broader literature on social capital and informal networks in finance. In public markets, personal connections have been shown to affect trading patterns, capital flows, and corporate decisions. Cohen, Frazzini, and Malloy (2008) document that mutual fund managers with shared educational ties exhibit similar trading behavior, while Engelberg, Gao, and Parsons (2012) show that social relationships influence capital allocation decisions among fund managers. In the corporate sphere, Shue (2013) finds that CEO networks, particularly those based on educational background, af-

fect corporate policy choices. These studies highlight the role of informal relationships in shaping economic behavior, even in relatively transparent markets. In contrast, venture capital is a private, opaque market in which trust and repeated interaction are even more critical, yet the role of informal social networks remains underexplored. This paper extends the logic of social capital into the VC context by quantifying the causal effects of unobserved social ties not captured by coinvestment data on fund performance.

Third, this paper builds on structural approaches to modeling networks and peer effects in economics and finance (Allen and Babus 2009). Foundational work by Acemoglu et al. (2012) and Elliott, Golub, and Jackson (2014) models how network-based spillovers contribute to aggregate outcomes and systemic risk. In terms of empirical implementations (Graham 2020), More recent contributions by Battaglini, Sciabolazza, and Patacchini (2020), Battaglini, Patacchini, and Rainone (2021), and Lewbel, Qu, and Tang (2023) develop structural frameworks that allow for endogenous peer effects and link formation. I adapt this approach to the venture capital setting, estimating both how performance depends on peers and how relationships are formed in equilibrium. A key innovation is the use of historical biographical data (shared education and prior employment) to instrument for unobserved social ties. This enables the recovery of latent social networks and the identification of their causal effect on performance. Despite widespread belief in their importance, such informal networks remain largely unmeasured in the VC literature and are rarely incorporated into models of financial intermediation.

2   Data Description

2.1   VC data

The VC data covers all recorded deal flows involving U.S.-based venture capital firms between 1990 and 2009. The cutoff year of 2009 is selected to ensure that the performance of each VC fund can be meaningfully assessed, given the typical life cycle of a VC fund is approximately ten years. Each funding round involves a target company and a syndicate of VCs, although the composition of the syndicate may change across rounds. Startups may receive multiple rounds of financing prior to an exit, which is classified as either an initial public offering (IPO), an acquisition by another firm, or a failure. Exit dates and modes are observed for completed cases. For firms still listed as active, a company is assumed to have failed if it has not received a new round of financing within the past five years, consistent with evidence that the operational life cycle of most startups does not exceed a

decade.

The sample is restricted to traditional VC firms, defined as small partnerships focused exclusively on early-stage investing. Institutions such as investment banks, large corporate investors, and healthcare companies are excluded due to their scale and diversified operations, which obscure meaningful identification of inter-firm connections. Furthermore, the analysis includes only those VC firms with at least one partner who has a publicly accessible LinkedIn profile, as professional and alumni networks constructed from these data serve as key sources of exogenous variation. The final sample comprises 15,777 funding rounds involving 670 VC firms. Summary statistics are reported in Table 1.

VC performance

Following prior studies (Das, Jo, and Kim 2011; Du 2016; Lindsey 2008; Hochberg, Ljungqvist, and Lu 2007), VC performance is defined as the proportion of a firm's portfolio companies that have successfully exited the market through either an initial public offering (IPO) or an acquisition. Throughout the paper, the terms "exit rate" and "VC performance" are used interchangeably. While direct data on fund-level returns would provide a more precise measure of financial performance, such information is generally unavailable due to the absence of regulatory disclosure requirements for private VC firms. Despite this limitation, exit rate serves as a credible proxy, as successful exits are a key determinant of realized returns in the industry. Moreover, exit rate is bounded between zero and one, which facilitates model estimation and improves numerical stability. Summary statistics on VC exit rates are reported in Table 1.

Coinvestment networks

The coinvestment network is constructed using observed VC deal activity. For each round of funding, the data identify all participating VCs. The adjacency matrix **G** is defined such that for any pair of VCs i and j, the element g_{ij} records the total number of coinvestments between them over the observed period. This formulation results in a weighted network, where g_{ij} reflects the intensity or strength of the connection. By convention, self-links are excluded, so $g_{ii} = 0$ for all i. For empirical implementation, two alternative measures of connection intensity are also considered: a binary indicator equal to one if i and j have ever coinvested, and a log-transformed version of the raw coinvestment count. Table 2 reports summary statistics for the coinvestment network. The unit of observation is an ordered VC pair, yielding $n(n-1)$ dyads in total for n VCs in the sample.

8

Table 1: Summary statistics of VC firms

	N	Mean	SD	Min	Max
No. rounds	6713	25.38	81.87	1	2373
No. startups	6713	13.80	37.39	1	989
Experience (years)	6713	6.35	7.66	0.00	39.23
No. coinvestments	6713	82.23	272.02	1	8075
No. coinvestors	6713	35.25	72.59	1	1327
Performance					
No. IPOs	6713	1.69	7.24	0.00	186.00
No. acquisitions	6713	5.09	15.87	0.00	375.00
No. write-offs	6713	4.29	10.91	0.00	285.00
No. private companies	6713	2.73	7.68	0.00	148.00
IPO rate	6713	0.082	0.19	0.00	1.00
Exit rate	6713	0.41	0.38	0.00	1.00
Attributes					
Pct business & financial services	6713	0.19	0.28	0.00	1.00
Pct consumer goods & services	6713	0.13	0.24	0.00	1.00
Pct healthcare	6713	0.22	0.34	0.00	1.00
Pct information technology	6713	0.38	0.36	0.00	1.00
Pct female	6713	0.022	0.11	0.00	1.00
Pct Asian	6713	0.023	0.13	0.00	1.00
Centrality					
Degree	6713	35.25	72.59	1.00	1327.00
Betweenness	6713	0.00029	0.0014	0.00	0.059
Harmonic	6713	0.34	0.068	0.00015	0.57
Eigenvector	6713	0.039	0.081	0.00	1.00

Notes: Summary statistics of VC characteristics based on VC deals data.

Given the adjacency matrix **G**, four centrality measures are computed to characterize the position of each VC within the coinvestment network, following standard concepts from network and graph theory. Each VC is treated as a vertex, and each connection as an edge. (1) The degree centrality of a vertex is defined as the number of distinct connections it has to other vertices. Since the network is undirected in this application, no distinction is made between in-degree and out-degree.[3] In the weighted version of the network, degree centrality can also incorporate the number of coinvestments as edge weights. (2) The betweenness centrality measures the number of shortest paths between all pairs of nodes that pass through a given vertex. For each pair of VCs in the network, there exists

---

[3]See Hochberg, Ljungqvist, and Lu (2007) for a discussion of directionality in the context of VC networks.

at least one shortest path that minimizes the number of intermediate steps (or the total edge weight in the case of weighted networks). Betweenness thus captures the extent to which a VC serves as a bridge within the network. (3) The harmonic centrality, closely related to closeness centrality, is the inverse of the average shortest path length from a node to all other reachable nodes in the network. This metric reflects how easily a VC can access the broader network of peers. (4) The eigenvector centrality measures a VC's influence based on the principle that connections to highly connected nodes contribute more to one's centrality. Formally, this metric is derived from the eigenvector associated with the principal eigenvalue λ in the linear system $\lambda \mathbf{x} = \mathbf{G}\mathbf{x}$. Summary statistics for these centrality measures are reported in Table 1.

It is also useful to introduce the concept of alpha centrality (sometimes also named after Katz 1953; Bonacich and Lloyd 2001), which will serve as a foundation for several structural formulations discussed later. Alpha centrality generalizes eigenvector centrality by incorporating external sources of influence. Formally, it is defined as the solution to the linear system:

$$\mathbf{x} = \delta \mathbf{G}\mathbf{x} + \varepsilon, \tag{1}$$

where $\mathbf{x}$ is the centrality vector, $\mathbf{G}$ is the adjacency matrix, ε is a vector of exogenous influence, and δ determines the relative weight of endogenous network effects versus external shocks. When ε is set to zero, this formulation reduces to eigenvector centrality. Alpha centrality can also be interpreted as a generalized form of degree centrality, where the influence of more distant nodes is discounted. The structure in equation (1) will reappear in later sections with different behavioral and economic interpretations.

Covariates

Several VC-level characteristics are included as covariates in the analysis. These variables are selected based on their potential influence on performance and their prominence in the venture capital literature.

First, performance is expected to correlate with fund size and industry specialization. Although direct observations of fund size are unavailable, two proxies are constructed: the number of startups backed by a VC and the number of funding rounds in which the VC has participated. These measures serve as reasonable indicators of investment capacity under the assumption that larger firms typically engage in more deals.

Table 2: Summary Statistics of Pairwise Connection Intensities

	N	Mean	SD	Min	Max
Coinvestment, **G**					
Having coinvested	45064369	0.01	0.07	0.00	1.00
No. coinvestments	236628	2.33	3.49	1.00	129.00
log(No. coinvestments)	236628	0.51	0.69	0.00	4.86
Professional connections, **H**$_p$					
Having professional connections	45064369	0.00	0.04	0.00	1.00
No. professional connections	87758	23.28	657.18	1.00	88977.00
log(No. professional connections)	87758	0.70	1.17	0.00	11.40
Alumni connections, **H**$_a$					
Having alumni connections	45064369	0.01	0.09	0.00	1.00
No. alumni connections	357024	9.92	113.69	1.00	20628.00
log(No. alumni connections)	357024	0.95	1.14	0.00	9.93

Notes: The unit of observation is a VC-VC pair. The number of coinvestments is calculated based on the common funding round that both VCs participated in. Professional connections and alumni connections are calculated at the individual level and aggregated at the VC level. For example, if partner A from VC 1 and partner B from VC 2 have both worked at the same company prior to joining their respective VCs, this is one professional connection.

Second, VCs often concentrate their investments within one or a few sectors to leverage expertise and avoid the inefficiencies associated with over-diversification. Four major sectors are identified in the data: business and financial services, consumer goods and services, healthcare, and information technology.

Third, demographic composition is captured by two variables: the share of female partners and the share of Asian partners within each VC firm. The venture capital industry remains predominantly white and male, making it important to understand the role of gender and racial diversity in shaping outcomes. Gender and ethnicity are imputed using first and last names extracted from LinkedIn profiles. The classification algorithm is conservative, resolving ambiguous cases in favor of male and non-Asian designations. The analysis focuses on the Asian versus non-Asian distinction for two reasons. Asian names are more reliably identified using this method, and the Asian presence in the industry is large enough to offer meaningful variation. Approximately 10 percent of partners in the sample are identified as Asian. Summary statistics for these covariates are reported in Table 1.

2.2 VC Partner Data

The VC data are supplemented with firm-level information on professional and alumni networks, constructed from the LinkedIn profiles of VC partners. LinkedIn is an online platform for professional networking where individuals voluntarily disclose their career history, educational background, and other credentials. The underlying dataset was assembled by a private data provider in 2017 through large-scale web scraping of publicly available LinkedIn profiles, capturing a range of attributes including employment history and education.

For the present study, the dataset is filtered to include individuals identified as partners or directors at the VC firms in the main sample. While LinkedIn data are self-reported and may contain inaccuracies, such concerns are mitigated by the incentives for senior professionals to maintain accurate public profiles. In addition, manual screening was conducted to remove spurious or clearly inconsistent entries.

A remaining limitation is that not all individuals maintain LinkedIn profiles, a gap more pronounced among smaller VC firms with fewer listed partners. As a result, the coverage of historical networks may be incomplete, potentially attenuating the estimated effect of alumni and professional ties. Although the LinkedIn data are at the individual level, all professional and alumni networks are aggregated to the firm level for empirical analysis.

Professional networks

Professional networks are constructed using work history data from the LinkedIn profiles of VC partners. An adjacency matrix $\mathbf{H}_p$ is defined such that each element h_{ij} represents the number of shared work experiences between partners at VCs i and j. A shared experience is defined as a case in which at least one partner from each VC has worked at the same company at some point in time. The value of h_{ij} is computed as the total number of such pairwise overlaps across all partners of the two firms.

It is acknowledged that not all shared affiliations reflect actual interpersonal relationships, as individuals may not have worked together directly. Consequently, the measure h_{ij} should be interpreted as a proxy for the potential basis of professional ties, rather than a direct measure of existing social links. Shared employment history lowers the barrier to future interaction and thus serves as a plausible foundation for informal networking. Summary statistics for pairwise professional connections are presented in Table 2. Three

forms of the variable are reported: the raw count of shared affiliations, a log-transformed version, and a binary indicator equal to one if at least one shared connection exists.

Alumni networks

Alumni networks are constructed analogously, based on educational background. The adjacency matrix $\mathbf{H}_a$ is defined such that h_{ij} denotes the number of alumni connections between VCs i and j. A connection is counted when one partner from each VC has attended the same educational institution. The final value of h_{ij} is the sum of all such pairwise overlaps across partners from both firms. This measure captures potential affinity or ease of networking that may arise from shared educational backgrounds. As with professional networks, the alumni network is used as a proxy for the potential basis of informal ties. Summary statistics for these connections are also reported in Table 2.

2.3   Evidence of network on performance

To establish a benchmark, the correlation between VC performance and network position is examined, following the empirical strategy of Hochberg, Ljungqvist, and Lu (2007). The econometric model is

$$ExitRate_i = \gamma Centrality_i + X_i\beta + \varepsilon_i, \tag{2}$$

where the dependent variable $ExitRate_i$ denotes the proportion of a VC's portfolio companies that successfully exited through either an IPO or an acquisition. The central explanatory variable is a measure of network centrality, constructed using various definitions outlined above. While Hochberg, Ljungqvist, and Lu (2007) address endogeneity by constructing time-lagged centrality measures based on coinvestment activity in the five years preceding each fund's vintage year, the present specification abstracts from the time dimension. The purpose of this reduced-form model is to document the strength of the correlation between performance and network position. Endogeneity concerns are addressed in subsequent sections using a structural framework.

Table 3 reports the estimation results from equation (2) and serves as a baseline for comparison with later structural estimates. All coefficients on the centrality measures are positive and statistically significant, consistent with theoretical expectations. For example, an additional connection—corresponding to a one-unit increase in degree centrality—is associated with a 0.2 percentage point increase in exit rate, holding other factors constant.

Table 3: Reduced-form evidence of network effect on VC performance

	Dependent variable:								
	Exit rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Degree	0.002*** (0.0002)					0.002*** (0.0004)			
Betweenness		9.857*** (2.068)					15.972*** (4.092)		
Harmonic centrality			8.517*** (0.850)					6.517*** (0.952)	
Eigenvector centrality				0.637*** (0.057)					0.929*** (0.121)
No. startups					0.001*** (0.0001)	−0.001*** (0.0003)	0.002*** (0.0003)	0.001*** (0.0001)	−0.001*** (0.0003)
Percent business & finance					0.306*** (0.093)	0.277*** (0.090)	0.289*** (0.092)	0.335*** (0.090)	0.261*** (0.089)
Percent consumer G&S					0.190* (0.099)	0.164* (0.097)	0.177* (0.098)	0.184* (0.096)	0.150 (0.095)
Percent healthcare					0.425*** (0.090)	0.394*** (0.088)	0.392*** (0.089)	0.419*** (0.087)	0.385*** (0.086)
Percent info tech					0.405*** (0.092)	0.364*** (0.090)	0.370*** (0.091)	0.379*** (0.089)	0.330*** (0.089)
Percent female					0.014 (0.039)	0.019 (0.038)	0.016 (0.038)	0.010 (0.038)	0.021 (0.037)
Percent Asian					−0.016 (0.034)	−0.017 (0.033)	−0.016 (0.033)	−0.026 (0.032)	−0.019 (0.032)
Constant	0.339*** (0.013)	0.398*** (0.012)	0.100*** (0.033)	0.332*** (0.013)	0.047 (0.083)	0.035 (0.080)	0.064 (0.082)	−0.168* (0.086)	0.049 (0.079)
Observations	670	670	670	670	670	670	670	670	670
R^2	0.133	0.033	0.131	0.156	0.143	0.190	0.163	0.200	0.214
Adjusted R^2	0.131	0.031	0.129	0.155	0.134	0.181	0.153	0.190	0.204

Notes: Estimates of equation (2) of various specifications are presented. Columns (1)-(4) only use centrality measure as the explanatory variable. Columns (5)-(8) include additional covariates. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels.

A one-standard-deviation increase in betweenness centrality (approximately 0.005) corresponds to a 5 percent increase in the exit rate. These magnitudes are comparable to those reported in Hochberg, Ljungqvist, and Lu (2007), reinforcing the robustness of the centrality-performance relationship.

Before turning to the structural framework, several limitations of the reduced-form approach merit discussion. First, the analysis is subject to significant endogeneity concerns arising from both omitted variables and reverse causality. Unobserved characteristics, such as partner ability or reputation, may influence both network position and fund performance. For instance, highly capable VCs may be more effective at securing attractive deals and cultivating strategic relationships. Additionally, reverse causality is plausible:

successful funds are more likely to attract coinvestors, and better-performing VCs may be selectively targeted for syndication by their peers. As a result, observed centrality may reflect performance outcomes rather than cause them.

Second, centrality measures, aside from degree, are abstract summaries of network position that are nonlinear in connections and difficult to interpret economically. These metrics do not provide insight into the marginal effect of adding a connection or improving the quality of an existing one. For example, while it is possible to estimate the effect of a one-standard-deviation increase in betweenness or eigenvector centrality, such changes do not map clearly onto intuitive or policy-relevant interpretations. Even degree centrality, which counts direct connections, captures only local network effects and ignores the broader influence of indirect ties. It also conflates the number and quality of connections, making it difficult to assess whether ties are formed with high- or low-performing peers. A more precise interpretation of network effects requires a model with explicit micro-foundations that links performance directly to the characteristics and outcomes of connected agents. This is the focus of the next section.

3 Structural Network Model

Following Battaglini, Sciabolazza, and Patacchini ([2020](#)) and Battaglini, Patacchini, and Rainone ([2021](#)), this section presents a condensed version of the structural model and its econometric specification. Full details are provided in the appendix and the referenced papers. The central feature of the model is the equilibrium determination of VC performance, where each VC's outcome depends on the networked interactions with peers. Two network models are introduced in sequence: an exogenous model, in which VCs are endowed with predetermined connections, and an endogenous model, in which VCs strategically form links in anticipation of their performance implications. The analysis begins with a micro-founded production function that captures how social connections contribute to performance, drawing on foundational ideas from information economics.

3.1 Baseline exogenous networks

Production function

Financial intermediary markets, including venture capital, are characterized by high levels of uncertainty, risk, and information asymmetry. To mitigate these frictions, VCs play two critical roles: screening and value-adding. The startup landscape is saturated with ven-

tures, but only a small subset will generate outsized returns. VCs must carefully screen opportunities before committing capital and, once invested, actively support portfolio companies through strategic guidance, operational expertise, and access to networks—until exit via acquisition or IPO becomes viable. Both functions rely heavily on information: screening hinges on the ability to identify high-potential startups amid noisy signals, while value-adding depends on the VC's access to relevant resources and connections. In a networked environment, the quality and reach of information are shaped by whom a VC is connected to, both in terms of the number (extensive margin) and productivity (intensive margin) of its peers. This motivates a simple structural framework in which a VC's performance depends on the performance of its connected peers, in addition to its own effort and characteristics.

Consider a market that consists of n VCs, indexed by $\mathcal{N} = \{1, \dots, n\}$. Each VC $i \in \mathcal{N}$ wants to maximize its *performance* P_i that follows the Cobb-Douglas production function with two inputs,[4] *social connectedness* s_i and own effort l_i:

$$P_i = \rho s_i^\alpha l_i^{1-\alpha} + \varepsilon_i, \tag{3}$$

where ε_i is an idiosyncratic shock and $\rho > 0$ is a productivity constant. Social connectedness s_i is defined as a weighted average of the performance of VC i's network peers:

$$s_i = \sum_{j \in \mathcal{N}} g_{ij} P_j, \tag{4}$$

where $g_{ij} \geq 0$ denotes the intensity of the unilateral social link from VC i to VC j. Let the matrix $\mathbf{G} = (g_{ij})$ denote the structure of the social network, with g_{ij} either binary (denoting the presence of a link) or continuous (capturing connection strength). For now, let us assume that the industry is exogenously endowed with the network $\mathbf{G}$. We will relax this assumption later. Because of the Cobb-Douglas functional form, α is the elasticity of P_i with respect to s_i, that is, the responsiveness of a VC's performance with respect to its social connectedness s_i. Intuitively, g_{ij} captures the quantity of VC i's social ties (the extensive margin), while P_j captures the quality of these ties (the intensive margin). Thus, performance depends not only on individual effort, but also on the productivity of connected peers and the structure of the underlying network.

---

[4]The performance term P_i represents the effectiveness of VC i in generating successful investment outcomes. Empirically, this can be proxied by the fund's historical exit rate—the proportion of portfolio companies that achieve a successful IPO or acquisition.

Before proceeding, it is useful to interpret the economic mechanism captured by equation (3), which links a VC's performance P_i to the performance of its peers through network-based interactions. First, in the screening stage, P_i can be interpreted as the strength or precision of VC i's signal about a startup's quality. Connections to other informed VCs improve signal accuracy through shared information and referrals. Second, in the value-adding phase, P_i captures the social and human capital that enables a VC to support portfolio companies—primarily through targeted advice, strategic connections, and operational expertise. Because VCs rarely engage in day-to-day operations, much of their contribution stems from informational and reputational capital. Third, in the matching process between startups and investors, P_i can also reflect a VC's reputation. Well-connected VCs are more visible and credible to entrepreneurs, and association with high-performing peers signals competence and enhances the likelihood of being selected by top startups.

Exogenous network equilibrium

To close the model, let the cost of effort be linear, given by l_i, so that the VC maximizes net performance $P_i - l_i$. Under mild regularity conditions, there exists a unique equilibrium in which all VCs simultaneously choose optimal effort levels. The resulting equilibrium performance vector P satisfies the following autoregressive system:

$$P = \delta \mathbf{G} P + \varepsilon, \tag{5}$$

where $\delta = \rho^{\frac{1}{\alpha}}(1-\alpha)^{\frac{1-\alpha}{\alpha}}$ is the social spillover (Battaglini, Sciabolazza, and Patacchini 2020).

While the derivation is relegated to the appendix, equation (5) is intuitive: being connected to high-performing peers enhances one's own performance. Comparing equation (5) with the definition of alpha centrality in equation (1), we see that the equilibrium performance vector P corresponds exactly to the alpha centrality measure, with network weight δ and exogenous influence ε. The key distinction is that, in this model, centrality is not an externally computed summary statistic used to explain performance—it is the equilibrium outcome of the model itself. In this sense, performance is not a consequence of centrality; it is centrality. The structural formulation thus provides a deeper behavioral interpretation: being effective is equivalent to being central in the flow of information and influence. Compared to the reduced-form model in equation (2), which regresses performance on precomputed centrality measures, the structural model explains performance as emerging directly from the pattern of connections.

3.2 Endogenous network formation

A key limitation of the exogenous network approach is that it treats link formation as driven solely by observable similarity (e.g., homophily), while ignoring strategic considerations in network formation. If a VC's performance improves due to exogenous factors, it is natural to expect other VCs to seek closer ties, anticipating spillovers from high-performing peers. This reverse causality is not captured in the baseline model (5) and is instead often absorbed into an endogeneity correction. Moreover, the analysis so far assumes that observed coinvestment ties fully reflect relevant connections, but informal and personal relationships—often unobserved—may also influence performance. These latent social ties introduce an additional layer of endogeneity that cannot be addressed through standard correction techniques. Together, these concerns motivate a structural model in which social networks are formed endogenously in equilibrium.

The structure of the model remains similar, but it now unfolds over two periods. In the first period, VC i chooses its network connections $g_i = (g_{i1}, \ldots, g_{in})$ in anticipation of how these links will affect its future performance. In the second period, given the realized network, it then selects effort level l_i, and performance outcomes are determined in equilibrium. Forward-looking VCs optimize their network formation decisions in the first period, internalizing the effect of their connections on equilibrium effectiveness. The equilibrium is defined by the pure strategy profile (g_i, l_i), where g_i maps the VC i's type to a vector of connections, and l_i maps both type and network structure to the chosen effort level.

The final component of the model is the cost of forming social links. In the first period, VC i incurs a cost $c(g_{ij}, \theta_{ij}; \lambda)$ to establish a connection of intensity g_{ij} with VC j. This cost is increasing in the connection strength g_{ij} and decreasing in the pairwise compatibility θ_{ij}, which captures how naturally VCs i and j are able to form a tie (to be specified below). I assume the following isoelastic cost function:

$$c(g_{ij}, \theta_{ij}; \lambda) = \frac{\lambda}{1 + \lambda} \left(\frac{g_{ij}}{\theta_{ij}} \right)^{1 + \frac{1}{\lambda}}, \tag{6}$$

where $\lambda > 0$ captures the curvature of the cost function. As will become clear, λ provides a convenient measure of the elasticity of link formation with respect to peer performance, i.e., how responsive VC i's optimal connection intensity is to the value of being linked to high-performing peers.

Endogenous network equilibrium

Given the setup, Battaglini, Patacchini, and Rainone (2021) defines a *network competitive equilibrium* (l, P, G) that satisfies three conditions: (1) In period 1, each VC chooses a vector of connections $g_i = (g_{i1}, \ldots, g_{in})$ optimally given P (VCs are "price-taking"); (2) In period 2, each VC chooses own effort l_i optimally given P and g_i; and (3) Performance P_i satisfies the production function given l_i and g_i (price must clear the market). Under mild regularity conditions, a unique pure-strategy equilibrium exists with interior solutions. The equilibrium performance P is characterized by

$$P_i = \varphi \sum_j (\theta_{ij} P_j)^{1+\lambda} + \varepsilon_i \tag{7}$$

for all i, where φ is a function of the structural parameters ρ, α, and λ.[5] In equilibrium, the social connectedness G is given by

$$g_{ij} = \theta_{ij}^{1+\lambda}(\alpha \delta P_j)^\lambda \tag{8}$$

for all $i \neq j$.

Equation (7) states that the resulting equilibrium performance is governed by a system of nonlinear equations. The parameter φ captures the strength of network spillovers. Comparing the endogenous system in equation (7) with the exogenous network equilibrium in equation (5), performance can be interpreted as a generalized form of alpha centrality, augmented by a nonlinear component driven by λ. This nonlinearity arises from endogenous network formation: because VC i optimally chooses its connection with j, g_{ij}, proportional to P_j^λ in equation (8), its own performance P_i becomes a function of $P_j^{1+\lambda}$. In this endogenous framework, centrality and performance are no longer separable; being central in the network reflects both connection strength and peer quality, jointly determined through forward-looking strategic behavior.

Finally, under the parametric specification in equation (6), the elasticity of link inten-

---

[5]The closed-form expression is $\varphi = \alpha^\lambda \delta^{1+\lambda}$, where $\delta = \rho^{\frac{1}{\alpha}}(1-\alpha)^{\frac{1-\alpha}{\alpha}}$ is a shorthand parameter in the model identical to that in equation (5). Details are provided in the appendix.

sity g_{ij} with respect to peer performance P_j is exactly equal to λ.[6] This gives λ a convenient and intuitive interpretation: it captures both the sensitivity of link formation to peer quality and the strength of the feedback between performance and network structure. When $\lambda = 0$, equations (7) and (8) reduce to the baseline model (5), in which the network is exogenously given and fixed. Thus, λ provides a structural measure of how much active, performance-driven network formation occurs in the VC industry, and how much the endogenous model improves upon the exogenous benchmark in explaining observed performance.

4 Estimation

4.1 Baseline specifications

For estimation, the unobserved component of performance is assumed to depend linearly on observable VC-level characteristics. Let $\mathbf{X} = [X_1, \cdots, X_n]'$ denote the matrix of covariates. The baseline empirical model takes the form:

$$P = \delta \mathbf{G} P + \mathbf{X}\beta + \varepsilon, \tag{9}$$

which corresponds to a spatial autoregressive (SAR) model commonly used in the network literature. This system can be estimated via maximum likelihood and allows for direct inference on δ, the reduced-form parameter capturing the strength of peer spillovers.[7] If $\delta = 0$, network spillovers are absent, and the model reduces to a standard linear regression on individual characteristics.

Estimation is implemented using multiple specifications of the adjacency matrix $\mathbf{G}$. The baseline specification relies on observed coinvestment ties, which reflect formal collaboration between VC firms. While standard in the literature, this approach is vulnerable to endogeneity arising from simultaneity and omitted variables. To address this, I also consider alternative network matrices based on historical affiliations—specifically, alumni

---

[6]The elasticity of a link g_{ij} with respect to the effectiveness of j is

$$\varepsilon_{g_{ij},P_j} = \frac{\partial g_{ij}}{\partial P_j}\frac{P_j}{g_{ij}} = \theta_{ij}^{1+\lambda}(\alpha\delta P_j)^{\lambda-1}\alpha\delta\frac{P_j}{g_{ij}} = \lambda.$$

[7]Recall that δ is a composite of the structural parameters ρ and α from the Cobb-Douglas production function in equation (3), which cannot be separately identified.

and professional networks, denoted $\mathbf{H}_a$ and $\mathbf{H}_p$, respectively—constructed from biographical information. These past connections are plausibly exogenous to current performance and help mitigate concerns related to unobserved heterogeneity. However, they do not capture the influence of ongoing, contemporaneous interactions. To reconcile this limitation, a two-step estimation procedure is introduced, linking past and current networks while accounting for selection into coinvestment relationships.

Instrumental variable (IV) approach

The baseline specification in equation (9) is subject to endogeneity concerns due to simultaneity and omitted variables. For example, a VC partner's intrinsic ability or socioeconomic background may influence both performance and network formation, leading to biased estimates of peer effects. To address these issues, a two-step Heckman-style correction is introduced to account for selection into network links. Historical affiliations—specifically, alumni and professional ties—serve as proxies for unobserved individual heterogeneity.

In the first stage, the probability of a coinvestment tie between VC i and VC j is modeled as a function of past connections between their partners, based on shared educational and employment backgrounds. This step controls for selection driven by characteristics correlated with both performance and network structure. The residual from this regression captures unobserved factors influencing link formation and is included in the second stage as a control function. The performance equation is then re-estimated, incorporating this correction to isolate the causal effect of peer performance while mitigating endogeneity bias.

The identification strategy using alumni and professional networks relies on two conditions: relevance and exogeneity. Relevance is relatively uncontroversial. Shared educational and occupational experiences are well-documented predictors of long-term professional relationships. In the venture capital setting, such historical affiliations plausibly influence the formation of coinvestment ties, a pattern confirmed in the first-stage regression.

The exogeneity condition is more demanding. It assumes that historical connections affect current performance only through their influence on network formation, not through any direct or unobserved channels. This assumption may be difficult to satisfy in environments like venture capital, where informal and persistent social ties often operate alongside formal partnerships. While these instruments help mitigate endogeneity concerns

related to unobserved ability and background characteristics, they may not fully account for latent social relationships. For this reason, alumni and professional networks are best viewed as useful but partial instruments. The structural model introduced in the next section addresses this limitation by endogenizing network formation directly.

The first step estimates a standard dyadic model of link formation, in which the presence or intensity of a coinvestment connection between VC i and VC j is explained by their historical ties and the distance between their observable characteristics. The specification is given by

$$g_{ij} = \gamma_0 + \gamma_1 h_{ij} + \sum_{\ell} \gamma^{\ell} d(X_i^{\ell}, X_j^{\ell}) + \eta_{ij}, \tag{10}$$

where h_{ij} denotes a past connection through shared educational or professional affiliation, and $d(X_i^{\ell}, X_j^{\ell})$ is a distance metric between VCs i and j along characteristic ℓ. The error term η_{ij} captures unobserved determinants of link formation. Intuitively, the probability or strength of a coinvestment tie increases with prior affiliation and decreases with dissimilarity in key attributes, reflecting homophily in network formation.

In the second step, two estimation strategies are available: a standard two-stage least squares (2SLS) instrumental variable (IV) approach using the predicted links $\hat{g}_{ij}$ from the first stage, or a control function method based on the residuals $\hat{\eta}_{ij}$, analogous to a Heckman correction. The second approach requires an assumption on the covariance of the residuals ε and η_{ij}, which are outlined in the appendix. In particular, it assumes that the correlation is the same between unobserved characteristics determining link formation η_{ij} and the unobserved characteristics driving the outcome ε_i for all VCs. Under this correction, the equilibrium performance equation is augmented as follows:

$$P = \delta GP + X\beta + \psi\xi + \varepsilon, \tag{11}$$

where $\xi_i = \sum_{j \neq i} \eta_{ij}$ aggregates the residuals from the first-stage link formation equation. The additional term $\psi\xi$ captures unobserved individual heterogeneity that influences both network formation and performance, correcting for selection bias in the estimation of peer effects.

4.2 Endogenous network formation

For econometric implementation of endogenous network formation, individual hetero-geneity is assumed to be a linear function of observable characteristics. The resulting equilibrium condition yields the main estimating equation:

$$P_i = \varphi \sum_j (\theta_{ij} P_j)^{1+\lambda} + X_i \beta + \varepsilon_i, \tag{12}$$

where φ and λ are structural parameters, X_i denotes the vector of characteristics for VC i, and ε_i is the idiosyncratic error term. The term θ_{ij} captures the compatibility between VCs i and j and governs the strength of their latent social connection.

Compatibility θ_{ij} is modeled as a Bernoulli random variable, where the probability of a tie is given by a logistic function of χ_{ij}, a latent connectivity index that depends on historical ties and pairwise distances in observable characteristics:

$$Pr(\theta_{ij} = 1 | \chi_{ij}) = \frac{\exp(\chi_{ij})}{1 + \exp(\chi_{ij})}, \tag{13}$$
$$\text{with } \chi_{ij} = \kappa_0 + \kappa_1 h_{ij} + \sum_\ell \kappa^\ell d(X_i^\ell, X_j^\ell),$$

where h_{ij} denotes a prior connection (e.g., shared educational or employment background), and $d(X_i^\ell, X_j^\ell)$ represents the distance between VCs i and j along characteristic ℓ. This formulation links observed historical data to the latent structure of social ties that influence equilibrium performance.

A final note distinguishes equation (13) from the link formation model used in the first step of the two-step network model in equation (10). In equation (13), the outcome variable is θ_{ij}, which represents latent compatibility between VCs and governs the probability of a social tie. In contrast, equation (10) models g_{ij}, the observed coinvestment connection itself, as a function of past ties and characteristic distances. While the two specifications are similar in form, their interpretation is fundamentally different. In the network formation model, social networks g_{ij} are not observed directly but are endogenously recovered from the equilibrium in equation (8). As will be shown in the empirical results, the recovered social networks share important features with observed coinvestment ties but also reveal distinct patterns of informal connectivity.

4.3 Estimation method

The estimation is implemented using Bayesian methods. The main estimation equations are the baseline network model (9), the two-step procedure (11), and the structural model with endogenous network formation (12). In each case, VC performance P appears on both sides of the equations through network interactions, introducing simultaneity that renders classical estimation approaches infeasible or inconsistent. Instead, all models are estimated via a Bayesian framework that accommodates the recursive structure of equilibrium and facilitates inference on the full posterior distribution of parameters.

Specifically, estimation proceeds via Approximate Bayesian Computation (ABC), a simulation-based method particularly suited to structural models with intractable likelihoods. The algorithm builds on the classic Metropolis-Hastings framework (see Metropolis et al. 1953; Hastings 1970) and follows the implementation in Marjoram et al. (2003) and Battaglini, Patacchini, and Rainone (2021). Starting from an initial value of the parameters ω, the algorithm proposes a candidate parameter vector ω' from a pre-specified transition kernel. If the proposed parameter ω' fits the observed data P better according to the equilibrium condition than the current parameter ω does, then the algorithm moves to the proposed parameter ω' with some probability. The algorithm generates a Markov chain with a limiting stationary distribution, which, under the assumption that the model is correctly specified, coincides with the true conditional distribution of the parameter $P(\omega|P)$, the object of our interest.

5 Results

5.1 Baseline specifications

Table 4 reports estimation results from the baseline exogenous network model. This specification relates VC performance to observed network connections, without correcting for endogeneity or accounting for latent social ties. Column (1) provides a benchmark OLS regression of exit rates on VC characteristics alone, ignoring network effects. As expected, performance is positively associated with fund size, measured by the number of supported startups. This finding aligns with prior work showing that larger VCs tend to outperform, likely due to greater resources and broader deal access. Industry specialization also plays a role: VCs focused on healthcare and information technology exhibit higher exit rates, reflecting strong growth in these sectors over the sample period.

Table 4: Estimation results of the baseline network model

	Dependent variable:			
	Exit rate			
	(1)	(2)	(3)	(4)
	No networks	Coinvestment networks	Professional networks	Alumni networks
δ (Social spillover)		0.00934***	0.00127**	0.0107***
		(0.00126)	(0.000548)	(0.000605)
No. startups	0.000997***	-0.00098***	0.000799***	-0.000469*
	(0.000131)	(0.000323)	(0.000159)	(0.000269)
Percent business & finance	0.306***	0.265***	0.295***	0.0694
	(0.0927)	(0.0882)	(0.0925)	(0.166)
Percent consumer G&S	0.19*	0.151	0.179*	-0.0839
	(0.0993)	(0.0945)	(0.0991)	(0.179)
Percent healthcare	0.425***	0.388***	0.415***	0.14
	(0.0898)	(0.0845)	(0.0896)	(0.162)
Percent information tech	0.405***	0.354***	0.392***	0.0845
	(0.0918)	(0.0873)	(0.0917)	(0.165)
Percent female	0.0143	0.0198	0.0097	-0.00277
	(0.0388)	(0.0374)	(0.0387)	(0.0695)
Percent Asian	-0.0163	-0.0212	-0.0195	0.049
	(0.0335)	(0.032)	(0.0334)	(0.06)
Constant	0.047	0.037	0.0444	-0.0701
	(0.0826)	(0.0782)	(0.0823)	(0.148)
Observations	670	670	670	670

Notes: Column (1) reports the OLS of exit rate on VC characteristics. Columns (2) through (4) report the estimates of equation (9), with the coinvestment networks, professional networks, and alumni networks, respectively. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels.

Columns (2) through (4) of Table 4 present results from the baseline network model in equation (9). Column (2) uses observed coinvestment ties **G** as the network matrix, while Columns (3) and (4) use professional and alumni networks, $\mathbf{H}_p$ and $\mathbf{H}_a$, respectively. Across all specifications, the estimated coefficient on the peer effect parameter δ is positive and statistically significant, consistent with the presence of network spillovers in VC performance.

The estimated magnitude of δ admits multiple interpretations. At the intensive margin, a 10 percentage point increase in the exit rate of a coinvestor is associated with a 0.1 percentage point increase in the VC's own exit rate, holding the network fixed. At the extensive margin, forming a new connection with a VC whose exit rate is 10% yields a similar improvement in expected performance. These effects are broadly consistent with the reduced-form estimates in Table 3, where an additional coinvestor is associated with a 0.2 percentage point increase in exit rate based on degree centrality. While the structural estimates are somewhat smaller in magnitude, they offer a more nuanced interpretation

by jointly capturing both the quantity and quality of connections, rather than aggregating ties through centrality alone.

5.2 IV approach

First step: link formation

As a first step in the instrumental variable (IV) strategy, I assess whether historical networks are predictive of current coinvestment ties. Table 5 reports OLS estimates from the dyadic link formation model in equation (10). Columns (1) and (2) use binary indicators for the presence of a coinvestment tie, while Columns (3) and (4) use the raw number of connections between VC pairs as the outcome variable.

The results provide strong evidence of homophily in network formation. Coefficients on the pairwise distances in fund size, industry specialization, and demographic characteristics are consistently negative and statistically significant, indicating that VCs are more likely to coinvest with similar peers. Demographic similarity appears particularly salient: VCs with greater representation of female or Asian partners are more likely to syndicate with others sharing these traits. This pattern may reflect preferences for in-group trust and collaboration, or alternatively, structural segmentation in an industry where informal networks shape access and opportunity.

Of particular interest for the IV strategy are the coefficients on professional and alumni networks. Both variables are positive and statistically significant across specifications, confirming that historical ties are predictive of current coinvestment behavior. In Columns (1) and (2), the presence of an alumni connection increases the probability of a coinvestment tie by approximately 5 percentage points, while a professional connection increases it by 13 percentage points. The stronger predictive power of professional ties holds in the continuous specifications as well: in Columns (3) and (4), each additional professional connection is associated with an increase of 0.03 coinvestments, compared to 0.015 for alumni connections. These results suggest that prior work experience plays a more substantial role than shared educational background in shaping current collaborative behavior among VCs.

Second step

Table 6 reports the second-stage estimation results using the IV strategy to address endogeneity in the baseline network model. As benchmarks, Column (1) presents OLS es-

26

Table 5: First step in the Heckman-corrected model: coinvestment network formation

	If coinvest		No. coinvestments	
	(1)	(2)	(3)	(4)
Professional connections	0.131***		0.0278***	
	(0.00125)		(0.000582)	
Alumni connections		0.0461***		0.0154***
		(0.00065)		(0.000214)
No. startups (absolute distance)	-0.00367**	-0.00349*	-0.013	-0.011
	(0.00178)	(0.00179)	(0.0134)	(0.0134)
Percent business & finance (absolute distance)	-0.0124***	-0.0128***	-0.0595***	-0.0556***
	(0.00106)	(0.00106)	(0.00799)	(0.00796)
Percent consumer G&S (absolute distance)	-0.0189***	-0.019***	-0.083***	-0.0864***
	(0.000877)	(0.000883)	(0.00662)	(0.0066)
Percent healthcare (absolute distance)	-0.0251***	-0.0242***	-0.107***	-0.0994***
	(0.000782)	(0.000788)	(0.0059)	(0.00588)
Percent information tech (absolute distance)	-0.0102***	-0.011***	-0.0416***	-0.041***
	(0.0013)	(0.00131)	(0.00981)	(0.00978)
Percent female (absolute distance)	-0.0182***	-0.0169***	-0.129***	-0.113***
	(0.000565)	(0.000575)	(0.00425)	(0.00424)
Percent Asian (absolute distance)	-0.015***	-0.0156***	-0.108***	-0.0946***
	(0.000572)	(0.000578)	(0.00429)	(0.00429)
Constant	0.052***	0.0461***	0.268***	0.241***
	(0.000456)	(0.000517)	(0.00332)	(0.00334)
Observations	448900	448900	448900	448900

Notes: Results from the estimation of equation (10). Columns (1) and (2) uses the binary outcome of coinvestment as the outcome variable. Columns (3) and (4) uses the number of coinvestments as the outcome variable. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels.

timates of exit rates on VC characteristics alone, and Column (2) replicates the baseline network model using coinvestment ties from Table 4. Columns (3) through (6) report estimates from equation (11), incorporating a control function term derived from the first-stage link formation model.

Across all specifications, the estimated network spillover effect δ remains positive and statistically significant, consistent with the presence of peer effects in VC performance. The magnitudes are similar to those in the baseline model, suggesting that the initial results are not driven by omitted variable bias. The coefficients on the control function term ξ, capturing unobserved individual-level factors affecting both performance and network formation, are statistically insignificant. This finding implies that the professional and

Table 6: Estimation results of the Heckman-correct network model

	(1) No networks	(2) Baseline model	(3) IV (professional, binary)	(4) IV (alumni, binary)	(5) IV (professional, count)	(6) IV (alumni, count)
			Dependent variable:			
			Exit rate			
δ (Social spillover)		0.00934***	0.00934***	0.00934***	0.00742***	0.00744***
		(0.00126)	(0.00126)	(0.00126)	(0.000394)	(0.000394)
No. startups	0.000997***	-0.00098***	-0.000979***	-0.00098***	-0.00466***	-0.00465***
	(0.000131)	(0.000323)	(0.000323)	(0.000323)	(0.000263)	(0.000262)
Percent business & finance	0.306***	0.265***	0.265***	0.266***	0.328***	0.332***
	(0.0927)	(0.0882)	(0.0884)	(0.0885)	(0.115)	(0.115)
Percent consumer G&S	0.19*	0.151	0.15	0.152	0.307**	0.309**
	(0.0993)	(0.0945)	(0.0947)	(0.0948)	(0.123)	(0.123)
Percent healthcare	0.425***	0.388***	0.388***	0.389***	0.155	0.159
	(0.0898)	(0.0845)	(0.0847)	(0.0847)	(0.108)	(0.108)
Percent information tech	0.405***	0.354***	0.353***	0.355***	0.361***	0.365***
	(0.0918)	(0.0873)	(0.0877)	(0.0877)	(0.113)	(0.114)
Percent female	0.0143	0.0198	0.0198	0.0198	0.0754	0.0748
	(0.0388)	(0.0374)	(0.0374)	(0.0374)	(0.0491)	(0.0491)
Percent Asian	-0.0163	-0.0212	-0.0211	-0.0212	0.0691*	0.0689*
	(0.0335)	(0.032)	(0.032)	(0.032)	(0.0405)	(0.0406)
Constant	0.047	0.037	0.0376	0.0361	0.0934	0.09
	(0.0826)	(0.0782)	(0.0785)	(0.0785)	(0.103)	(0.103)
Observations	670	670	670	670	670	670

Notes: Column (1) reports the OLS of exit rate on VC characteristics. Column (2) reports the estimates of equation (9). Columns (3) to (6) report the estimates of equation (11), with the professional networks and alumni networks, either binary or raw count, respectively. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels.

alumni networks used as instruments do not contain additional explanatory power beyond what is already captured by observed coinvestment ties.

5.3 Endogenous network formation

This section presents the estimation results from the endogenous network formation model. Table 7 and Table 8 report the median values of the posterior distributions. Table 7 summarizes the posterior medians for the key structural parameters of the network competitive equilibrium in equation (12), including the peer spillover parameter φ, the link formation elasticity λ, and the coefficients β on VC characteristics. Table 8 reports posterior medians of the parameters γ from the first-stage network formation equation (13).[8]

Before turning to the results, it is important to emphasize that both VC performance and network structure are jointly determined in equilibrium. The following interpreta-

---

[8]Instead of standard errors, the tables report empirical *p*-values for the null hypothesis that the parameter equals zero. These are computed as the proportion of posterior draws on the opposite side of zero from the posterior median. A *p*-value close to 0 or 1 indicates strong posterior support for a parameter being strictly negative or strictly positive, respectively. For instance, a *p*-value of 1 implies that the entire posterior support lies above zero, suggesting statistical significance at conventional levels.

tions therefore rely on the assumption that small changes in peer performance or network links have limited general equilibrium effects, i.e., they do not meaningfully alter the broader network architecture.

The analysis begins with the parameter φ in equation (12), which captures the strength of peer spillovers in the endogenous network setting. The estimate is positive and statistically significant, consistent with the presence of social externalities in performance. While the magnitude of φ is not directly interpretable due to the nonlinear structure of the model—specifically, the dependence on $P_j^{1+\lambda}$—the implied effects are economically meaningful. For example, an increase in the exit rate of a connected peer from 10% to 20% raises a VC's own expected exit rate by approximately 0.09 percentage points. Similarly, forming a new connection with a peer whose exit rate is 10% increases the VC's own performance by roughly 0.08 percentage points.

The parameter λ admits a more direct interpretation given the structure of the model. As shown in equation (8), λ represents the elasticity of connection intensity g_{ij} with respect to peer performance P_j. That is, a 1% increase in a peer's exit rate leads to a 0.74 percentage point increase in the intensity of the connection to that peer, holding all else constant and assuming negligible general equilibrium feedback. This result suggests that VCs are highly responsive to peer quality and strategically adjust their networks to strengthen ties with more effective partners.

Table 7: Results from the endogenous network model

	Dependent variable:
	Exit rate
φ (Social spillover)†	0.0002***
	[1.0000]
λ (Elasticity of network formation)†	0.7411***
	[1.0000]
No. startups	0.0010***
	[1.0000]
Percent business & finance	0.3395***
	[1.0000]
Percent consumer G&S	0.2165***
	[1.0000]
Percent healthcare	0.4897***
	[1.0000]
Percent information tech	0.4705***
	[1.0000]
Percent female	0.0148***
	[1.0000]
Percent Asian	-0.0232***
	[0.0000]
Pseudo-R^2	0.8352
Penalized pseudo-R^2	0.8341
MSE	0.1648
MASD	0.4320
Observations	670

Notes: λ is the elasticity of link g_{ij} with respect to the performance of j, E_j. φ is calculated based on the estimates of ρ, α, and λ. Estimates of parameters in equation (12) are reported in column (1). The median of the posterior distribution estimated with the ABC algorithm is reported for each parameter. The empirical p-value of zero on the estimated posterior is reported in the brackets. p-value is equal to 1 if the support of the empirical posterior distribution is greater than zero, whereas p-value is equal to 0 if the support of the empirical posterior distribution is less than zero. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels based on empirical p-values.

Table 8 reports the estimated determinants of social connections in the VC industry based on the posterior medians from the first-stage network formation model. Consistent with prior results, there is strong evidence of homophily: VCs exhibit a pronounced tendency to form ties with peers who share similar demographic characteristics and industry specializations.

One notable exception is the positive coefficient on the distance in the number of supported startups, suggesting that VCs may be more likely to form connections with partners of different fund sizes. This result points to a potential complementarity in network formation. VCs may seek to diversify their information sets or mitigate risk by engaging with firms that differ in scale. Smaller VCs might benefit from the reach and experience of larger funds, while larger VCs may gain access to niche expertise or localized knowledge from smaller peers. These heterogeneous connections could enhance the value of social ties beyond what homophilous relationships alone can offer.

A key finding from Table 8 is the positive and statistically significant coefficient on past connections. While the magnitude is not directly interpretable due to the logistic specification in equation (13), the direction and significance of the estimate are noteworthy. Importantly, these parameters are inferred from VC performance data within the structural model, rather than being derived from observed coinvestment networks. This alignment with intuitive expectations underscores the model's capacity to recover meaningful social structures from performance outcomes alone. Notably, the social networks inferred from the model differ from the observed coinvestment networks, suggesting that the model captures latent relational dynamics not immediately evident in direct investment ties. This distinction opens avenues for further research into the nature and implications of these latent social networks in the VC industry.

To place the endogenous network formation model in context, I compare it with two benchmark specifications. The first is a benchmark without network effects, in which VC performance depends solely on observable characteristics. This is equivalent to imposing $\rho = 0$ in the production function (3), and consequently $\varphi = 0$ in equation (12). The second benchmark allows performance to depend on networks, but treats connections as exogenously given, i.e., VCs do not choose their links strategically. This corresponds to setting the network formation elasticity $\lambda = 0$, such that g_{ij} becomes equal to θ_{ij} in equation (8), and equation (12) reduces to the baseline exogenous network model in equation (9).

Several key findings emerge. First, the estimated social spillover parameter φ is positive and statistically significant in both the exogenous and endogenous models, confirm-

31

Table 8: Results of link formation in the endogenous network model

	Dependent variable:
	Compatibility
Professional connection	1.3400***
	[1.0000]
No. startups (absolute distance)	0.0039***
	[1.0000]
Percent business & finance (absolute distance)	-4.1258***
	[0.0000]
Percent consumer G&S (absolute distance)	-3.1104***
	[0.0000]
Percent healthcare (absolute distance)	-0.8625***
	[0.0000]
Percent information tech (absolute distance)	-1.9955***
	[0.0000]
Percent female (absolute distance)	-0.1731***
	[0.0000]
Percent Asian (absolute distance)	-0.3480***
	[0.0000]
Constant	-1.8462***
	[0.0000]
Observations	448,900

Notes: Estimates of parameters in equation (13) are reported in column (1). The median of the posterior distribution estimated with the ABC algorithm is reported for each parameter. The empirical p-value of zero on the estimated posterior is reported in the brackets. p-value is equal to 1 if the support of the empirical posterior distribution is greater than zero, whereas p-value is equal to 0 if the support of the empirical posterior distribution is less than zero. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels based on empirical p-values.

ing that peer performance influences a VC's own success. However, the magnitude is notably smaller in the endogenous case, likely reflecting the model's adjustment for strategic link formation. Second, the elasticity of network formation λ is estimated at 0.7411 and is highly significant, indicating that VCs actively respond to peer quality when forming social connections. Across all models, fund size and industry specialization are strong predictors of exit performance, particularly in healthcare and information technology. Interestingly, the coefficient on percent Asian becomes increasingly negative and significant as network structure is more fully modeled, suggesting potential segmentation in network-driven access to high-quality deals.

Table 9: Comparison between the main estimation and two benchmarks

		Dependent variable:	
		Exit rate	
	(1)	(2)	(3)
	No networks	Exogenous networks	Endogenous networks
φ (Social spillover)†	-	0.0012***	0.0002***
		[1.0000]	[1.0000]
λ (Elasticity of network formation)†	-	-	0.7411***
			[1.0000]
No. startups	0.0010***	0.0011***	0.0010***
	[1.0000]	[1.0000]	[1.0000]
Percent business & finance	0.3539***	0.3300***	0.3395***
	[1.0000]	[1.0000]	[1.0000]
Percent consumer G&S	0.2403***	0.2298***	0.2165***
	[1.0000]	[1.0000]	[1.0000]
Percent healthcare	0.4730***	0.4323***	0.4897***
	[1.0000]	[1.0000]	[1.0000]
Percent information tech	0.4546***	0.4581***	0.4705***
	[1.0000]	[1.0000]	[1.0000]
Percent female	0.0154	0.0108***	0.0148***
	[0.6571]	[1.0000]	[1.0000]
Percent Asian	-0.0158	-0.0172***	-0.0232***
	[0.3185]	[0.0000]	[0.0000]
Observations			

Notes: λ is the elasticity of link g_{ij} with respect to the performance of j, E_j. φ is calculated based on the estimates of ρ, α, and λ. Estimates of parameters in equation (7) are reported in column (3). Column (1) reports the estimates with the constraint $\lambda = 0$. Column (2) reports the estimates with the constraint $\rho = 0$. The median of the posterior distribution estimated with the ABC algorithm is reported for each parameter. The empirical p-value of zero on the estimated posterior is reported in the brackets. p-value is equal to 1 if the support of the empirical posterior distribution is greater than zero, whereas p-value is equal to 0 if the support of the empirical posterior distribution is less than zero. *, **, and *** indicates statistical significance at the 10, 5, and 1% levels based on empirical p-values.

6 Conclusion

The venture capital industry operates in an environment defined by high uncertainty, asymmetric information, and limited transparency. In such settings, networks, both formal and informal, play a critical role in reducing informational frictions, shaping investment decisions, and influencing performance outcomes. While it is widely acknowledged that networks matter in VC, most empirical work has focused on observed coinvestment ties using reduced-form methods, often leaving unresolved the underlying endogeneity

and the role of latent social relationships.

This paper introduces a structural approach to studying VC networks, offering a unified framework that links network formation, information flow, and fund performance. To the best of my knowledge, this is the first study to estimate VC networks using a structural equilibrium model rooted in microeconomic foundations. The results provide robust evidence that VCs with stronger connections to high-performing peers achieve better outcomes, measured by the proportion of successful portfolio exits. More significantly, the analysis shows that much of this effect can be attributed to unobserved social connections that are not captured by formal coinvestment data.

The paper makes three main methodological contributions. First, I develop a micro-founded production model that connects peer performance to own performance through information diffusion, grounding the analysis in theories of financial intermediation and organizational learning. Second, I use historical professional and alumni networks to construct instruments that help address endogeneity in link formation. These results highlight the persistent influence of background affiliations and suggest that the VC industry is shaped by relationship-driven dynamics that may limit access for outsiders. Third, I propose an endogenous network formation model that recovers latent social networks directly from performance outcomes, past affiliations, and firm-level characteristics. This final contribution offers a novel way to study informal networks and shows that the recovered social structure, while overlapping with observed coinvestments, contains meaningful and distinct differences.

Taken together, the findings demonstrate that VC success depends not only on capital and skill, but also on access to and integration within the right networks. Informal ties and shared histories can be just as influential as formal syndication partnerships in determining who gets access to deals and resources. Future work could build on this framework by examining the dynamics of network evolution, the role of geographic and sectoral clustering, or the interaction between social capital and innovation outcomes. Understanding these forces can help explain the deeper social architecture that drives performance in entrepreneurial finance.

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012). "The Network Origins of Aggregate Fluctuations". *Econometrica* 80.5, pp. 1977–2016.

Allen, Franklin and Ana Babus (2009). "Networks in Finance". *The Network Challenge: Strategy, Profit, and Risk in an Interlinked World*. Vol. 367.

Battaglini, Marco, Eleonora Patacchini, and Edoardo Rainone (2021). "Endogenous Social Interactions with Unobserved Networks". *The Review of Economic Studies*, rdab058.

Battaglini, Marco, Valerio Leone Sciabolazza, and Eleonora Patacchini (2020). "Effectiveness of Connected Legislators". *American Journal of Political Science* 64.4, pp. 739–756.

Bonacich, Phillip and Paulette Lloyd (2001). "Eigenvector-like Measures of Centrality for Asymmetric Relations". *Social Networks* 23.3, pp. 191–201.

Brander, James A., Raphael Amit, and Werner Antweiler (2002). "Venture-Capital Syndication: Improved Venture Selection vs. The Value-Added Hypothesis". *Journal of Economics & Management Strategy* 11.3, pp. 423–452.

Bubna, Amit, Sanjiv R. Das, and Nagpurnanand Prabhala (2020). "Venture Capital Communities". *Journal of Financial and Quantitative Analysis* 55.2, pp. 621–651.

Cochrane, John H. (2005). "The Risk and Return of Venture Capital". *Journal of Financial Economics* 75.1, pp. 3–52.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy (2008). "The Small World of Investing: Board Connections and Mutual Fund Returns". *Journal of Political Economy* 116.5, pp. 951–979.

Da Rin, Marco, Thomas Hellmann, and Manju Puri (2013). "A Survey of Venture Capital Research". *Handbook of the Economics of Finance*. Ed. by George M. Constantinides, Milton Harris, and Rene M. Stulz. Vol. 2. Elsevier, pp. 573–648.

Das, Sanjiv R., Hoje Jo, and Yongtae Kim (2011). "Polishing Diamonds in the Rough: The Sources of Syndicated Venture Performance". *Journal of Financial Intermediation* 20.2, pp. 199–230.

Du, Qianqian (2016). "Birds of a Feather or Celebrating Differences? The Formation and Impacts of Venture Capital Syndication". *Journal of Empirical Finance* 39, pp. 1–14.

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson (2014). "Financial Networks and Contagion". *American Economic Review* 104.10, pp. 3115–3153.

Engelberg, Joseph, Pengjie Gao, and Christopher A. Parsons (2012). "Friends with Money". *Journal of Financial Economics* 103.1, pp. 169–188.

Ewens, Michael (2023). "Gender and Race in Entrepreneurial Finance". *Handbook of the Economics of Corporate Finance*. Ed. by B. Espen Eckbo, Gordon M. Phillips, and Morten

Sorensen. Vol. 1. Private Equity and Entrepreneurial Finance. North-Holland, pp. 239–296.

Graham, Bryan S. (2020). "Network Data". *Handbook of Econometrics*. Ed. by Steven N. Durlauf, Lars Peter Hansen, James J. Heckman, and Rosa L. Matzkin. Vol. 7. Handbook of Econometrics, Volume 7A. Elsevier, pp. 111–218.

Hastings, W. K. (1970). "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". *Biometrika* 57.1, pp. 97–109.

Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu (2007). "Whom You Know Matters: Venture Capital Networks and Investment Performance". *The Journal of Finance* 62.1, pp. 251–301.

Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu (2010). "Networking as a Barrier to Entry and the Competitive Supply of Venture Capital". *The Journal of Finance* 65.3, pp. 829–859.

Huang, Can (2022). *Networks in Venture Capital Markets*. SSRN Scholarly Paper. Rochester, NY.

Kaplan, Steven N. and Antoinette Schoar (2005). "Private Equity Performance: Returns, Persistence, and Capital Flows". *The Journal of Finance* 60.4, pp. 1791–1823.

Katz, Leo (1953). "A New Status Index Derived from Sociometric Analysis". *Psychometrika* 18.1, pp. 39–43.

Lerner, Josh, Hilary Shane, and Alexander Tsai (2003). "Do Equity Financing Cycles Matter? Evidence from Biotechnology Alliances". *Journal of Financial Economics* 67.3, pp. 411–446.

Lerner, Joshua (1994). "The Syndication of Venture Capital Investments". *Financial Management* 23.3, pp. 16–27.

Lewbel, Arthur, Xi Qu, and Xun Tang (2023). "Social Networks with Unobserved Links". *Journal of Political Economy* 131.4, pp. 898–946.

Lindsey, Laura (2008). "Blurring Firm Boundaries: The Role of Venture Capital in Strategic Alliances". *The Journal of Finance* 63.3, pp. 1137–1168.

Marjoram, Paul, John Molitor, Vincent Plagnol, and Simon Tavaré (2003). "Markov Chain Monte Carlo without Likelihoods". *Proceedings of the National Academy of Sciences* 100.26, pp. 15324–15328.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller (1953). "Equation of State Calculations by Fast Computing Machines". *The Journal of Chemical Physics* 21.6, pp. 1087–1092.

Nahata, Rajarishi (2008). "Venture Capital Reputation and Investment Performance". *Journal of Financial Economics* 90.2, pp. 127–151.

Rider, Christopher I. (2012). "How Employees' Prior Affiliations Constrain Organizational Network Change: A Study of U.S. Venture Capital and Private Equity". *Administrative Science Quarterly* 57.3, pp. 453–483.

Shue, Kelly (2013). "Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers". *The Review of Financial Studies* 26.6, pp. 1401–1442.

Sørensen, Morten (2007). "How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital". *The Journal of Finance* 62.6, pp. 2725–2762.

Sorenson, Olav and Toby E. Stuart (2001). "Syndication Networks and the Spatial Distribution of Venture Capital Investments". *American Journal of Sociology* 106.6, pp. 1546–1588.

Sorenson, Olav and Toby E. Stuart (2008). "Bringing the Context Back In: Settings and the Search for Syndicate Partners in Venture Capital Investment Networks". *Administrative Science Quarterly* 53.2, pp. 266–294.

Tian, Xuan (2011). "The Causes and Consequences of Venture Capital Stage Financing". *Journal of Financial Economics* 101.1, pp. 132–159.

Appendix A   Details of the Structural Model

This section reproduces results in Battaglini, Sciabolazza, and Patacchini (2020) and Battaglini, Patacchini, and Rainone (2021).

Exogenous network equilibrium

The production function for VC performance is given by

$$P_i = \rho s_i^\alpha l_i^{1-\alpha} + \varepsilon_i, \tag{A.1}$$

where s_i denotes the social connectedness of VC i, defined as

$$s_i = \sum_{j \in \mathcal{N}} g_{ij} P_j. \tag{A.2}$$

g_{ij} is the intensity of the social link between VCs i and j, and P_j represents the performance of peer j.[9]

Each VC chooses effort l_i to maximize performance net of effort cost:

$$\max_{l_i} \rho s_i^\alpha l_i^{1-\alpha} + \varepsilon_i - l_i. \tag{A.3}$$

The first-order condition yields the optimal effort level:[10]

$$l_i^* = \left(\rho(1-\alpha)\right)^{\frac{1}{\alpha}} s_i. \tag{A.4}$$

Substituting this optimal choice back into the production function gives the equilibrium performance:

$$P_i^* = \delta \sum_{j \in \mathcal{N}} g_{ij} P_j^* + \varepsilon_i, \tag{A.5}$$

---

[9]The model imposes the following parameter restrictions. Effort is bounded such that $l_i \in [0, \bar{l}]$ for some $\bar{l} > 0$, and the cost of effort is normalized to l_i. Link intensity is similarly bounded, with $g_{ij} \in [0, \bar{g}]$ for some $\bar{g} > 0$, and self-connections are ruled out by assumption, i.e., $g_{ii} = 0$ for all i. Individual heterogeneity enters additively through ε_i, which is assumed to lie in the interval $[\underline{\varepsilon}, \bar{\varepsilon}]$ with $\underline{\varepsilon} > 0$ and $\bar{\varepsilon} \in (0, 1)$. Finally, assume that $\rho \bar{g}^\alpha \bar{l}^{1-\alpha} + \bar{\varepsilon} < 1$. This provides a sufficient condition that guarantees $P_i \in (0, 1)$.

[10]Assume that $\bar{l} > \left((1-\alpha)\rho\right)^{\frac{1}{\alpha}}$. This guarantees interior solutions of $l_i < \bar{l}$.

where $\delta = \rho^{\frac{1}{\alpha}}(1-\alpha)^{\frac{1-\alpha}{\alpha}}$ is a reduced-form parameter that summarizes the strength of social spillovers. Because the system in (A.5) is linear in P, it admits a unique closed-form solution. Letting $\mathbf{G}$ denote the matrix of link intensities and ε the vector of individual shocks, the equilibrium is given by

$$P(\mathbf{G}, \varepsilon; \delta) = [\mathbf{I} - \delta\mathbf{G}]^{-1}\varepsilon. \tag{A.6}$$

Exogenous network equilibrium

The cost of establishing a social connection between VCs is modeled by the following functional form:[11]

$$c(g_{ij}, \theta_{ij}; \lambda) = \frac{\lambda}{1+\lambda}\left(\frac{g_{ij}}{\theta_{ij}}\right)^{1+\frac{1}{\lambda}}, \tag{A.7}$$

where g_{ij} is the intensity of the social connection from i to j, θ_{ij} captures compatibility or ease of forming the link, and $\lambda > 0$ governs the curvature of the cost function. The parameter λ thus determines the elasticity of connection formation with respect to peer performance and plays a key role in shaping equilibrium link choices.

The model is set in two periods. In period 1, VCs choose their network connections; in period 2, they select effort levels conditional on the realized network. Each VC is characterized by a type $\omega_i = (\varepsilon_i, (\theta_{ij})_j, \mathcal{M}_i)$, where ε_i represents idiosyncratic heterogeneity, θ_{ij} denotes compatibility with each potential peer VC j, and $\mathcal{M}_i$ is the set of VCs such that $\theta_{ij} > 0$. Let Ω denote the space of types.

A strategy profile consists of a pair of functions (g, l). The connection strategy $g : \Omega \to [0, \bar{g}]^{n-1}$ maps each VC's type to a vector of connection intensities, specifying how strongly they link to each other peer. The effort strategy $l : \Omega \times G \to [0, \bar{l}]$ maps each VC's type and the realized network G to an effort level in period 2. A pure-strategy equilibrium is defined as a fixed point (g, l) in which all VCs optimize given their expectations over peer performance, network structure, and the cost of forming and maintaining social connections.

---

[11]The cost of link formation c_{ij} is incurred solely by VC i, implying an asymmetric cost structure. That is, c_{ij} is borne by i alone, while c_{ji} is borne by j. This assumption simplifies the exposition and can be generalized to a symmetric or shared-cost formulation without affecting the core results.

We solve the game by backward induction. In period 2, VC i chooses its own effort l_i to maximize its performance net of effort cost. This problem is identical to the baseline model analyzed earlier, with equilibrium performance P^* determined by the autoregressive system in equation (A.5). Ignoring discounting and substituting the period-2 optimal effort into the production function, the continuation value for VC i is given by

$$P_i^*(\mathbf{G}, \varepsilon) - l_i^*(\mathbf{G}, \varepsilon) = \alpha\delta \sum_{j \in \mathcal{N}} g_{ij} P_j^*(\mathbf{G}, \varepsilon) + \varepsilon_i. \tag{A.8}$$

In period 1, VC i chooses its connections $\mathbf{g}_i = (g_{i1}, \ldots g_{in})$ to maximize its expected continuation value net of connection costs. Using the parametric cost function from equation (A.7), the link formation problem becomes

$$\max_{\mathbf{g}_i} \sum_{j \in \mathcal{N}} \left(\alpha\delta g_{ij} P_j^*(\mathbf{G}, \varepsilon) + \varepsilon_i - \frac{\lambda}{1+\lambda} \left(\frac{g_{ij}}{\theta_{ij}} \right)^{1+\frac{1}{\lambda}} \right). \tag{A.9}$$

The first-order condition of equation (A.9) yields the following characterization:

$$g_i^* \leq \theta_{ij}^{1+\lambda} (\alpha\delta P_j^*)^\lambda. \tag{A.10}$$

Together, equations (A.5) and (A.10) characterize the network competitive equilibrium $(l^*, P^*, \mathbf{G}^*)$. If an interior solution exists, then the two conditions collapse to the following system:

$$P_i^* = \varphi \sum_{j \in \mathcal{N}} (\theta_{ij} P_j^*)^{1+\lambda} + \varepsilon_i, \tag{A.11}$$

where $\varphi = \alpha^\lambda \delta^{1+\lambda}$. In other words, the equilibrium performance P^* is characterized by a system of nonlinear equations.[12]

Control function approach

To account for selection bias using the control function approach in the second stage, assume the unobserved components (ε, η) have the following joint distribution. $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)'$ and $\eta_i = (\eta_{i1}, \cdots, \eta_{in})'$ are jointly normally distributed with mean zero. The covariance has the following structure: $E(\varepsilon_i^2) = \sigma_\varepsilon^2$, $E(\eta_{ij}^2) = \sigma_\eta^2$, $E(\varepsilon_i\eta_{ij}) = \sigma_{\varepsilon\eta}$ for all $i \neq j$, and

---

[12] Assume that $\bar{g} > (\alpha\delta)^\lambda \bar{\theta}^{1+\lambda}$, where $\bar{\theta} = \max \theta_{ij}$. If $\delta \leq \frac{1}{\bar{\theta}} \left(\frac{1}{(1+\lambda)\alpha^\lambda \bar{m}} \right)^{\frac{1}{1+\lambda}}$, then the equilibrium is unique.

$E(\eta_{ij}\eta_{ik}) = 0$ for all $j \neq k$. Under these assumptions, the expected value of the second-stage error conditional on the first-stage residuals is given by $E(\varepsilon_i|\eta_{i1}, \cdots, \eta_{in}) = \psi \sum_{j \neq i} \eta_{ij}$, where $\psi = \sigma_{\varepsilon\eta}/\sigma_{\eta}^2$. Incorporating this selection term yields the corrected model:

$$P = \delta GP + X\beta + \psi\xi + \varepsilon, \tag{A.12}$$

where $\xi_i = \sum_{j \neq i} \eta_{ij}$ captures unobserved factors influencing the likelihood of forming links.