Nested Syndication Networks: Community Structure and Hierarchical Organization in Venture Capital Ecosystems

Disclaimer: this is a intermediary article over the results gotten until now. Colored rectangles represent "to do" or "ongoing" activities to be done in subsequent work until the final version of this article is complete.

Abstract

This study analyzes venture capital syndication networks using network theory to understand how investors organize and collaborate in funding startups. We examine 104,618 investment records from Crunchbase involving 38,843 investors and 16,932 companies to identify structural patterns in investor communities.

Using greedy modularity optimization, we discovered 170 distinct investor communities, with three large communities containing over 12,000 investors (75% of the network). These communities exhibit power-law degree distributions typical of scale-free networks but differ significantly in investment activity and organization.

Our key finding is the emergence of significantly nested structures within Community 2, which is dominated by Silicon Valley investors. This hierarchical organization enables substantially higher transaction volumes (33.6% of all investments) compared to similarly-sized communities. The nested structure suggests that informal investment hierarchies systematically influence funding accessibility for entrepreneurs.

The concentration of nested structures specifically within Silicon Valley indicates that geographic clustering in innovation hubs may facilitate hierarchical investor relationships. High-degree investors in the nested community, particularly early-stage firms like SV Angel, function as network hubs that create efficient capital allocation pathways.

These findings challenge assumptions about random mixing in venture capital markets and suggest that network topology, rather than community size alone, determines investment ecosystem efficiency. The hierarchical organization may enhance stability against random investor departures but creates vulnerability to targeted removal of highly connected nodes.

Our results provide insights for ecosystem development strategies and suggest that understanding network structure is crucial for predicting investment patterns and improving access to capital for entrepreneurs.

Contents

1 Introduction

Improve introduction, it is still embryonic

Venture capital syndication, where two or more firms co-invest in the same enterprise, represents a fundamental mechanism in innovation financing [5]. This collaborative behavior emerges as investors seek to reduce risks associated with early-stage companies that lack complete validation or functional products. In such contexts, reputation and network centrality influence investment decisions, as venture capitalists use their co-investors' characteristics as signals that affect their own investment choices.

The increasing prevalence of syndicated investments in recent decades demonstrates that innovation networks are highly socialized systems where agents do not act in isolation but rather form communities influenced by their peers. This phenomenon reflects the concept of embeddedness, where economic actions are shaped by social structures and relationships [5].

Extensive literature documents how heavy-tailed degree distributions emerge from such interactions in social networks. Innovation networks follow similar patterns, particularly regarding investor connectivity, where concentrated hubs of highly connected agents coexist with a majority of participants having few connections. This heterogeneous distribution of connectivity follows a power-law degree distribution pattern.

The widespread presence of power-law degree distributions has motivated numerous studies investigating the mechanisms behind their emergence and their impact on processes such as spreading dynamics [8] and network robustness [1].

Regarding spreading dynamics, social and economic scientists have explored how novel ideas propagate through networks and how bridge edges with high betweenness centrality affect the likelihood of innovation diffusion. Established theories such as the Strength of Weak Ties [6] provide theoretical foundations for these analyses.

In the other hand, measuring robustness in social and innovation networks remains both a theoretical and practical challenge. Ecological network analysis has provided insights for addressing this challenge, with metrics such as nestedness and their ecological consequences being applied to social networks [10].

This paper applies network theory to represent syndicated investments as edges in a network where investors serve as nodes with diverse characteristics (geographic, financial, sectoral). This mathematical representation enables visualization and interpretation of syndication network structures through both ecological and economic perspectives.

Particular attention is given to the observation of nestedness within a large community of early and late-stage investors in California. The implications and potential consequences of this phenomenon are explored in the *Discussion and Implications* section of this paper.

2 Methodology

2.1 Data Source and Preprocessing

This study uses data from Crunchbase, a comprehensive database containing information about startups, venture capital firms, and investment rounds. The dataset includes information about companies, investors, investments, and funding rounds in the United States market. International venture capital firms from other countries also appear in the dataset when they participate in US startup investments.

The data preprocessing follows established methodologies from entrepreneurship literature [4]. The cleaning process includes several steps: (1) removal of companies with incomplete information, (2) exclusion of companies founded after 2017 to allow sufficient time for investment patterns to emerge, (3) removal of companies with exit status (bankruptcy, acquisition, or IPO), and (4) application of a minimum funding threshold of $150,000 to focus on substantive investment relationships.

2.2 Investment Network Construction

The analysis focuses on venture capital co-investment patterns across different funding stages. Investment stages are categorized into two main groups:

- Early stages: angel, pre-seed, seed, and Series A
- Late stages: Series B through Series I

A bipartite network is constructed where nodes represent venture capital firms and edges represent co-investment relationships in the same company. The network is bipartite because it connects two distinct sets of investors: those participating in early-stage rounds (right nodes) and those participating in late-stage rounds (left nodes).

This approach allows us to study how early-stage and late-stage investors interact in the investment ecosystem.

The bipartite graph $G = (U \cup V, E)$ consists of:

$$U = \{u_1, u_2, \ldots, u_m\} \text{ (late-stage VCs)} \tag{1}$$

$$V = \{v_1, v_2, \ldots, v_n\} \text{ (early-stage VCs)} \tag{2}$$

$$E \subseteq U \times V \text{ (co-investment relationships)} \tag{3}$$

To prevent spurious connections from related entities, investor pairs where the first five characters of their names match are filtered out, reducing the likelihood of including different funds from the same parent organization. Furthermore, investors that participated in both early and late stages receive a suffix so they can be treated as distinct agents for each phase.

Clearly show the overlap or number of connections made between the same investors but in distinct phases ex. VC1_serieA-VC1_serieC

2.3 Community Detection

Community structure in the bipartite network is identified using the greedy modularity optimization algorithm [3]. This method iteratively merges communities to maximize the modularity score, which measures the density of connections within communities compared to connections between communities.

For a bipartite network, modularity Q is defined as:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \tag{4}$$

where A_{ij} is the adjacency matrix, k_i is the degree of node i, m is the total number of edges, c_i is the community of node i, and $\delta(c_i, c_j)$ is 1 if nodes i and j are in the same community, 0 otherwise.

The algorithm identifies communities of venture capital firms that frequently co-invest together, revealing structural patterns in the investment ecosystem that may not be apparent from individual investment decisions.

2.4 Nestedness Analysis

Nestedness is a structural property commonly observed in ecological networks [2] that describes the tendency for specialists to interact with a subset of the partners of generalists. In the context of venture capital networks, nestedness would indicate that investors with fewer connections tend to co-invest with a subset of the partners of more connected investors.

We measure nestedness using the NODF (Nestedness based on Overlap and Decreasing Fill) metric [2]. For a bipartite adjacency matrix M with rows and columns sorted by decreasing degree, NODF is calculated as:

$$NODF = \frac{NODF_{rows} + NODF_{columns}}{2} \tag{5}$$

where:

$$NODF_{rows} = \frac{100}{R(R-1)/2} \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \frac{|N_i \cap N_j|}{k_j} \text{ if } k_i > k_j \tag{6}$$

$$NODF_{columns} = \frac{100}{C(C-1)/2} \sum_{i=1}^{C-1} \sum_{j=i+1}^{C} \frac{|N_i \cap N_j|}{k_j} \text{ if } k_i > k_j \tag{7}$$

Here, R and C are the number of rows and columns, N_i represents the set of connections for node i, and k_i is the degree of node i.

Using this method, NODF values range between 0 and 1 (perfect nestedness).

2.5 Statistical Significance Testing

To determine whether observed nestedness values are significantly higher than expected by chance, we employ a null model approach using the Curveball algorithm [9]. This algorithm generates randomized matrices that preserve the degree sequence of both node sets while randomizing the connection patterns.

For each community, we generate 100 null matrices using 10,000 Curveball iterations. The statistical significance is assessed by comparing the observed NODF score against the distribution of null model scores:

Generate 1000 null matrices instead

The standardized Z-score is calculated to quantify how many standard deviations the observed nestedness differs from the null expectation:

$$Z = \frac{NODF_{observed} - \mu_{null}}{\sigma_{null}} \tag{8}$$

where μ_{null} and σ_{null} are the mean and standard deviation of the null distribution, respectively.

The p-value is calculated empirically from the null distribution as the proportion of randomized matrices that exhibit nestedness equal to or greater than the observed value:

$$p = \frac{1 + \sum_{i=1}^{N} I(NODF_{null,i} \geq NODF_{observed})}{N+1} \tag{9}$$

where N is the number of null matrices (100 in our case), $NODF_{null,i}$ is the nestedness score of the i-th null matrix, and $I(\cdot)$ is an indicator function that equals 1 when the condition is true and 0 otherwise. The addition of 1 in both numerator and denominator provides a conservative estimate that avoids p-values of exactly zero.

The p-value represents the probability of observing nestedness as high as or higher than the observed value under the null hypothesis of random co-investment patterns. Communities with $p < 0.05$ are considered to have significantly high nestedness, indicating that the observed nested structure is unlikely to have arisen by chance alone.

While both Z-scores and p-values assess statistical significance, they provide complementary information: the Z-score quantifies the magnitude of deviation from the null expectation in standardized units, while the p-value provides the probability of observing such deviation under the null hypothesis. In our analysis, we primarily rely on p-values for significance testing as they directly quantify the statistical evidence against the null hypothesis of random network structure.

3 Results

3.1 Data Mining and Clustering Overview

The Crunchbase dataset, following the cleaning processes described in the *Methodology* section, yields 147,832 investment records, representing transactions among 22,527 companies and 38,843 investors.

Exclusion of non-venture capital investors reduces the dataset to 104,618 investment records and 16,932 unique companies with venture capital funding.

The division of venture capital firms into early-stage and late-stage investor groups results in 169,679 investment pairs comprising 3,666 unique startups.

Add explanation on why investment pairs are higher than investment records

Add network visualization showing bipartite structure

Community detection using greedy modularity optimization identifies approximately 170 distinct communities (the number of communities oscillates between 167 and 175 across different trials), with the largest communities containing over 4000 investors each, followed by 1 community with almost 1000 agents, 4 communities with more than 100 agents, and then several smaller groups, as summarized in Table 1.

Initially, analysis focuses on communities with at least 150 nodes to ensure statistical power for nestedness analysis. This threshold yields 5 communities.

Add rationale for threshold

The largest three communities (0, 1, and 2) contain over 12,000 investors combined, representing approximately 75% of all investors in the network. This concentration suggests a highly centralized structure within the venture capital ecosystem, with most investment activity occurring within a small number of large communities [3].

Community ID	Number of Investors
0	4,248
1	4,089
2	3,959
3	979
4	188
5	155
6	137
7	122

Table 1: Size distribution of the largest investor communities identified through greedy modularity optimization

Mention literature, as this phenomenon is well-documented

Add figure of community size distribution

3.2 Communities Characterization

Community boundaries are defined at the node level, meaning each investor belongs to exactly one community. However, edges (investment relationships) can span community boundaries when investors from different communities co-invest in the same startup.

To analyze investment patterns, we classify each syndicated investment as either: (1) intra-community if all participating investors belong to the same community, or (2) cross-community if investors from multiple communities participate together. Table 2 presents the resulting investment distribution across communities.

The analysis reveals important patterns in investment activity distribution. Community 2 accounts for the largest share of investments (33.6%), containing approximately 50% more investments than Community 0 and over three times more than Community 1. This concentration of investment activity suggests that structural features of Community 2 may facilitate higher transaction volumes. Notably, cross-community investments represent over

Community	**Co-investments Pairs**	**Relative Proportion**
Community 0	32,164	19.4%
Community 1	17,301	10.4%
Community 2	55,863	33.6%
Cross-community	58,329	35.1%

Table 2: Distribution of syndicated investments across investor communities

one-third of all transactions, indicating substantial interconnectedness across community boundaries.

3.2.1 Degree Distribution

Before further examining community-level investment patterns in-depth, we analyze the degree distribution characteristics across the three largest investor communities. This analysis reveals that all communities exhibit power-law patterns typical of scale-free networks, with notable differences in magnitude and scale parameters.

Communities 0 and 2 demonstrate remarkably similar degree distribution magnitudes, while Community 1 exhibits consistently lower magnitude values across all degree ranges. This pattern becomes particularly evident when examining the degree distributions on a logarithmic scale, where Community 2 shows the highest volume of nodes across all degree ranges, suggesting greater overall connectivity within this community.

Compare connectance among communities

Analysis of high-degree nodes (95th percentile) reveals distinct patterns across communities. Community 0 demonstrates a mixed composition of hub nodes, with early-stage investors like Techstars-seed (degree: 582) and 500 Global-seed (degree: 564) dominating the highest positions, alongside significant late-stage players such as Gaingels-series_b (degree: 509).

Community 1 exhibits a more balanced distribution with Intel Capital maintaining strong presence across multiple investment stages, while Com-

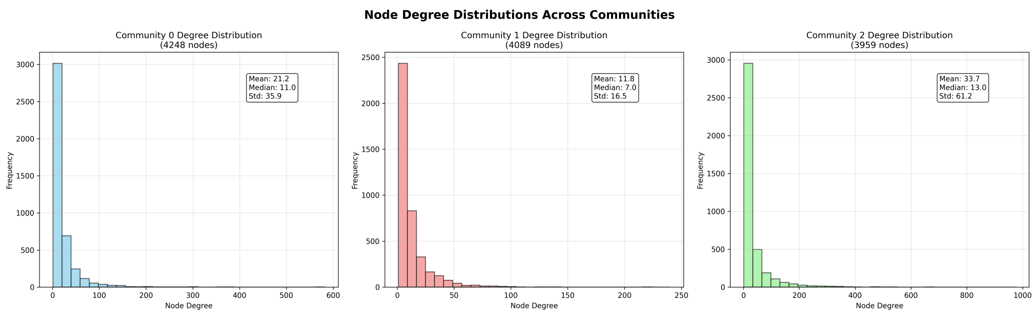


Figure 1: Node degree distributions for the three largest investor communities. The figure shows both the individual degree distributions and their overlap on a logarithmic scale. Communities 0 and 2 display similar heavy-tailed, power-law-like patterns, while Community 1 has consistently lower degree magnitudes. The log-scale overlay highlights that Community 2 maintains the highest node volume across all degree ranges, indicating greater overall connectivity.

munity 2 shows remarkable concentration among early-stage Silicon Valley investors, with SV Angel-seed achieving the highest connectivity (degree: 974) followed by other prominent early-stage firms including Andreessen Horowitz and Khosla Ventures.

Table 3 presents the top high-degree nodes for each community, highlighting the structural differences in hub organization.

As per Figure 3, the relationship between degree centrality and investment activity demonstrates a positive correlation across all communities, with higher-degree nodes exhibiting greater investment frequency. This pattern suggests that network position, as measured by degree centrality, serves as a reliable predictor of investment activity levels within the venture capital ecosystem.

Add more centrality metrics

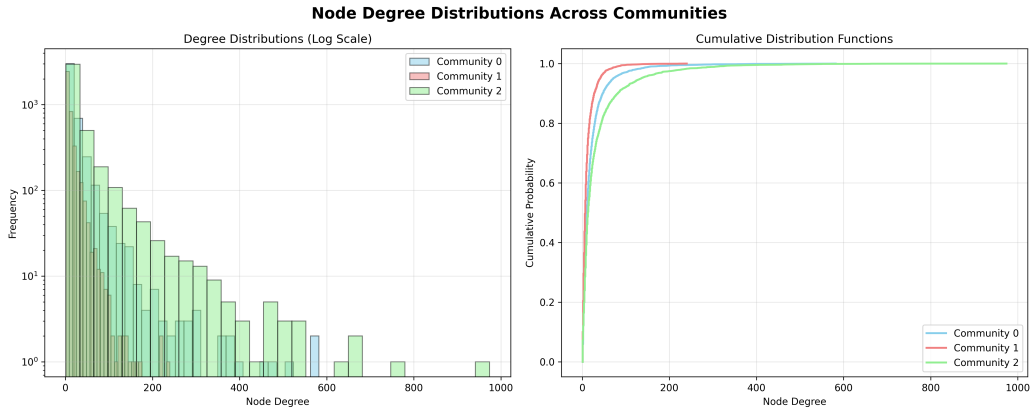


Figure 2: TBD

3.2.2 Geographic Distribution

Geographic analysis reveals distinct spatial clustering patterns across the three largest communities. Figure 5 illustrates the asymmetric geographic distributions between early-stage and late-stage investment networks.

Better format geographic distribution figure

Communities 0 and 2 exhibit similar geographic profiles with predominantly American investors, reflecting the dominance of U.S.-based venture capital (important to remember our dataset contains only American startups' investments, which include international investors).

However, regional analysis within the United States reveals a striking pattern: Community 2 demonstrates exceptional concentration in California, particularly Silicon Valley, with approximately 50% more California-based investors than either Community 0 or Community 1 for both late-stage and early-stage investor categories.

This Silicon Valley concentration in Community 2 is particularly notable given the region's status as the world's premier innovation ecosystem. The dominance of California investors in this community, which also exhibits the highest transaction volumes, suggests potential advantages conferred by

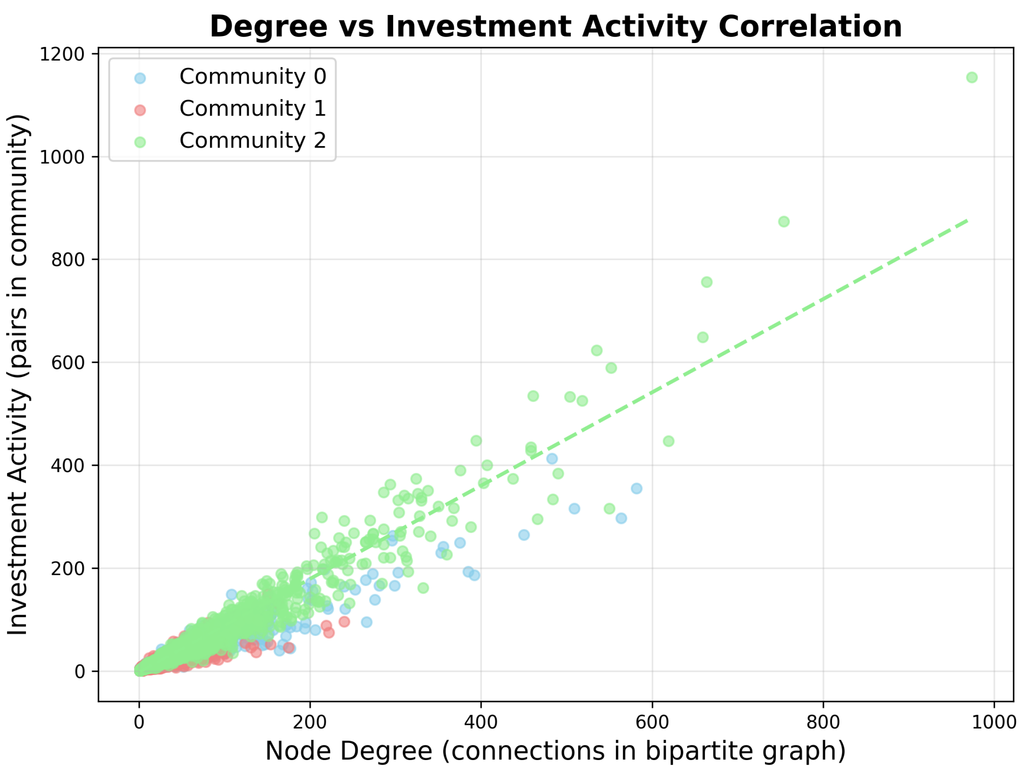
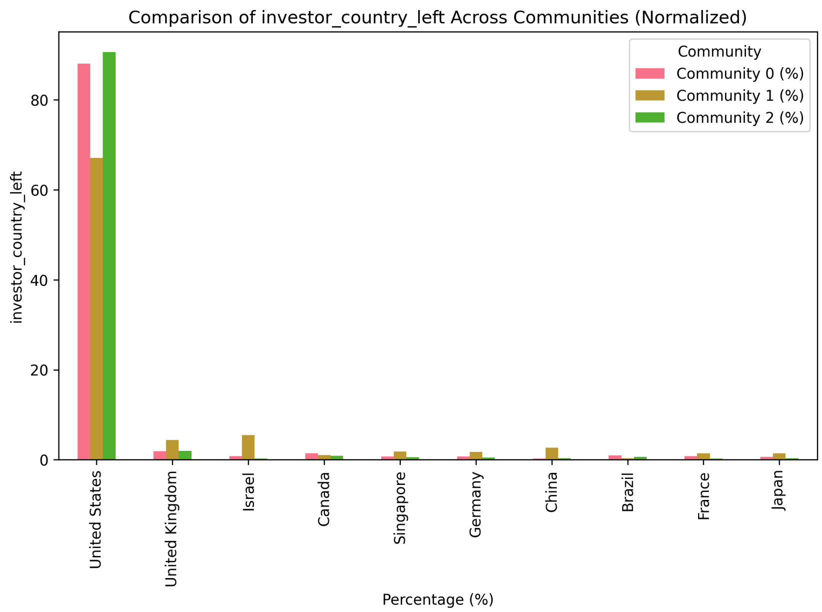
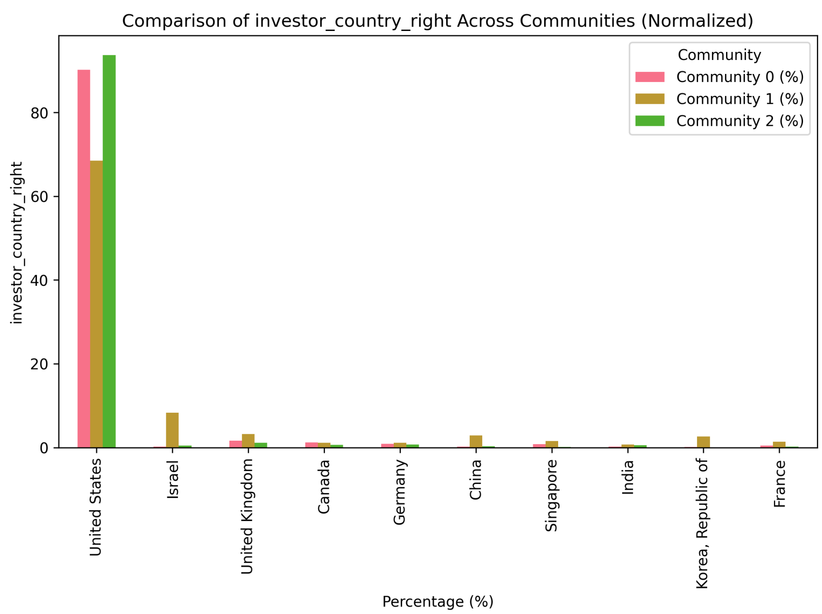


Figure 3: Relationship between node degree and investment activity for the three largest investor communities. The scatter plot demonstrates a positive correlation: higher-degree nodes tend to participate in more investment activities. This pattern is consistent across all communities, supporting the interpretation that network centrality is a strong predictor of investment frequency and influence within the venture capital ecosystem.

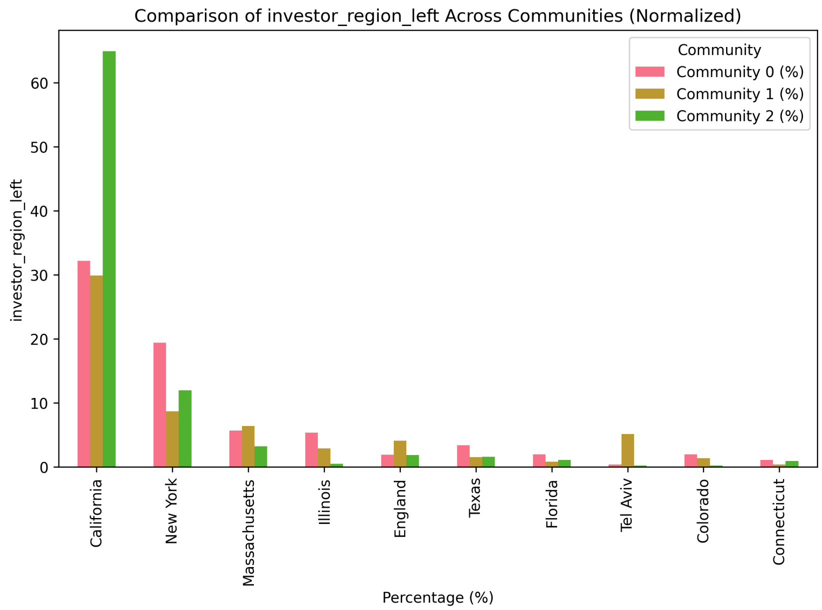


(a) Late-stage investors geographic distribution (countries)

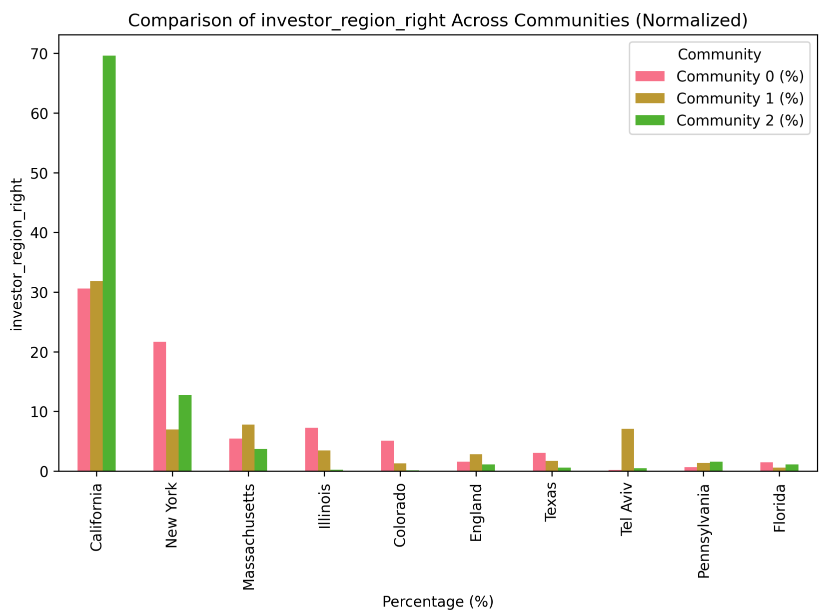


(b) Early-stage investors geographic distribution (countries)

Figure 4: Geographic distribution of venture capital investors across the largest communities. The bipartite structure reveals differential geographic clustering between late-stage (top) and early-stage (bottom) investor networks, with Community 1 exhibiting greater international diversification compared to the U.S.-concentrated Communities 0 and 2.



(a) Late-stage investors geographic distribution (regions)



(b) Early-stage investors geographic distribution (regions)

Figure 5: Geographic distribution of venture capital investors across regions

Community	Investor	Degree	Type
0	Techstars-seed	582	Early-stage
	500 Global-seed	564	Early-stage
	Gaingels-series_b	509	Late-stage
	Greycroft-series_a	483	Early-stage
	Bossa Invest-series_b	450	Late-stage
1	Intel Capital-series_b	240	Late-stage
	Norwest Venture Partners-series_a	222	Early-stage
	Canaan Partners-series_a	219	Early-stage
	Intel Capital-series_c	175	Late-stage
	SOSV-series_b	163	Late-stage
2	SV Angel-seed	974	Early-stage
	SV Angel-series_a	754	Early-stage
	Andreessen Horowitz-series_a	664	Early-stage
	Khosla Ventures-series_a	659	Early-stage
	New Enterprise Associates-series_a	619	Early-stage

Table 3: Top 5 high-degree nodes (95th percentile) for the three largest investor communities

geographic clustering within innovation hubs.

In contrast, Community 1 demonstrates significantly greater international diversification, with substantial representation from Israel, the United Kingdom, China, South Korea, Singapore, and France. This international composition in Community 1 may reflect different risk tolerance profiles, regulatory environments, or access to cross-border deal flow compared to the more domestically concentrated communities.

3.2.3 Funding Characteristics

Analysis of funding patterns reveals that Community 2 exhibits substantially higher funding frequency and larger investment amounts compared to the other communities, as demonstrated in Figure 6.

Add attachment with statistical proofs

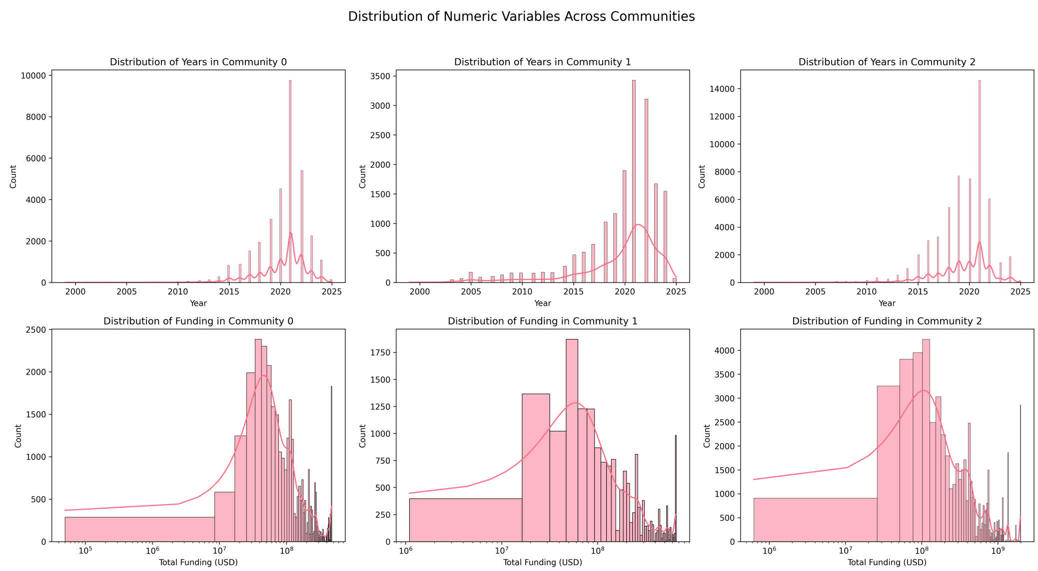


Figure 6: Funding characteristics across the three largest investor communities. The analysis reveals systematic differences in investment amounts, round frequency, and funding patterns between communities, with different organizational structures exhibiting distinct capital deployment strategies.

plot a graph of distribtuion of invesments among degrees of investors

Furthermore, communities exhibit concentrated capital deployment patterns, with higher-degree investors participating in larger funding rounds while maintaining broader portfolio diversification. This suggests that certain organizational structures may create more efficient capital allocation mechanisms compared to other network configurations.

Despite superficial similarities between Communities 0 and 2 in terms of size and geographic concentration within the United States, Community 2 appears to confer distinctive advantages in organizational efficiency. While both communities share comparable scales and American investor bases, Community 2's organizational structure enables it to achieve substantially higher transaction volumes (33.6% vs. 19.4% of total investments) and more comprehensive funding coverage across all investment stages.

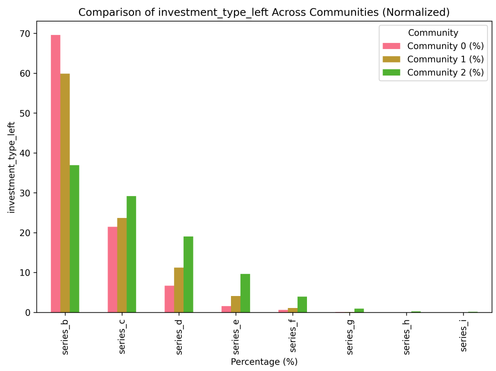
This pattern suggests that specific network topologies, rather than community size or geographic distribution alone, may be critical determinants of investment ecosystem efficiency.

add literature base for this strong assumption

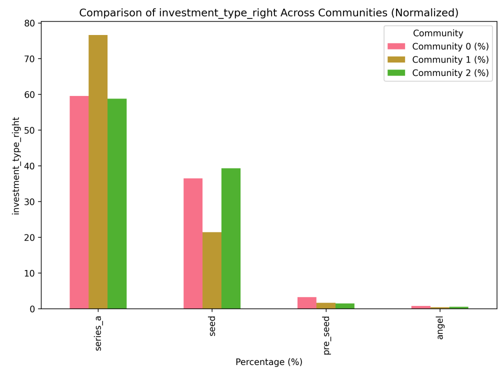
3.2.4 Investment Stage Preferences

The investment stage distributions reveal systematic specialization patterns across communities that align with their geographic profiles and transaction volumes. Figure 7 demonstrates the distribution patterns of investment types within the bipartite network structure.

Late-stage investment patterns: Community 2 dominates Series C and later funding rounds, demonstrating its role in growth-stage capital deployment. Community 1, despite its smaller transaction volume, shows strong representation in Series B and later stages, with participation rates exceeding Community 0 in Series C and beyond. Community 0 exhibits particular strength in Series B rounds while maintaining lower participation in



(a) Late-stage investment types distribution



(b) Early-stage investment types distribution

Figure 7: Investment stage distribution across the three largest communities. The distribution patterns reveal stage-specific specialization within investor communities.

later stages compared to Community 2.

Early-stage investment patterns: Community 2 shows prominence in seed-stage investments while maintaining comparable levels to Community 0 in Series A funding. Both Communities 0 and 2 participate actively in angel and seed rounds, though Community 0 shows relatively higher pre-seed activity. Community 1 demonstrates concentrated focus on Series A investments, aligning with its international profile and suggesting specialization in cross-border early-growth funding.

These stage-specific patterns suggest that Community 2's organizational structure facilitates participation across the entire funding spectrum, from seed to late-stage rounds, potentially enabling more comprehensive support for portfolio companies throughout their development lifecycle.

3.2.5 Sectoral Focus

The sectoral analysis reveals distinct specialization patterns that align with each community's structural and geographic characteristics. Figure 8 illustrates the distribution of investment focus across technology sectors, demonstrating how different communities exhibit varying degrees of sectoral concentration.

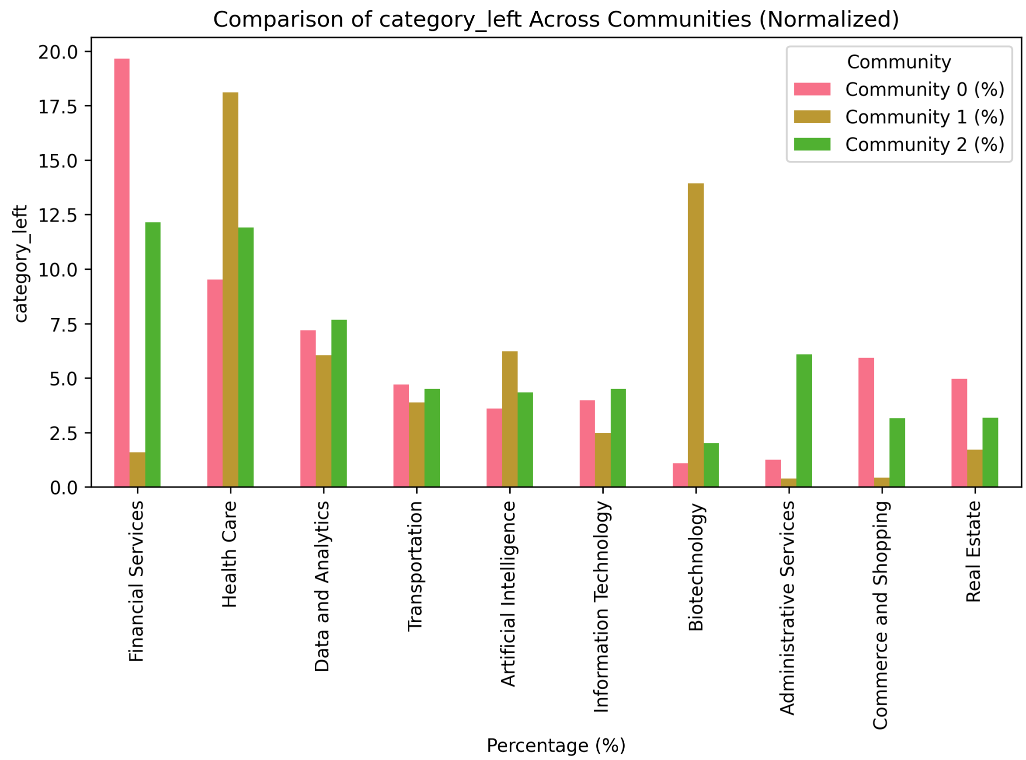


Figure 8: Sectoral distribution across the three largest investor communities. The analysis reveals differential industry focus patterns, with certain communities demonstrating concentrated investment strategies in specific technology sectors while others maintain broader sectoral diversification.

Comment sectorial distribution

Community 0: Concentrates in real estate, commerce and shopping, and financial services, while maintaining standard representation in information technology and artificial intelligence. Notably absent from administrative services and biotechnology, suggesting specialized expertise in consumer-facing and financial technology sectors.

Community 1: Specializes significantly in biotechnology and healthcare, with enhanced artificial intelligence focus compared to other communities. Shows reduced interest in real estate, commerce and shopping, and administrative services. The biotechnology concentration aligns with its international composition, potentially reflecting access to global biotech innovation hubs.

Community 2: This community exhibits strong representation in administrative services while maintaining comparable levels to Community 0 in information technology and artificial intelligence. Despite its larger transaction volume, it shows lower concentration in real estate and financial services than Community 0, suggesting that its organizational structure facilitates broader sectoral participation rather than concentrated specialization. This sectoral breadth, combined with the community's Silicon Valley concentration, indicates potential advantages of diversified investment strategies within innovation-rich geographic clusters.

Community 0 serves as an effective structural baseline for comparison with Community 2, given their similar geographic profiles and certain sectoral overlaps, while differing significantly in network organization and transaction volumes. The systematic differences observed between these structurally similar communities highlight the potential impact of organizational structures on investment behavior and capital deployment efficiency.

3.3 Overall Nestedness Findings

Nestedness analysis across investor communities reveals heterogeneous structural patterns. Among the 5 communities examined, one exhibits statistically significant nestedness ($p < 0.01$) relative to degree-preserving null models generated through the Curveball algorithm.

Figure 9 presents the comparison between observed and null model nestedness scores, where each plot represents a distinct community's null model NODF distribution accompanied by its observed NODF value.

Community 2 demonstrates the most pronounced nestedness, exhibiting

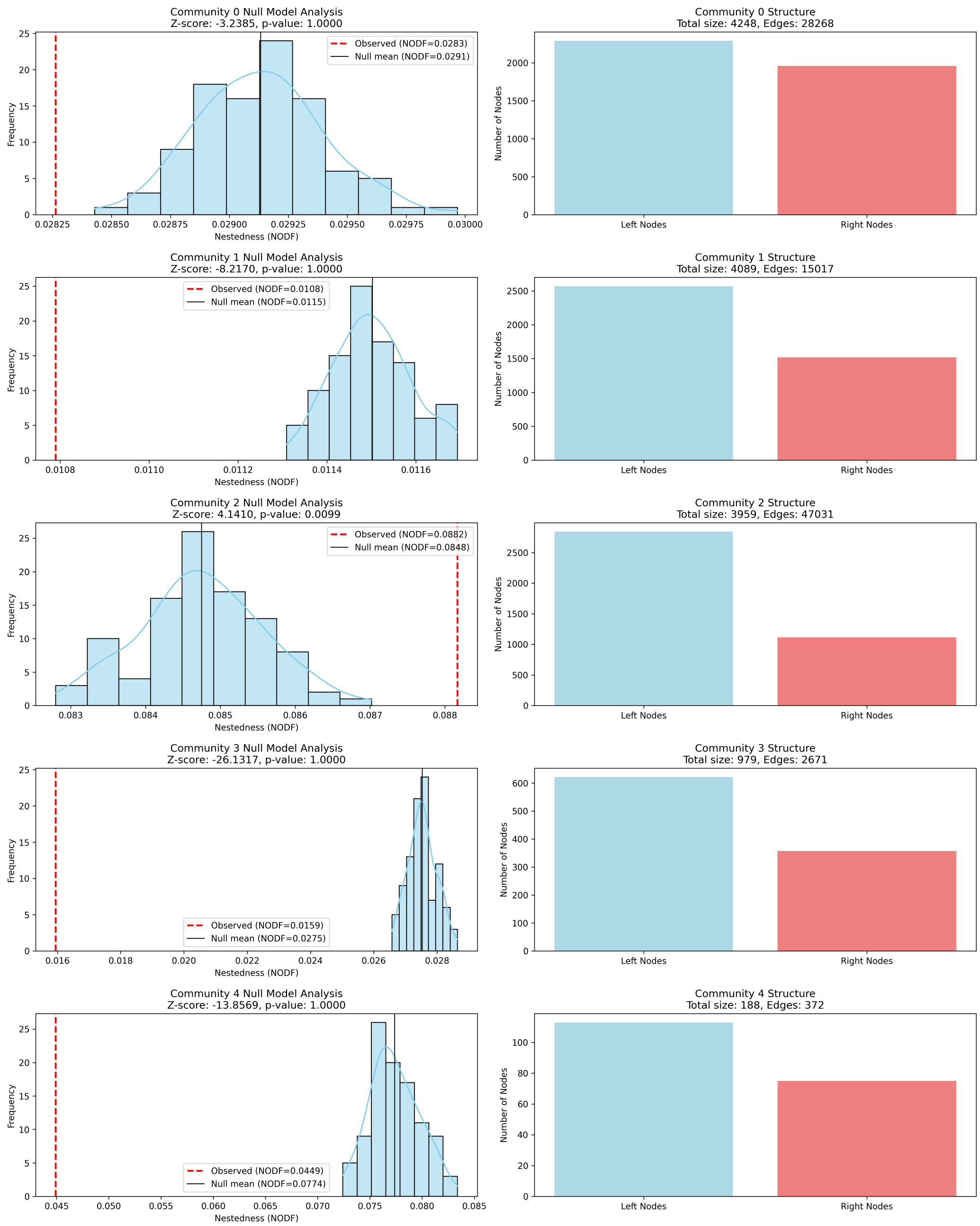


Figure 9: Comparison of observed versus null model nestedness scores for the five largest investor communities. The diagonal line represents equal observed and expected values, with points above the line indicating higher-than-random nestedness. "Left" refers to late-stage investors, and "Right" refers to early-stage investors.

an NODF score of 0.088 with statistical significance of $p = 0.00001$. This indicates a non-random behavior where less-connected investors maintain co-investment relationships with subsets of partners associated with highly-connected investors, creating a hierarchical investment pattern.

Additionally, Community 2 exhibits an asymmetric composition with a pronounced ratio favoring late-stage investors over early-stage investors. This imbalance may contribute to the nested structure by creating hierarchical dependencies between investor types.

Explore how imbalance contribute to nestedness according to literature

The following sections provide detailed characterization of the 3 most relevant communities in terms of number of investors, while analyzing in-depth Community 2 (where nestedness was observed) through comparison with Communities 0 and 1, which serve as contrasting examples of similar-sized but differently structured investor networks.

3.4 Evolution of Nestedness in Silicon Valley Community

To understand the emergence and development of nested structure, we conducted a temporal analysis of Community 2's nestedness evolution using cumulative window. This analysis reveals how the significantly nested structure observed in the static analysis developed over time and when it became statistically distinguishable from random network configurations.

The temporal analysis spans 18 years (2007-2024) with sufficient data for nestedness calculation. Early periods (2004-2006) contained insufficient investment activity for meaningful analysis. The evolution demonstrates three distinct phases: an early non-significant period (2007-2018), a transition period (2019), and a sustained significant period (2019-2024).

During the non-significant period (2007-2018), observed nestedness scores ranged from 0.15 to 0.38, consistently failing to exceed null model expectations. Z-scores remained predominantly negative, indicating that the observed network structure was less nested than expected under random configuration preserving the degree sequence. This suggests that early venture capital network organization in Community 2 followed patterns indistinguishable from random syndication among investors with equivalent activity levels.

The transition occurred in 2019, marking the first year when Community 2 achieved statistically significant nestedness (Z-score: 2.91, p-value < 0.001). This transition coincided with substantial network growth, reaching 2,364 total nodes and 33,805 edges.

Notably, the 2019 transition corresponds to the network achieving a connectance threshold of 0.0264, suggesting that nestedness emergence may be related to specific density conditions within large-scale investment networks.

The sustained significant period (2019-2024) demonstrates consistent statistical significance with progressively strengthening Z-scores, reaching a maximum of 4.76 in 2022.

Interestingly, while absolute nestedness scores decreased from 0.38 in early periods to 0.088 in 2024, the statistical significance increased significantlly. Table 4 presents key statistics from the temporal evolution analysis.

This apparent paradox reflects the fundamental principle that nestedness significance depends on comparison with appropriate null models rather than absolute magnitude [7].

Period	**Years**	**Mean NODF**	**Mean Z-score**	**Significant Years**
Early (2007-2018)	12	0.2547 ± 0.1158	-1.4442 ± 1.1108	0
Transition (2019)	1	0.1177	2.9125	1
Significant (2019-2024)	6	0.1012 ± 0.0121	3.8950 ± 0.7182	6

Table 4: Summary statistics for Community 2 nestedness evolution across temporal periods

The analysis reveals several notable patterns in network structure evolution. Late-stage investor participation increased consistently over time, while early-stage investor numbers showed initial growth followed by stabilization

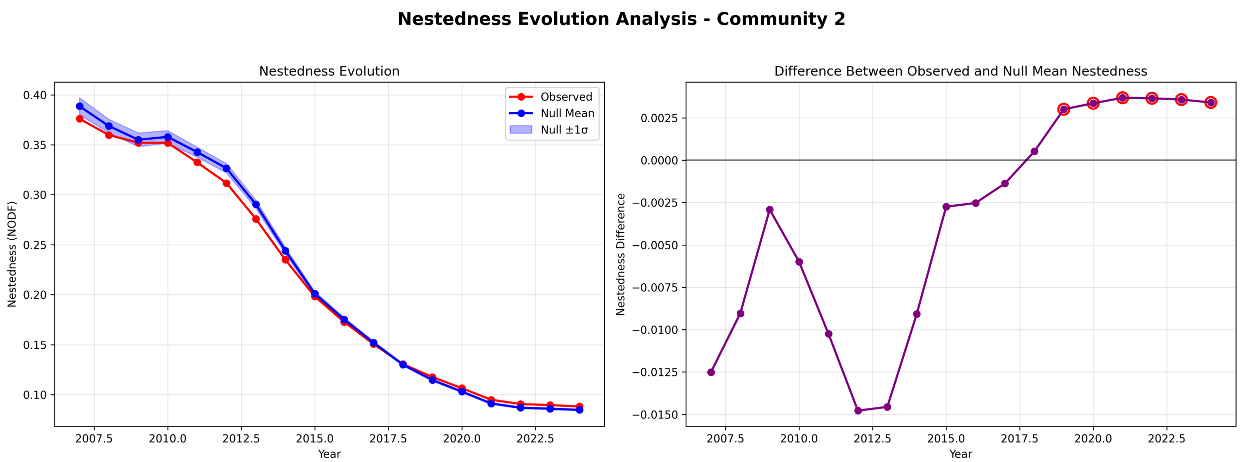


Figure 10: !!!UPDATE COMMENT Temporal evolution of nestedness in Community 2. The left subplot shows the observed nestedness (red) and null model mean (blue) over time, with shaded areas representing the null model standard deviation. The right subplot displays the evolution of network size, with total nodes (green) and total edges (purple) over the same period. The figure highlights the phase transition in 2019, where observed nestedness becomes statistically significant and the network grows rapidly.

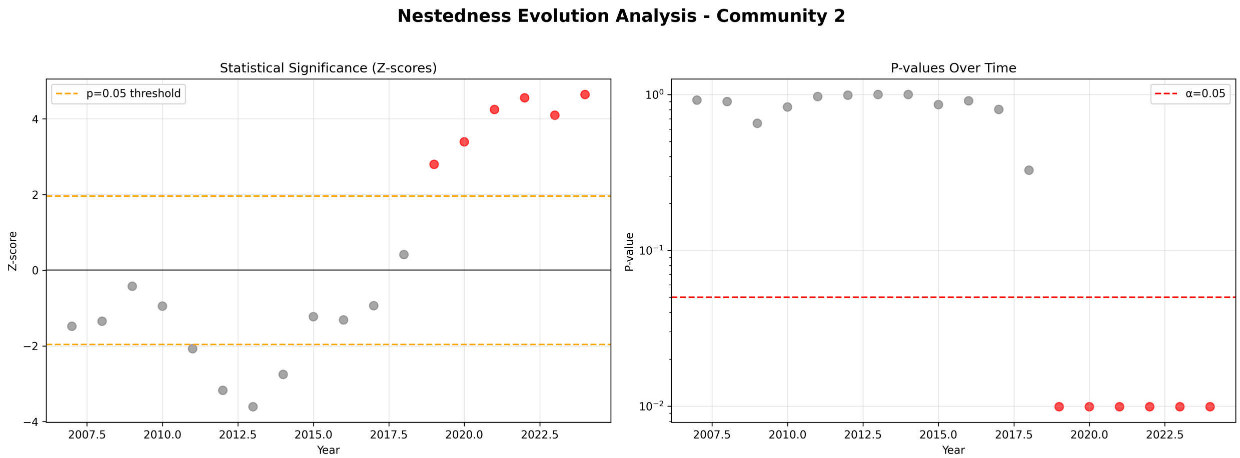


Figure 11: !!!UPDATE COMMENT Connectance and statistical significance evolution in Community 2. The left subplot shows the decreasing trend in connectance (network density) over time, while the right subplot presents the evolution of p-values on a logarithmic scale, indicating the emergence of statistically significant nestedness after 2019. The figure demonstrates how the network becomes sparser yet more hierarchically organized, with significance emerging as the network reaches critical size and density thresholds.

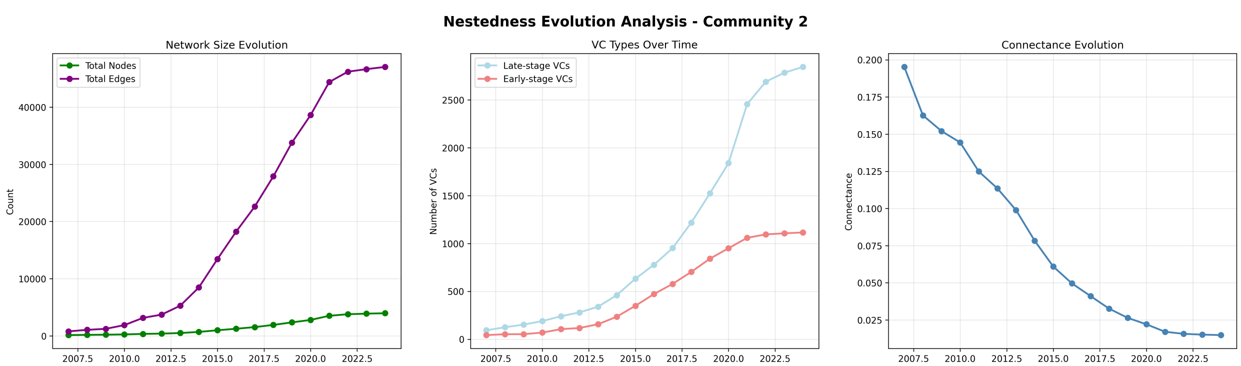


Figure 12: TBD

after 2020. This asymmetric growth pattern contributed to the development of the nested structure by creating increasingly hierarchical relationships between investor types.

Connectance exhibited a systematic decline from 0.195 in 2007 to 0.015 in 2024, reflecting the network's evolution toward sparser but more strategically organized connections. Despite this decreased density, the emergence of statistical significance suggests that the remaining connections became increasingly hierarchically organized, with less-connected investors maintaining relationships with subsets of the partners of highly-connected investors.

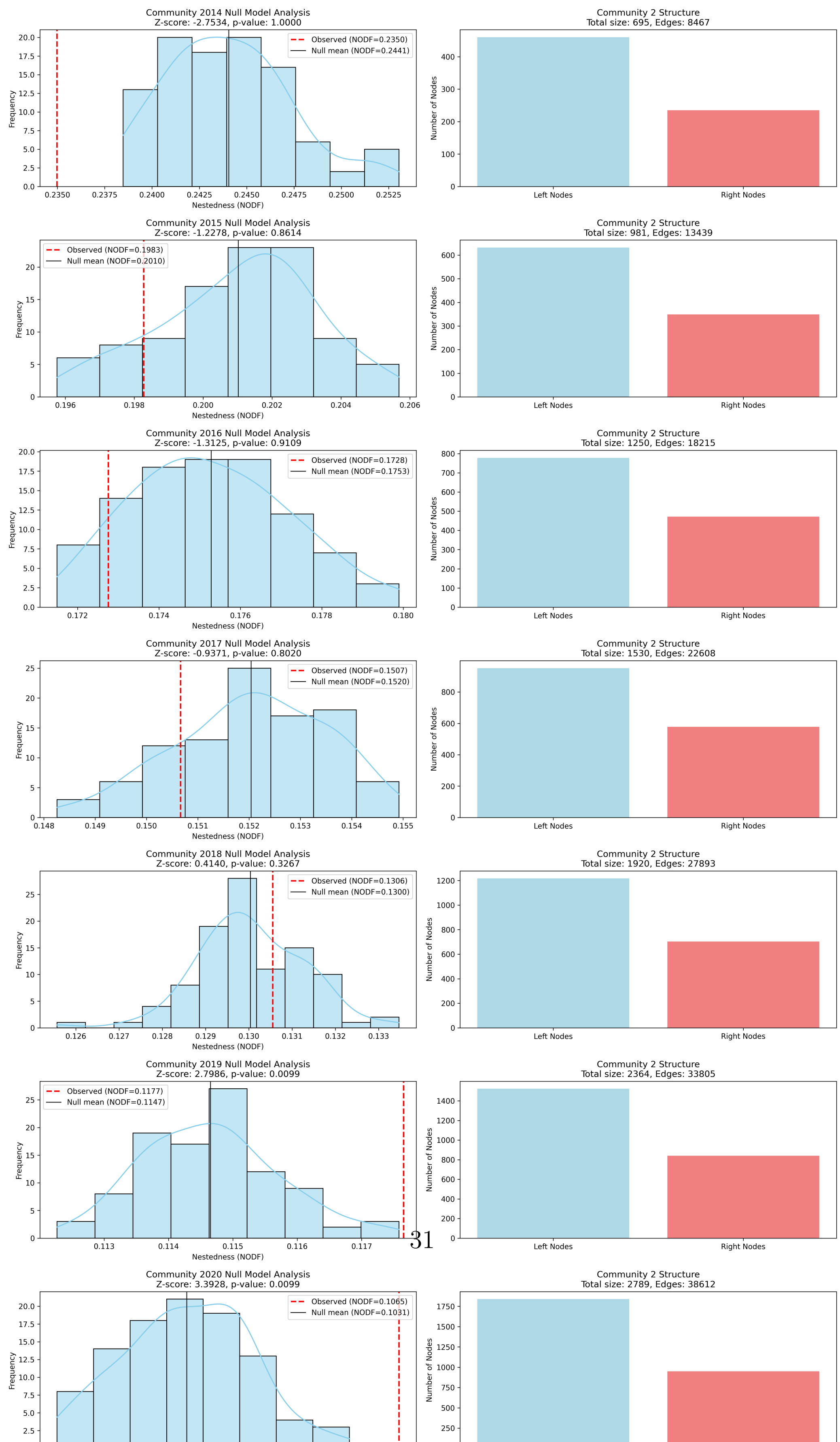
Detailed analysis of the significant periods (2019-2024) reveals consistent patterns in network organization. Each significant year demonstrates similar structural characteristics: large networks ($>$2,300 nodes), substantial edge counts ($>$33,000), low connectance ($<$0.027), and strong statistical significance (Z-scores $>$2.9). This consistency suggests that the nested structure represents a stable organizational state that persists once established.

The temporal analysis provides evidence that nestedness in venture capital networks emerges through a phase transition process rather than gradual development. The sharp transition from non-significant to highly significant nestedness in 2019, followed by sustained significance, suggests threshold effects in network organization.

This pattern aligns with theoretical frameworks from complex network theory indicating that certain topological properties emerge discontinuously as networks reach critical size or density parameters [7].

Make reflection about possible link with network effects from economics and industrial organization

The period of nestedness emergence (2019-2024) corresponds with an increase in late-stage investor participation and a relative decrease in early-stage investor numbers within the community. This asymmetric evolution may contribute to the hierarchical structure by creating conditions where early-stage investors increasingly depend on relationships with a subset of


Community 2014 Null Model Analysis
Z-score: -2.7534, p-value: 1.0000
Observed (NODF=0.2350)
Null mean (NODF=0.2441)
Frequency
Nestedness (NODF)
Community 2 Structure
Total size: 695, Edges: 8467
Number of Nodes
Left Nodes
Right Nodes
Community 2015 Null Model Analysis
Z-score: -1.2278, p-value: 0.8614
Observed (NODF=0.1983)
Null mean (NODF=0.2010)
Frequency
Nestedness (NODF)
Community 2 Structure
Total size: 981, Edges: 13439
Number of Nodes
Left Nodes
Right Nodes
Community 2016 Null Model Analysis
Z-score: -1.3125, p-value: 0.9109
Observed (NODF=0.1728)
Null mean (NODF=0.1753)
Frequency
Nestedness (NODF)
Community 2 Structure
Total size: 1250, Edges: 18215
Number of Nodes
Left Nodes
Right Nodes
Community 2017 Null Model Analysis
Z-score: -0.9371, p-value: 0.8020
Observed (NODF=0.1507)
Null mean (NODF=0.1520)
Frequency
Nestedness (NODF)
Community 2 Structure
Total size: 1530, Edges: 22608
Number of Nodes
Left Nodes
Right Nodes
Community 2018 Null Model Analysis
Z-score: 0.4140, p-value: 0.3267
Observed (NODF=0.1306)
Null mean (NODF=0.1300)
Frequency
Nestedness (NODF)
Community 2 Structure
Total size: 1920, Edges: 27893
Number of Nodes
Left Nodes
Right Nodes
Community 2019 Null Model Analysis
Z-score: 2.7986, p-value: 0.0099
Observed (NODF=0.1177)
Null mean (NODF=0.1147)
Frequency
Nestedness (NODF)
Community 2 Structure
Total size: 2364, Edges: 33805
Number of Nodes
Left Nodes
Right Nodes
Community 2020 Null Model Analysis
Z-score: 3.3928, p-value: 0.0099
Observed (NODF=0.1065)
Null mean (NODF=0.1031)
Frequency
Community 2 Structure
Total size: 2789, Edges: 38612
Number of Nodes

the partners associated with highly-connected late-stage investors.

Add reference to figure of nestedness vs null models between 2019 and 2024

However, whether this temporal correlation reflects causal mechanisms or represents coincidental market dynamics requires further investigation with appropriate theoretical frameworks [4].

4 Discussion and Implications

As outlined in the Results section, this article first presents a detailed characterization of investor communities and then examines nestedness within those communities; the discussion below follows that ordering and builds directly on the community-level descriptions provided earlier.

The discovery, documented in the Results section, of significantly nested communities within the late-early stage venture capital network provides new insights into investor behavior and startup access to capital. The hierarchical structure in Community 2 suggests that informal investment hierarchies may systematically influence funding accessibility for entrepreneurs.

The concentration of nested structure specifically within Silicon Valley investors adds a significant geographic dimension to these findings. Community 2's exceptional dominance by California-based investors, particularly those in Silicon Valley, suggests that geographic clustering within the world's premier innovation ecosystem may facilitate the emergence of hierarchical investment structures.

This pattern may indicate that dense information networks, frequent face-to-face interactions, and shared risk assessment practices characteristic of innovation hubs naturally give rise to nested investor relationships. The dominance of California investors in the only statistically nested community suggests a potential relationship between geographic clustering in innovation hubs and the emergence of hierarchical investment structures.

Furthermore, the comparative analysis between Communities 0 and 2 reveals that nestedness functions as an organizational catalyst that transforms otherwise similar investor communities. Despite comparable sizes and geographic concentration within the United States, Community 2's nested structure enables substantially higher transaction volumes and more comprehensive funding coverage across all investment stages.

This suggests that network topology, rather than community size or geographic distribution alone, may be the critical determinant of investment ecosystem efficiency. The hierarchical organization in Community 2 appears to create more efficient capital allocation mechanisms, with higher-degree investors participating in larger funding rounds while maintaining broader portfolio diversification compared to randomly structured networks.

4.1 Degree Distribution Patterns and Hub Organization

The analysis of degree distributions across communities reveals fundamental organizational differences that complement the nestedness findings. All three communities exhibit power-law degree distributions characteristic of scale-free networks [3], but with distinct magnitude and structural parameters that reflect different organizational strategies.

The similarity in degree distribution magnitudes between Communities 0 and 2, contrasted with Community 1's consistently lower values, suggests that network scale alone does not determine investment efficiency. Community 2's superior performance in investment volumes and nested organization occurs despite degree distribution patterns nearly identical to Community 0.

This finding reinforces that the specific arrangement of connections, rather than their quantity or distribution, drives the observed efficiency advantages.

Analysis of high-degree nodes (network hubs) reveals distinct organizational philosophies across communities. Community 2's hub structure demonstrates exceptional concentration among early-stage Silicon Valley investors, with SV Angel achieving remarkable connectivity across both seed and Series A stages.

This pattern contrasts with Community 0's more diversified hub structure and Community 1's balanced early-late stage distribution.

The concentration of hubs within the early-stage investment category in Community 2 may facilitate the nested structure by creating clear hierarchical pathways from early-stage to late-stage investment relationships.

The positive correlation between degree centrality and investment activity across all communities validates network position as a predictor of investor influence. However, the strength of this relationship appears amplified within the nested Community 2, suggesting that hierarchical organization may enhance the efficiency of high-degree investors in deploying capital and identifying investment opportunities.

4.2 Temporal Dynamics and Phase Transition Emergence

Add observation that literature on other formations were already explored, but VC-VC is rare

The temporal evolution analysis of Community 2 provides unprecedented insight into how nested structures emerge within VC-VC investment networks. The identification of a sharp phase transition in 2019, rather than gradual nestedness development, challenges assumptions about evolutionary network organization and suggests threshold-dependent emergence mechanisms.

The three-phase evolution pattern (non-significant period between 2007-2018, transition in 2019, and sustained significance during 2019-2024) indicates that nested organization in venture capital networks may represent a distinct organizational state that emerges discontinuously when specific conditions are met.

This finding aligns with theoretical predictions from complex network theory regarding critical transitions in topological properties [7].

The apparent paradox of decreasing absolute nestedness scores (from 0.38 to 0.088) coinciding with increasing statistical significance reflects sophisticated changes in network organization. As the network grew substantially larger, maintaining even modest levels of hierarchical organization became increasingly difficult under random formation processes, making the observed nested patterns more statistically remarkable.

The correspondence between nestedness emergence and specific connectance thresholds (approximately 0.026) provides practical insights for ecosystem development.

This threshold may represent a critical density where hierarchical organization becomes sustainable within large-scale investment networks, offering guidance for policy interventions aimed at fostering similar organizational efficiency in other innovation ecosystems.

The asymmetric evolution of investor types—increasing late-stage participation coupled with stabilizing early-stage numbers—may have facilitated nested structure emergence by creating conditions favoring hierarchical relationships.

This pattern suggests that the development of nested organization may require specific demographic conditions within investor communities, rather than simply network growth or density changes.

4.3 Network Robustness and Resilience

The nested structures challenge assumptions of random mixing in venture capital markets, suggesting that certain investors function as "gatekeepers" who control access to broader investment networks. This finding aligns with social network theories about structural holes and brokerage positions [3].

Following insights from nestedness research in complex networks [7], the hierarchical organization observed in Community 2 may confer distinct robustness properties to the venture capital ecosystem.

In mutualistic networks, nestedness typically enhances stability against random node removal but creates vulnerability to targeted elimination of highly connected nodes. Applied to venture capital, this suggests that while nested investor communities may be resilient to random investor departures, they could be particularly vulnerable to the exit of key hub investors.

The concept of "mutualistic trade-offs" from ecological network theory provides a framework for understanding these dynamics. In nested venture capital communities, less-connected investors maintain relationships with subsets of the partners associated with highly-connected investors, creating dependencies that could influence network stability.

Future research should investigate whether less-connected venture capital firms exhibit higher exit probabilities, which would support the hypothesis that nestedness creates hierarchical fragility patterns.

5 Conclusion and Future Directions

5.1 Individual Nestedness Contributions

An important avenue for future research involves analyzing individual nestedness contributions within these communities. Rather than treating nestedness as a global network property, examining how specific investors contribute to the overall nested structure could reveal mechanisms driving community formation and persistence.

This approach could help predict which network positions are most vulnerable to disruption and identify critical nodes whose removal would significantly alter community structure.

Understanding individual contributions to nestedness could also inform strategies for network intervention and ecosystem development. If certain investor positions disproportionately contribute to nested stability, targeted support or policy interventions could enhance overall ecosystem resilience.

5.2 Dynamic Network Evolution

The temporal analysis reveals several critical insights about the dynamic processes underlying venture capital network organization. The sharp phase transition observed in Community 2's nestedness evolution suggests that network topology may be subject to discontinuous organizational changes rather than gradual evolution. This finding has important implications for understanding how investment ecosystems develop and potentially collapse.

The three-phase temporal pattern—extended non-significant periods, rapid transition, and sustained significance—may represent a general framework for understanding organizational emergence in investment networks. The identification of specific connectance thresholds (approximately 0.026) associated with nestedness emergence provides quantitative targets for ecosystem development strategies.

Future research should investigate whether similar threshold effects exist in other geographic markets and whether policy interventions can facilitate reaching these critical organizational states.

The asymmetric evolution of investor types during the transition period offers insights into the demographic conditions that may facilitate nested organization. The pattern of increasing late-stage investor participation coupled with stabilizing early-stage numbers suggests that hierarchical organization may require specific ratios between investor types.

This finding could inform strategies for ecosystem development in emerging markets, where the balance of early-stage and late-stage capital availability is often suboptimal.

The temporal analysis also reveals that nestedness persistence appears robust once established. The sustained significance observed from 2019-2024, despite substantial network growth and changing market conditions, suggests that nested organization may represent a stable attractor state in investment network evolution.

This stability has important implications for long-term ecosystem planning and suggests that successful development of nested organization may provide lasting competitive advantages for innovation hubs.

5.3 Causal Mechanisms and Economic Outcomes

The identification of these nested communities opens several avenues for future research into the social and economic mechanisms that drive venture capital ecosystem organization. [7] provides theoretical frameworks for understanding why nestedness emerges in complex systems, including factors such as heterogeneous node fitness, temporal constraints on link formation, and spatial or industry-specific constraints on partnership formation.

Investigating whether nested communities provide superior or inferior outcomes for portfolio companies compared to randomly organized investor groups represents a research opportunity. The concentrated capital deployment patterns observed in Community 2 suggest potential efficiency advantages, but these must be weighed against potential risks from reduced diversity and increased systemic vulnerability.

5.4 Policy and Ecosystem Development Implications

The concentration of nested structures within Silicon Valley suggests that geographic proximity within innovation hubs may be a prerequisite for the emergence of hierarchical investor relationships. This finding has important implications for ecosystem development strategies in other regions seeking to replicate Silicon Valley's success.

Understanding these patterns may inform policy discussions about startup ecosystem development and investor network formation. If nested structures facilitate higher transaction volumes and more comprehensive funding support, as observed in Community 2, policies that encourage the formation of such hierarchical investor relationships might enhance ecosystem efficiency.

However, the geographic specificity of this pattern suggests that simply replicating formal structures may be insufficient—the dense information networks and shared practices of established innovation hubs appear to be necessary conditions for nested organization to emerge.

Conversely, if nestedness creates barriers to entry for new investors or reduces access for certain entrepreneur populations, regulatory interventions might be warranted to promote more equitable network organization.

The dominance of Silicon Valley investors in the nested community raises questions about geographic bias in capital allocation and whether hierarchical structures may inadvertently concentrate investment opportunities within established innovation centers.

The possibility of network rewiring , analogous to ecological community adaptation, might confer additional robustness to venture capital ecosystems. Policies that facilitate investor mobility and relationship reformation could enhance system-wide resilience while maintaining the efficiency benefits of

mention literature about rewiring

nested organization.

Investigate whether the Silicon Valley concentration in nested communities reflects unique geographic advantages, information network density, or institutional factors that could be replicated in other innovation ecosystems.

Examine the relationship between geographic clustering in innovation hubs and the emergence of nested investor structures across different global venture capital markets.

Investigate the economic consequences of nested community structure on startup success rates and funding efficiency.

Analyze individual nestedness contributions to identify critical nodes and understand how specific investor positions contribute to overall community stability and structure.

Investigate the robustness properties of nested venture capital communities, particularly vulnerability to targeted removal of highly connected investors versus resilience to random investor departures.

Examine the relationship between investor position within nested hierarchies and probability of network exit, testing whether less-connected VCs exhibit higher departure rates.

Apply social network theories of structural holes to understand the role of highly connected investors in nested communities and their function as potential gatekeepers.

Investigate whether the nested structure reflects information asymmetries, risk-sharing mechanisms, or industry-specific constraints among investors.

Develop theoretical models to explain the emergence of nested structures in investment networks, incorporating insights from Chapter 5 of "Nestedness in complex networks" regarding heterogeneous node fitness and temporal constraints.

Compare nestedness patterns across different geographic markets and time periods to understand generalizability and cultural influences on network organization.

Investigate the relationship between degree distribution patterns and nestedness emergence, examining whether specific scale-free parameter ranges facilitate hierarchical organization.

Explore the role of hub investor strategies in nested community formation, particularly investigating how early-stage hub concentration may facilitate hierarchical pathway development.

Analyze the predictive power of connectance thresholds for nestedness emergence in other venture capital markets and innovation ecosystems.

Examine the stability mechanisms that maintain nested organization once established, investigating whether demographic balance between investor types is necessary for persistence.

References

[1] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[2] Mário Almeida-Neto, Paulo Guimarães, Paulo R. Guimarães Jr, Rafael D. Loyola, and Werner Ulrich. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, 117(8):1227–1239, 2008.

[3] Stephen P. Borgatti and Daniel S. Halgin. On network theory. *Organization Science*, 22(5):1168–1181, 2011.

[4] Jean-Michel Dalle et al. Accelerator-mediated access to investors among early-stage start-ups. *Research Policy*, 2025. In press.

[5] Mark Granovetter. Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3):481–510, 1985.

[6] Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[7] Sebastian Mariani. *Nestedness in complex networks*. Springer, 2019. ISBN: 978-3-030-29984-6.

[8] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, 2001.

[9] Giovanni Strona, Domenico Nappo, Francesco Boccacci, Simone Fattorini, and Jesús San-Miguel-Ayanz. A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, 5(1):4114, 2014.

[10] Theophile. Complex networks in entrepreneurial ecosystems. 2024. @todo add info.