Who brings investments? Venture Capital through reputation, networks, and influence

Marta Zava [*†], Stefano Caselli [‡]

July 2024, work in progress

Abstract

In the venture capital ecosystem, objectively quantifying startup quality is challenging. Thus, reputation and networks of influence play a crucial role in determining access to resources. This study examines the impact of investors on startup funding outcomes. We find that a company's likelihood of securing subsequent funding is positively correlated with the influence network of its previous investors. This effect is most pronounced in early-stage deals and diminishes in later stages, though its impact on the amount of capital raised persists. To analyze network influence, we introduce the funding attraction index, a novel time-sensitive metric that quantifies an investor's ability to attract follow-on investments from co-investors to their portfolio companies. We also develop an extension of the Sorenson-Heckman model, allowing us to differentiate between an investor's role in selecting promising companies and their contribution to post-investment success. Our findings shed light on the complex dynamics of venture capital networks and their impact on startup trajectories.

[*]Bocconi University, Goethe University, MUSA Program.
[†]Corresponding author: marta.zava@unibocconi.it
[‡]Bocconi University. stefano.caselli@unibocconi.it

We are extremely grateful for the comments of Uwe Walz, Laura Bottazzi, Josh Lerner, Rebel Cole, Sabrina Howell, Rose Liao, Giulio Cimini, Hernan Makse, and Alessandro Previtero. We also thanks the audiences of Bocconi University, Elsevier Future of Finance Conference 2023, Centro Ricerche Enrico Fermi Seminar Series, ENTFIN 2023 conference, Conference in Complex Systems 2023, AFA-AFFECT Brown-Bag Series, Netsci 2023, Complex Networks 2023, Netsci-X 2024, and Conference in Complex Systems 2022. Most importantly, we thank Bernardo Principi, Lucrezia Beretta, Giovanni Valle, Pietro Benedetti, and Efe Gulay for their exceptional research assistance.

1 Introduction New

Venture capital plays a crucial role in fostering innovation and economic growth by funding early-stage companies with high growth potential. However, the industry faces significant challenges in identifying promising startups and allocating resources effectively. High-profile failures like Theranos and Jawbone, which raised hundreds of millions before collapsing, highlight the difficulty of assessing startup quality. With around 40% of funded ventures failing in the first year and 90% ultimately ceasing operations (Timmons, 1990; Manigart et al., 2010), understanding the factors that influence startup success is of paramount importance.

This paper examines the role of investor networks and influence in determining startup outcomes, focusing on how the reputation and connections of early investors impact a company's ability to raise follow-on funding rounds. We introduce a novel concept, the Funding Attraction Index (FAI), to quantify an investor's ability to attract capital to their portfolio companies. Our research aims to fill a critical gap in the literature by providing a dynamic, network-based measure of investor influence that captures the ever-changing nature of the early-stage investment ecosystem.

To illustrate the power of the FAI compared to traditional metrics, consider the case of Andreessen Horowitz (a16z), a venture capital firm founded in 2009. By 2012, traditional cumulative reputation measures would have ranked a16z relatively low due to its young age and limited track record. However, our FAI captured a16z's rapidly growing influence in real-time, correctly predicting its outsized impact on subsequent funding rounds for its portfolio companies. This dynamic measure identified a16z as a top-tier influencer years before traditional metrics caught up, demonstrating FAI's ability to capture the fast-paced, cutting-edge nature of the VC ecosystem. In contrast to static, backwards-looking measures that inherently penalize younger investors, FAI recognizes that these newer entrants often have a better grasp of the latest innovations

- the very targets of VC investment. This example underscores how FAI's dynamic approach aligns with the rapidly evolving landscape of venture capital, where yesterday's metrics may fail to capture today's rising stars and tomorrow's kingmakers.

We develop a methodology combining network analysis, reputation metrics, and econometric techniques. Our approach involves: 1) constructing a temporally sensitive bipartite network of investors and startups; 2) formulating the Funding Attraction Index (FAI) to capture investors' funding attraction capacity; 3) projecting FAI scores onto companies seeking follow-on funding; 4) employing a Sorensen-Heckman approach (Bottazzi et al., 2008) to disentangle selection effects from influence effects. We analyze a comprehensive dataset of venture capital deals from 2010-2021, funding companies founded in California from 2010-2017, focusing on the transitions between seed, Series A, Series B, and Series C funding rounds.

Our analysis yields several results. First, a company's access to follow-on funding is strongly positively related to the network influence of its previous round investors, particularly in early stages. Second, the time-varying FAI predicts a company's ability to access subsequent rounds significantly better than traditional reputation measures, putting relatively young but effective funds in the spotlight. Third, while the FAI's impact on funding probability decreases in later stages, its relationship with the amount of money raised persists. Fourth, the influence effect is distinct from and more powerful than mere selection effects, indicating that influential investors actively shape startup outcomes. These findings fundamentally reshape our understanding of the venture capital ecosystem, demonstrating that network influence is a critical, dynamic factor in startup success.

Our work contributes to several key strands of literature. Considering network effects in venture capital, we extend the work of Hochberg et al. (2007, 2015) by focusing on the forward-looking influence of investors rather than static syndication networks. As comes to investor reputation and startup success, by building on Hsu (2005), Nahata (2008), and Nanda et al. (2020), we introduce a dynamic measure of reputation that captures real-time market developments and embraces the cutting-edge spirit of the venture capital scene. Then, taking into account venture capital decision-making, our findings complement the survey results of Gompers et al. (2016), providing quantitative evidence for the critical role of networks in deal flow and investment success. Lastly and more broadly, we contribute novel methodologies for analyzing temporal bipartite networks and quantifying influence in financial ecosystems, facilitating the blend between network science and finance.

The remainder of the paper is organized as follows: Section 2 discusses the relevant literature and hypotheses. Section 3 describes our methodology in detail. Section 4 specifies the data and variables used. Section 5 outlines our econometric model. Section

6 presents the empirical results and robustness checks. Finally, Section 7 concludes with a discussion of implications and future research directions.

2 Literature and Hypothesis Development

The early-stage investment ecosystem has garnered significant attention from researchers and practitioners alike due to its crucial role in fostering innovation, entrepreneurial ventures, and economic growth. This sphere of funding, encompassing angel investing, pre-seed, and seed funding, is known for its inherent high uncertainty and risk, posing challenges for comprehensive analysis. A dual game occurs: investors aim at identifying a profitable deal, rich in financial incentives, and preferably with a shareholder structure that would allow them to have some sort of influence over management in the upcoming future, albeit entrepreneurs seek to sell their equity dearly to either hand-off investors or to investors equipped with specific non-financial benefits (as expertise or connections) (Dixit and Pindyck, 1998). The current study intends to shed light on early-stage investors' impact over a company's business life, and how a company can position itself in a more favorable way when aiming at accessing follow-on funding rounds and greater amounts of capital.

One of the primary obstacles in studying early-stage investments lies in the scarcity of exhaustive datasets and the intricate interplay of undetectable and unobservable factors that influence investment decisions. As a company emerges, information about its potential and viability is limited and often held privately by the entrepreneur until subsequent funding rounds (Parra and Winter, 2022). Investors, however, assume a pivotal role in shaping the growth trajectory of a company (Inderst and Mueller, 2009), as their financial support and strategic guidance can significantly impact its development. Their expertise and resources can provide crucial opportunities for expansion, market penetration, and innovation. By making informed investment decisions and fostering a growth-enabling environment, investors act as catalysts for unlocking a company's full potential. Notably, angels have been found to play a particularly relevant role in fastening the growth of rapidly evolving companies with broad economic impact, whereas their influence appears less significant in more stable-paced businesses (Croce et al., 2021). This can be attributed to the angels' successful efforts in stimulating innovation within these high-growth ventures (Dutta et al., 2016). Furthermore, angels' presence has intuitively been associated with facilitating the subsequent rounds of capital raising, albeit primarily in the early stages of a company's life cycle, with its significance diminishing as the business matures (Kerr, Lerner and Schoar, 2014; Hellman and Thiele, 2015).

As concerns picking investments, VCs see the management team as more important than business related characteristics such as product or technology (Gompers et al., 2016). But what happens when entrepreneurs decide which investors to have on board? Indeed, companies are often attracted by an investor's reputation and are willing to offer them a 10-14% price discount for their affiliation (Hsu, 2007). This kind of investors acts as a certification for the company (Piacentino, 2019). Indeed, companies backed by reputable VCs are more likely to exit successfully, access public markets faster, and have higher asset productivity at IPO (Nahata, 2008). Investors' reputation is usually defined by their previous successful exists (Nanda et al., 2020; Amor and Kooli, 2019; Atanasov et al., 2012; Hsu, 2005), or by investors' experience, quantified as number of portfolio companies, of previous investments or simply by age (Krishnan et al., 2011; Amor and Kooli, 2019; Atanasov et al., 2012; Nahata, 2008).

Additionally, authors as Hsu (2004) included as reputation measure also the investors' network connections as estimated by Sorenson and Stuart (2008). Nevertheless, the first remarkable effort in connecting VCs' networks with the growth of their portfolio companies is attributable to Hochberg et al. (2007), who claim that the portfolio companies of better-networked VCs are significantly more likely to survive to subsequent financing and eventual exit. Indeed, VC syndication networks bring several advantages to the backed companies, as expanding the range of launch customers or strategic alliance partners (Bygrave, 1988), and generally providing better value-added services. As a matter of fact, ties among investors, as well as between investors and companies, or companies and companies, or companies and employees, have been extensively explored, often in relation with the company's successful exit (Zava and Caselli, 2023). But, as outlined, the studies regarding early-stage networks effects on follow-on investments have been limited. Among the few, using traditional network measures, Werth and Böert (2011) focused on business angel networks effects on raising series A. The current study intervenes to fill the gap in the literature.

We argue that receiving seed financing from high *funding attraction* investors has an impact in determining the chances raising series A funding, while the effect decreases at later stage in business life, namely raising series B funding. Specifically, the investors have a "funding attraction" factor determined by the business growth (defined as raising follow-on investments) of the startups they backed and the evolution of their early stage portfolio. Therefore, the hypotheses to be tested are the following: companies backed by investors with a high funding attraction factor are (1) more likely to raise follow on funding, but the mechanism fades in its importance as the quality of the company becomes clearer, and (2) more likely to raise higher amounts of capital in the follow-on round.

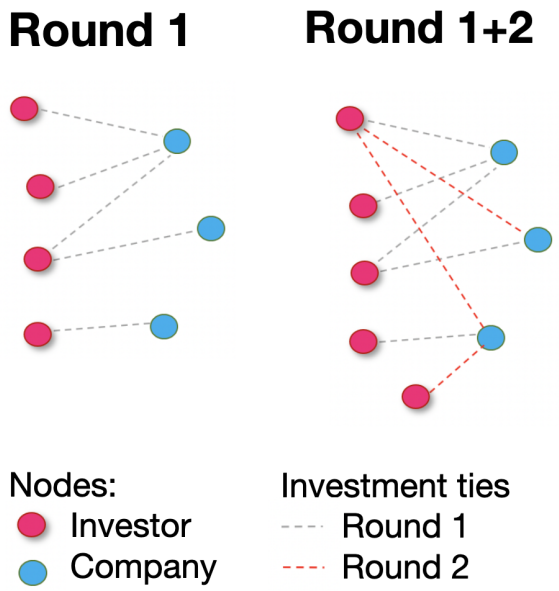


Figure 1: Dynamic Bipartite Network: Theoretical Construction

3 Network of Influence: Methodology

3.1 Dynamic Bipartite Network

The investments dynamics that occur between early investors and companies have a structure than can be naturally described by means of a bipartite network. As displayed in Figure 1, the vertices are of two types: early investors i and companies c. For every investor i and company c we draw an oriented edge from i to c, if investor i participates in a funding round of company c. In this case, nodes i feature exclusively outgoing edges, called out-degrees k_i, while nodes c display only incoming edges, called in-degrees k_c.

The investments are originated by the fact that every company needs funding in order to develop their products and move forward in their business cycle. Investing in a venture is possible only when a company decides to open the funding round. In network terms, an edge between i and c can be set only when c is actively looking to receive it and the window of opportunity is constrained in time. In the bipartite network we built, each company is open to investments twice (round 1 and round 2), while investors can participate multiple times according to their endowment. In studying the passage from pre-series A to series A, round 1 comprises all the pre-series A rounds (angel, pre-seed, seed), while round 2 corresponds to series A. Similarly, when the network is

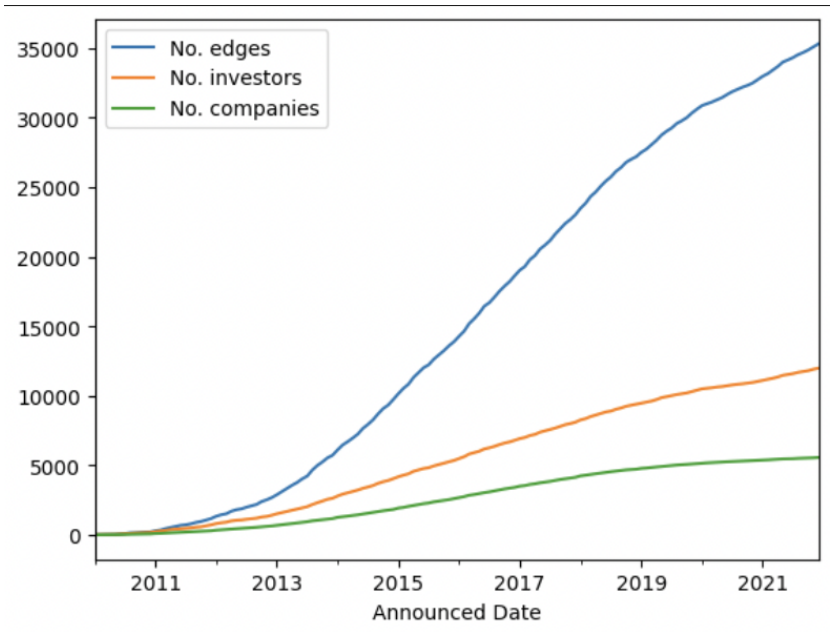


Figure 2: Dynamic Bipartite Network: growth in number of funding deals (edges), of companies, and of investors (nodes), over time

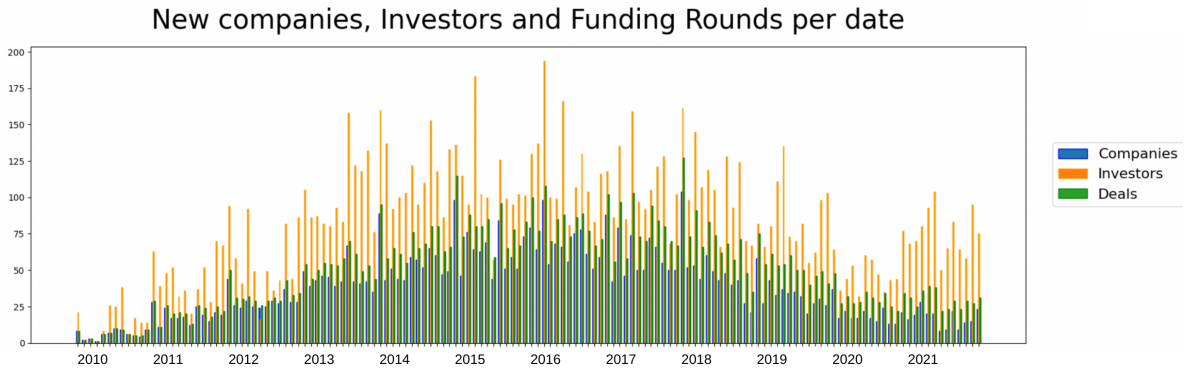


Figure 3: Dynamic Bipartite Network: monthly appearance of edges (deals) and nodes (companies and investors)

re-created to analyse the advancement from series A to series B, round 1 displays the seeding round plus series A while round 2 represents series B.

The network grows in a discrete fashion from 2010 to 2021. Funding interactions and new nodes are inserted at the beginning of every month, creating a snapshot sequence of graphs. The growth in number of nodes and edges can be seen in graph 2. The discrete time dynamic allows to represent complex temporal patters not captured by static networks, as burstiness, memory, and non-stationarity. Burstiness refers to the tendency of nodes or edges to appear in bursts over time. Indeed, as displayed in 3, the deal flow is uneven during the 11 years considered, with abrupt chances occasionally occurring between contingent months. The discrete time dynamic model developed allows us to explicitly include the edges resulting from the funding rounds when entered, thus modelling bursty activity. Another important pattern that can be captured by discrete time dynamic models is memory, which refers to the tendency of nodes or edges to be influenced by their past states. In our set-up, the memory aspect is especially important, as it plays a key role in determining who is attracting who and who is following who. Lastly, non-stationarity refers to the fact that the statistical properties of a system may change over time. The situation is often observed in systems that undergo phase transitions, which in our analysis are the passages between funding stages. Indeed, our model allows the network structure to change over time, therefore allowing for the quantification of the adiabatic approximation. The discrete time dynamic model explicitly ensembles complex temporal patterns that are not captured by static networks, and has the potential to significantly enhance our understanding of the dynamics underlying early stage funding.

3.2 Novel Funding Attraction Index

To analyse the role of investors and specifically their ability to catalyze investments on the companies they back, we developed a novel network measure, able to capture the overtime evolution of an investor's influence over its community. We based the creation of the innovative index on the topological overlap matrix (Barabási et al., 2002), a similarity matrix built on the overlap of neighbors between two nodes. Indeed, the matrix reflects how many neighbour nodes are connected to both the vertices and weights them by the minimum number of degrees between the vertices. It measures pairwise similarities of nodes: the greater the similarity of two objects, the greater the value of the measure.

In its original version, the matrix is static, refers to a univariate network, and to pairwise relations. Initially, we built the matrix in a bivariate setting, having the investors' nodes i in the axes and considering companies' nodes c as the connecting neighbours.

Consequently, we added a temporal dimension, shifting the values of one axis to the following moment in funding, therefore including investors j, which participated at the second round. Lastly, the communal aspect was taken into account, enabling the computation of a score which captures a one-to-many leader-followers dynamic, defined as the influence that the single investor i has on the community of investors j. The topological overlap measure would then be represented by the following formula:

$$T_i = \frac{1}{n} \sum_{j}^{n} \frac{c_{ij}}{min(k_i^{[1]}, k_j^{[2]})}$$

where c_{ij} is the number of companies to which both investors i and j are connected, $k_i^{[1]}$ is the node degree of investor i at round 1, and $k_j^{[2]}$ is the node degree of investor j at round 2.

Applied to our case, the T index considers each investor at round 2, j, which is funding the company c in which i invested at round 1 and scale its investment by the minimum number between investments in first rounds made by i and investments at second round made by j. Consequently, the ratios attributed to all investors j related to investor i are averaged, resulting in the T_i *score*. The T_i *score* is to be intended as an intermediary step between the topological overlap matrix and the funding attraction index M_i, as it lacks the economic significance of the latter, but it enriches with the bipartite, the temporal, and the communal dimension the former.

Based on it, we introduce the *funding attraction index (M_i)* that quantifies the number of times that the investments of an early investor at round 1 have been replicated by other early investors at round 2. Therefore, an investor i participating in round 1 is assigned a score based on how many investors j participating at round 2 are funding the company c in which i invested at round 1, scaled by the number of investments made by i at round 1. In formulas:

$$M_i = \frac{\sum_j^n c_{ij}}{k_i^{[1]}}$$

The result is then normalized to assure that the funding attraction index lies in the range [0, 1]. By itself, measure M_i quantifies the capability of an investor i to attract, at a subsequent round, further investors in the company they invested in. It is defined as an investment multiple assigned to the investor and perfectly captures the one-to-many leader-followers dynamic which it meant to ensemble.

An alternative index, considering how fully an investor at round 2 copied the behavior of an investor at round 1, was considered, since it would represent the leader-follower dynamics in a more literal fashion. However, such an index would need to focus on

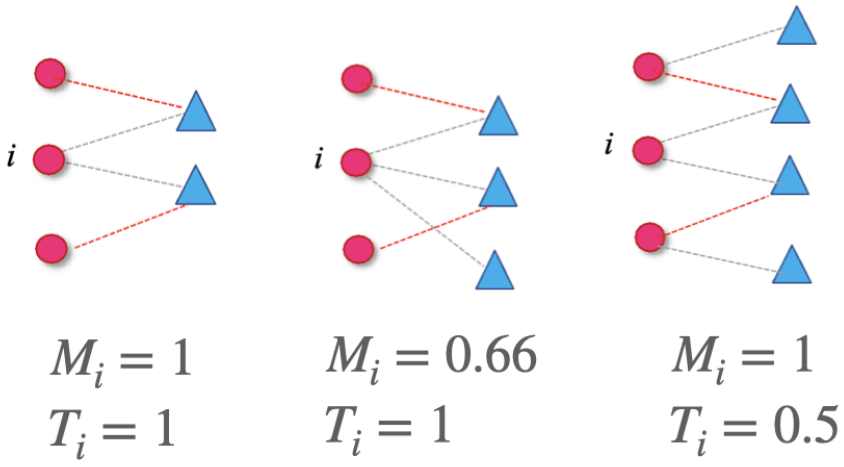


Figure 4: Scores: computation example

pairwise relations by definition, thus preventing its application on the whole investor community, which is the focus of the current analysis.

Figure 4 offers a graphical comparison of how to compute the scores, while 4 displays the evolution of M_i and T_i over time for the fund "Airbus Ventures". According to the temporal choice in methodology, the indexes are computed at the beginning of every month, thus gradually updating the changes in the score, contrarily to the most common reputation measures used in the literature.

3.3 Transitive Fitness Measure

In complex network theory, the fitness model is a model of the evolution of a network: how the links between nodes change over time depends on the fitness of nodes, defined as an intrinsic characteristic of the node itself (Caldarelli et al., 2002).

We introduce a novelty in fitness models by determining the fitness measure (Z) of a node c through the characteristics of the nodes i to which c was connected at the previous point in time. From a financial perspective, Z_c represents the appeal to investors. The model is described by Figure 6 and the following formula:

$$Z_c = f(V_i, t-1)$$

In our analysis, when company c raises the first investment round, we assign it the probability of attracting investors as a function of the specific investors nodes' properties V_i which previously funded it, i.e. are already connected to it. Therefore, the measure

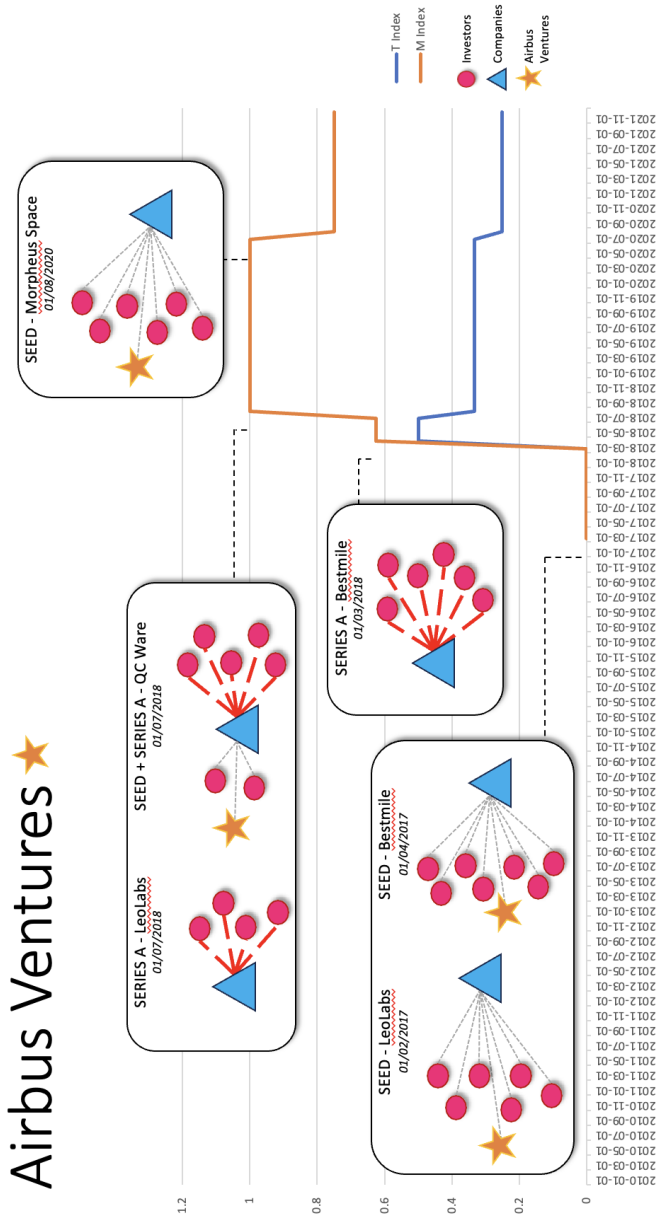


Figure 5: Scores evolution over time for the selected investor "Airbus Ventures"

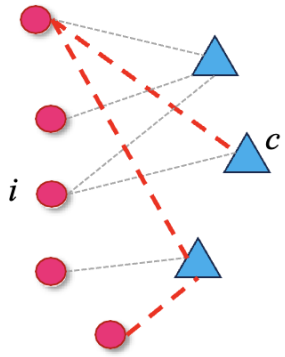


Figure 6: Transitive Fitness Model

$ZFAI_c$ represents company c easiness of attracting investors at the following round. In the key version of the proposed model, the intrinsic properties measure V_i of the investors' vertices i is the funding attraction index introduced in the previous section. We considered the maximum value among the funding attraction scores FAI_i belonging to nodes i connected to node c, thus equalizing the funding attraction score related to the company c to the one of the most attractive investor backing it. The rational underlying this choice was driven by the fact that often the most prominent participant in a group is the one symbolizing it, thus the same is applied to publicly disclosed investors.

It is worth stressing that the moment in which the score is computed is especially important to prove the relevance of specific investors involvement and disentangle it from the characteristics, actions and business success of the company itself. By assigning the funding attraction score to the company at the very same time in which the seed round is raised, the consequent individual occurrences and actions of the company or the investors are excluded from the analysis (industry wide instances are controlled through fixed effects).

In order to make the specifications including the novel score comparable to the state of the literature in the field of venture capital reputation, alternative measures are computed considering classical reputation measures as investors nodes' properties V_i. Namely, we considered experience (proxied by number of investments (among many, Atanasov et al., 2012) or years from the first investment (Nahata, 2008)) and success (proxied by the rate of successful exits (among others, Nanda et al., 2020)). Following the same rationale as in the funding attraction score scenario, the maximum value among investors backing the company was associated to venture c. As regards industry notoriety, the company takes value 1 if one of the investors backing it in seed funding

appeared in the prestigious Forbes' Midas list.

4 Disentangling Selection and Treatment: Methodology

Are influential VCs simply better at selecting in which companies to invest in or is the company's better performance a consequence of their influence? To disentangle the effects, we develop on the Sorensen-Heckman approach introduced by Bottazzi et al. (2008 and 2016). By using as the unit of observation the potential (realized and unrealized) matches between an individual company and an individual investor, the model estimates a selection equation (does the VC invest in the company?) and an outcome equation (is the company successful? Sorenson (2007) considers success as valuation, Bottazzi et al. (2008 & 2016) as exit and survival). We are interested in understanding the effects of influence on a company business life, thus we consider as outcome measure the company's access to the following funding round (Werth and Boert, 2011). Thus, we focus on three jumps: the passage from seed to series A, from series A to series B, and from series B to C. We take as an example the first rounds, the passage from seed funding to series A funding.

The first selection equation will be the following:

$$dealhappened_{seed} = \beta_0 + \beta_1 ZFAI_{seed} + \underline{\beta_2 TotalExperience_i} + \underline{\beta_3 IndustryExperience_i} + \beta_4 NumberofEmployees_c + \beta_5 EstimatedRevenueRange_c + \beta_6 Industry_c + \beta_7 DealYear + \epsilon_{c,t}$$

The first selection equation estimates the likelihood of being selected for seed funding. It allows us to compute the Mills ratio, which accounts for the unobservable selection effect. The unit of observation is constituted by the potential (realized and unrealized) matches between an individual company and an individual investor. Our exclusion restriction is composed of the two variables $TotalExperience_i$ and $IndustryExperience_i$.

Reducing the dataset to the seed deals which were finished, we insert the Mills ration into the second selection equation, outlined by the following formula:

$$dealhappened_{seriesA} = \beta_0 + \beta_1 ZFAI_{seed} + \underline{\beta_2 InverseMillsRatio_{seed}} + \underline{\beta_3 TotalExperience_i + \beta_4 IndustryExperience_i} + \beta_5 NumberofEmployees_c + \beta_6 EstimatedRevenueRange_c + \beta_7 Industry_c + \beta_8 DealYear + \epsilon_{c,t}$$

While this equation is a results itself, as it disentangles the effect that the FAI_{seed} has between being selected in seed and in series A, it is also considered a selection equation. Indeed, from it we derive the Mills ratio for series A, which accounts for the

unobservable selection effect faced by companies present in the series A sample. The inclusion of both the inverse Mills ratio and the exclusion restriction might create issues which we discuss in the sensitivity analysis section. Further, we developed the model into a more complex specification, as more data become available when the sample is reduced to exclusively deals that were realized. The specification is as follows:

$$dealhappened_{seriesA} = \beta_0 + \beta_1 ZFAI_{seed} + \underline{\beta_2 InverseMillsRatio_{seed}} + \beta_5 NumberofEmployees_c + \beta_6 EstimatedRevenueRange_c + \beta_7 Industry_c + \beta_8 DealYear + \beta_4 LogSeedMoneyRaised_{c,t} + \beta_5 NumberofSeedInvestors_{c,t} + \beta_5 NumberofFundingRounds_{c,t} + \epsilon_{c,t}$$

Lastly, we compute the effect of $ZFAI_{seed}$ on money raised in series A. In this case, we face the two-step selection for which we iterate the selection equations in seed and series A. Indeed, to take into account for hidden selection effects, we insert the estimation of the second selection equation in the new specification.

Thus, the money raised in series A equation is:

$$moneyraised_{seriesA} = \beta_0 + \beta_1 ZFAI_{seed} + \underline{\beta_2 InverseMillsRatio_{seed}} + \underline{\beta_2 InverseMillsRatio_{seriesA}} + \beta_5 NumberofEmployees_c + \beta_6 EstimatedRevenueRange_c + \beta_7 Industry_c + \beta_8 DealYear + \beta_4 LogSeedMoneyRaised_{c,t} + \beta_5 NumberofSeedInvestors_{c,t} + \beta_5 NumberofFundingRounds_{c,t} + \epsilon_{c,t}$$

The whole process is then repeated, through the adequate adaptation of the outlined formulas, for the passage from series A to series B and from series B to series C.

5 Dataset

5.1 California 2010-2021: data collection and cleaning process

The data has been collected from the academic crunchbase.com Web API. Crunchbase was founded in 2007, by TechCrunch, an online news outlet focused on the technology industry, as a way to easily keep track of startups in the news. The data provided by the Crunchbase website are manually recorded and managed by several contributors (e.g., incubators, venture funds, individuals) affiliated with the Crunchbase platform. Moreover, the data are further enriched by Web crawlers that scrape the Web, on a daily basis, in search for news about IPOs, acquisitions, and funding rounds, and checked by the platform's data analysts.

To date, Crunchbase is widely regarded as the world's most comprehensive open data set about startup companies (Retterath and Braun, 2020). It contains detailed information on organizations from all over the world discerning them into four categories, namely companies, investors, schools, and groups. In addition to people's professional activity, the data track information about their educational paths, and consequently their access to academic knowledge, and the participation to event elated to the startup and venture capital ecosystem. Having a profile in Crunchbase is a strong signal of legitimacy for entrepreneurs and their companies. Entrepreneurs, companies, funds, and investors listed in Crunchbase are motivated to ensure that their profiles are accurate.

The data is conveniently timestamped and hyperlinked, permitting capture of network structure and dynamics. Indeed, organizations and people are uniquely identified by alphanumeric IDs. All data are time-stamped, and an accurate reconstruction of historical events was made possible by the use of trust codes, i.e., numerical codes provided by Crunchbase to indicate the reliability of a certain timestamp. The timestamps indicate the dates of foundation, funding rounds, acquisitions, and IPOs, as well as the start and the end times of job roles.

The main issue with data is that a large number of entries related to companies contains very limited information. Accordingly, we cleaned the data keeping only the organizations and people for which enough information is provided, and for which such information is reliable. We solved and removed inconsistencies such as funding rounds registered at a date prior to the company's foundation. In these cases, we retained the most reliable information according to the trust code value. Inconsistencies have been removed by adopting a strong self-penalising data cleaning strategy. As a result, we do not retain in the graph links whose trust code value cannot be determined in a reliable way. This approach to data cleaning strengthens the validity of our results because it ensures that companies and investors do not get assigned a different position in the network as a result of connections that were forged in the database. In this way we avoid biases that could artificially inflate the success rate of the method, and accordingly our results can safely be seen as conservative lower bounds. As a note, we would like to point out that information about startups is most typically made available when the startup is in the process of raising funds. Hence, Crunchbase is limited primarily to startups that have already received funding or are actively looking for them.

The articulate cleaning process limits the number of companies, financial organizations and people. Indeed, from the Crunchbase platform, we extracted the data related to startups founded in California between 2010 and 2017 which received at least one round of funding. For each organization we added additional information such as details on firms' foundation dates, locations of the firms' headquarters, industry, estimated revenue range, number of employees, founding rounds, acquisitions, and IPOs.

Considering the funding rounds, we minutely collected all the deals companies raised between 2010 and 2021. For each round we identified all the financial organizations and angel investors involved and the general partners responsible for the deal, together with other information, as amount raised and funding stage. Consequently, we access information on all the investors involved, being them angel investors, accelerators or VC firms, with particular attention regarding their funding date, number of investments made, companies in the portfolio, investments lead, IPOs, and acquisitions, as well as more general information as stage of expertise, description, and headquarter location. As angels and individual investors are concerned, information regarding gender, job title in the current and previous jobs, social media accounts, and attended schools has also been collected.

5.2 Variable definition and correlations

5.2.1 Company founding and financing variables

We focus on startups founded in California as the State is the most vibrant area in venture capital investments worldwide, the home to the Silicon Valley, and the cradle to many of the most well known startups' success stories (Dorfman, 1983; Gompers et al., 2016). Focusing on a single area can eliminate geographically-related heterogeneity. Our primary dependent variable, **Series A**, is binary, taking value of 1 if the company raises series A funding and zero otherwise: 38% of the firms in our cleaned sample reach series A funding. The proportion is higher (51%) if we instead consider companies raising series B funding after having reached series A.

A further dependent variable, **Log Money Raised in Series A (or Series B, or Series C)**, is taken from Crunchbase - we use the natural logarithm of a startups' Series A funding to account for a right-skewed distribution. Lastly, we consider **Log Number of Investors in Series A (or Series B, or Series C)**. Similarly, the raw variable is available on Crunchbase, and it has been processed in order to account for the left-skewed distribution.

Since we collected deals over 11 years, we control for time considering **Year of Seed (or Series A, or Series B) Funding**, representing the year in which the company raised its first funding round, in the form of fixed-effects. The variable brings the added benefit of offering a shield against possible macroeconomic events affecting the venture capital industry as a whole in a given period, and therefore introducing biases in the analysis. Similarly, we controlled for the **Number of Seed (or Series A, or Series B) Investors** and **Seed (or Series A, or Series B) Money Raised**. Indeed, a higher amount of funding or of investors in the seed rounds can also positively influence the

perception of the startup in the investment community and attract attention from later-stage investors, as shown in the correlation table 10. It can signal that the startup has gained traction, demonstrated promising results, and is worthy of further investment. Contrarily, high participation, if not balanced by a reasonable amount of capital raised and especially corresponding equity distributed, can unintentionally exclude investors seeking high returns, as the situation lacks of financial incentive and likely leads to the impossibility of impacting management in a later moment. Crowded deals makes harder for funds to outperform and, in parallel, to be in the spotlight in the LPs seek.

Additionally, we also add **Industry Fixed Effects** to account for industry-specifics trends, **Estimated Revenue Range** to consider the company's economic and financial relevance and **Number of Employees Range** to recognize in the analysis the firm's size. Lastly, we control for **Number of Funding Rounds** (minus 1, in order to exclude the current round from the count).

5.2.2 Network variables

The primary independent variable and most prominent network measure used in the analysis is the **Funding Attraction Factor** (Z_i) which is assigned to the company as a function of the *Funding Attraction Index* of its seed (or series A, or series B) investors at the day preceding the investment in the company. In its normalized form, takes values spanning from 0 to 1. Its characteristics and peculiarities are described extensively in sections 3.2 and 3.3. Additionally, another unconventional network measure, based on the adaptation of the **Topological Overlap Score** (T_i), is used. Also its specifics are described in sections 3.2 and 3.3.

For the sake of providing a complete comparison with traditionally used network metrics in financial networks, other more widely used centrality measures are computed. Specifically, we projected the bipartite connection in a univariate setting, creating two univariate networks: the investors' networks, in which investors constitute the nodes and are connected by having a company in common (Hochberg et al., 2007, among others); and the company ones, in which firms are the nodes and are tied by having an investor in common (Bonaventura et al., 2020, Wang et al., 2023, among others). As occurs in the bipartite setting, also the new networks are updated at the beginning of each month. In both networks we computed the betweenness centrality measure assigned to each node, resulting therefore in **Company Betweenness Centrality** and **Investor Betweenness Centrality**. In network science, betweenness centrality measures the extent to which a vertex lies on paths between other vertices. Indeed, in financial networks, it is usually regarded as a measure of information transfer in the net, with nodes displaying higher values having enhanced access to information, by virtue of their control over information passing between others. High betweenness nodes are also

the ones whose removal from the network will most disrupt communications between other vertices because they lie on the largest number of paths taken by messages.

Lastly, the traditional **Company Degree Centrality** and **Investor Degree Centrality**, quantifying the number of nodes connected to the node for which the score is computed, are measured. In network science, degree centrality is a classic measure of general relevance, as it computes the number of directly linked nodes.

5.2.3 Investors related variables

For each investors in our dataset we focused on creating variables in line with the current literature on reputation and with contemporary industry practices. Experience is proxied by the variable **Number of Investments** (Atanasov et al., 2012), which considers the total number of investments made by a fund over its business life, and the variable **Years from First Investment** (Nahata, 2008), self-explanatory, updated yearly. Similarly, success is proxied by **Rate of Successful Exits**, computed as number of IPOs made by the fund over total number of investments.

Additionally, the desire to take into account industry notoriety required to match the investors in our dataset with Forbes' Midas investors list. The **Midas Investors** variable is binary and takes value of 1 if the investor or the fund appeared in the list in the year prior to the deal we observe in our dataset. All lists published between 2010 and 2020 are considered. The variable is a novelty in literature, and appears to have low or no correlation with the other variables used in the model, as displayed in Table 10.

5.2.4 Summary statistics

In the initial phase of the study, all the organizations and investors involved in angel, pre-seed, seed, and series A rounds have been included in the construction of the networks. The total number of pre-IPO equity funding rounds amounts to 32,392, of which 17,742 are pre-series A (angel, pre-seed, and seed rounds) and 5,584 are Series A. Respectively, they corresponded to 11,585 and 4,742 unique companies. The control scores included in the first control sample are computed on the full sample. Indeed, it could be argued that by inserting the most exhaustive possible information in a network, the network becomes more representative of the real-world and has the ability to grasp the dynamics occurring. By capturing a higher level of detail, the network can reflect more accurately the behaviors and outcomes observed in the real world. But, if inaccurate information is included in a network, it can significantly affect the validity and reliability of the network model and its ability to capture real-world dynamics.

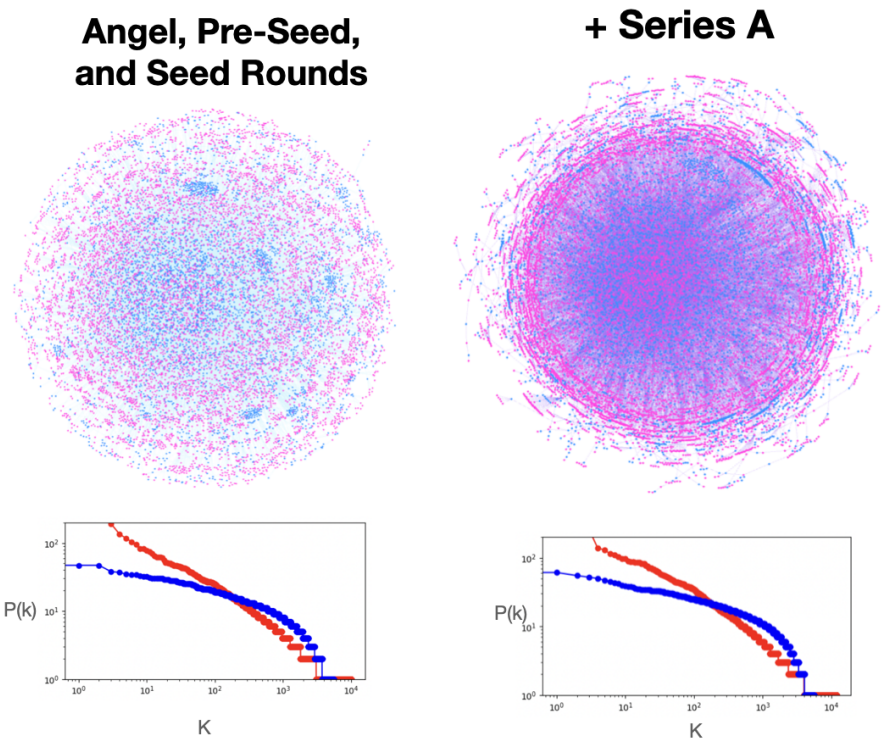


Figure 7: Network representations of Seeding Funding (Angel, Pre-Seed, and Seed) and Startup Funding (Series A), with nodes being investors (in red) and companies (in blue) connected by the investment ties, and the respective degree distribution

Defective data can introduce biases and distortions into the network, leading to erroneous interpretations of relationships, interactions, and patterns. This can hinder our understanding of the system being studied and impair the network analysis. Furthermore, inaccurate information can have a cascading effect, where errors in one part of the network can propagate to other areas, leading to a domino effect of inaccuracies, affecting the connectivity, structure, and dynamics of the nodes and edges. Indeed, to mitigate the impact of inaccurate information, the data cleaning strategy was executed.

Following the articulated and strong self-penalising data cleaning strategy detailed in the previous section, the final dataset resulted in a total of 5,564 unique companies involved in pre-series A rounds, of which 1,762 raised series A. Investors involved are 10,167 and 11,988 in seed funding and series A, respectively. Pre-series A rounds amount to 9,016 and reach the amount of 10,778 when series A are considered. They correspond to 25,391 and 33,712 investor-company links. The network representation are displayed in figure 7, with the corresponding degree distribution, while other statis-

tics are available in figure 8, focused on deals, and 9, focused on investors.

Companies in our sample take on average two years to raise their first funding round and, as described in the methodological section, investors build their scores over time. Therefore, bearing in mind the data collection includes companies founded starting from 2010, the final dataset is formed by companies which raised series A funding from 2014 onwards. The time shift was deemed necessary to allow investors' and companies' initial score computation on a real-world clean proprietary dataset, as to allow initial score computation on the previous data. Moreover, only investors which can be considered as professional investors, here defined as investors who participated in at least 5 funding rounds (Gafni et al., 2021), are considered. Indeed, the presence of occasional investors could impair the results, assigning unrealistic high attraction scores to casual angel investors who had the luck to pick a performing company early on, consequently undermining the final analysis. Similarly, companies which raised investments exclusively by occasional investors, considered as family-and-friends rounds, could not have a score assigned and thus had to be excluded from the final analysis. Lastly, the funding attraction score assigned to investors and consequently to companies is not stationary: as explained in the methodological section, it changes according to the new investments entered in the network and it incorporates a temporal dimension. Lastly, companies which did not disclose the fundamental variables considered in the analysis also had to be removed for the sake of adding fixed-effect and ensuring comparability *ceteris paribus*. The final dataset used to gain the empirical estimates resulted in 2,126 companies, as outlined in the summary statistics displayed in panel B, table **??** and **??**.

Independent Variables	Not Raised Series A (0)			
	average	mean	median	max
Network Score Max Norm	1,46	0,00	1,09	15,00
F Score Max Norm	0,10	0,00	0,06	0,92
T Score Max Norm	0,37	0,00	0,20	1,00
Network Score Betweeness Max	0,04	0,00	0,01	0,14
IPO Inv	0,13	0,01	0,12	0,14
Age Norm	15,27	1,00	12,00	187,00
Company New Score Betweeness	0,00	0,00	0,00	0,01
Number of Investors r1	5,21	1,00	3,00	61,00
Money Raised r1	2.819.152,03	5.000,00	1.120.000,000	700.000.000,000
Numb of Investments Norm	1.269,36	0,00	286,50	10.312,00
N obs	1.334	1.334	1.334	1.334

Table 1: Descriptive Statistics Panel A

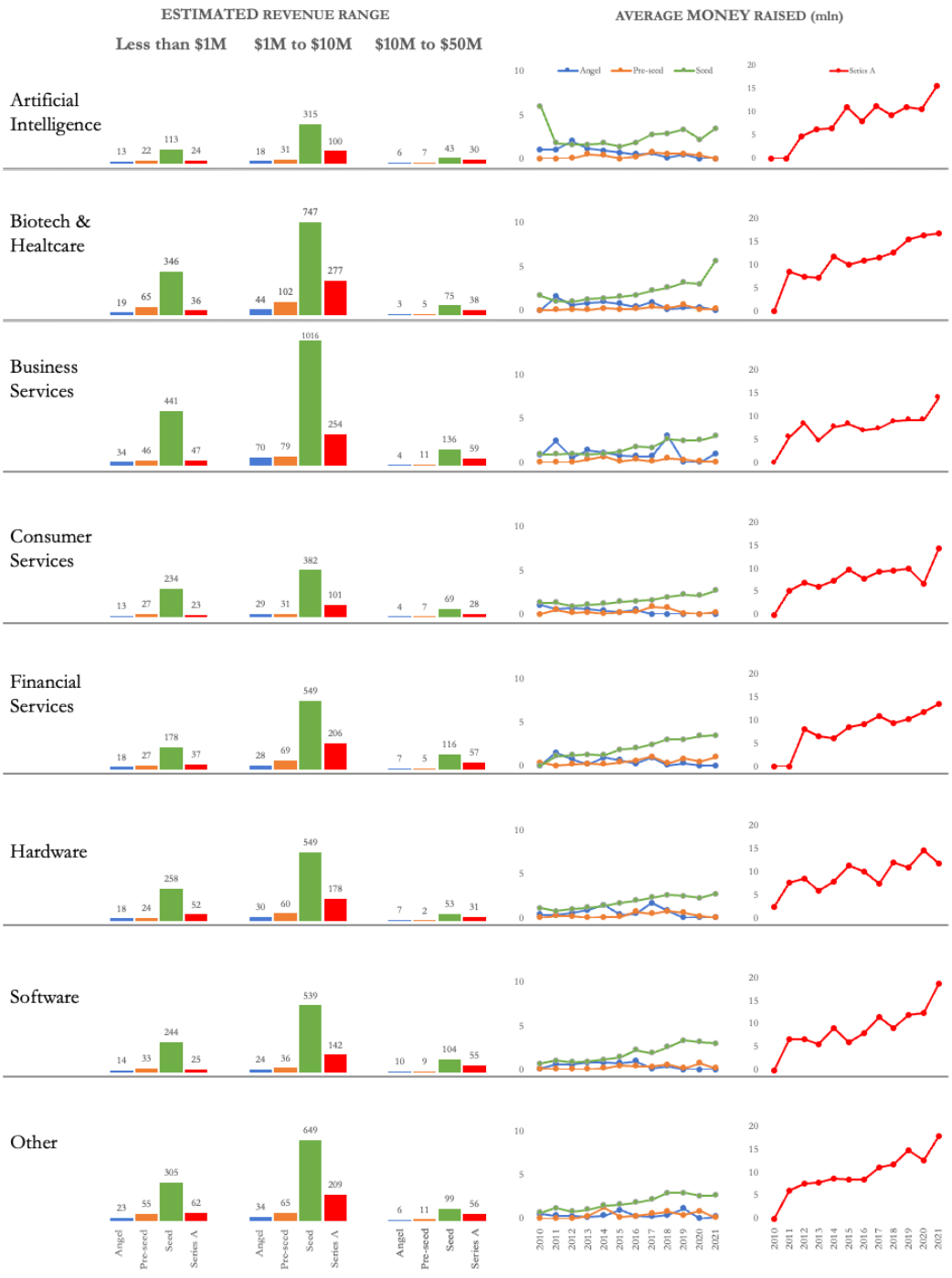


Figure 8: Descriptive Statistics Panel A

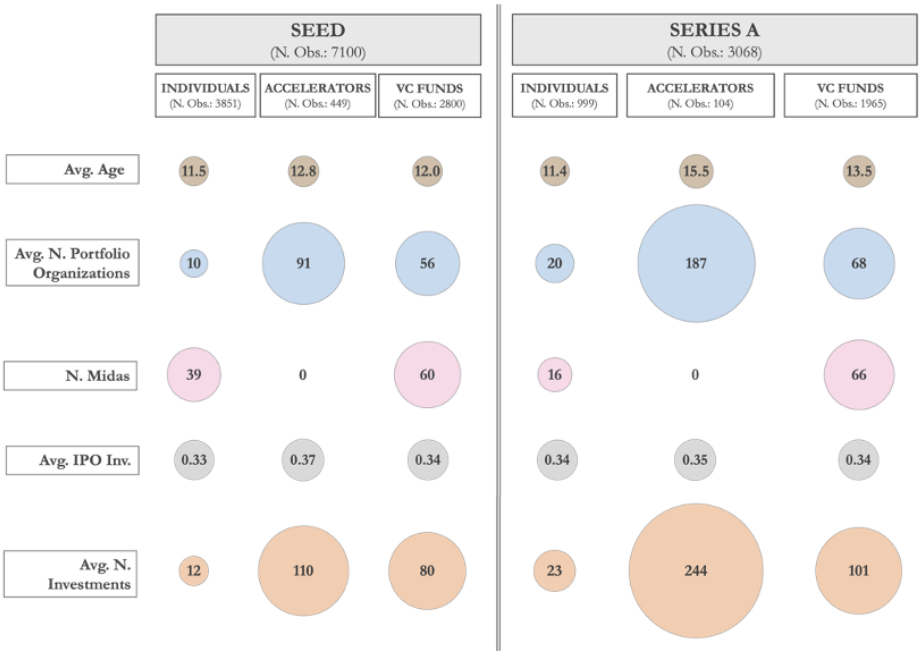


Figure 9: Descriptive Statistics Investors

Independent Variables	Raised Series A (1)			
	average	mean	median	max
Network Score Max Norm	2,14	0,00	1,88	16,00
F Score Max Norm	0,16	0,00	0,11	0,94
T Score Max Norm	0,43	0,00	0,33	1,00
Network Score Betweeness Max	0,03	0,00	0,01	0,15
IPO Inv	0,16	0,01	0,15	1,00
Age Norm	17,87	2,00	12,00	148,00
Company New Score Betweeness	0,00	0,00	0,00	0,02
Number of Investors r1	7,32	1,00	6,00	49,00
Money Raised r1	3.710.965,60	25.000,00	2.944.000,00	222.234.000,00
Numb of Investments Norm	960,87	0,00	325,00	8,650,00
N obs	824	824	824	824

Table 2: Descriptive Statistics Investors

Further summary statistics are reported in panel C, which considers the jump from series A to series B, Tables 13 and 14, in the appendix.

		Network Score Max Norm		Money Raised r2		Number of Investors r2	
		N Obs	average	N Obs	average	N Obs	average
Estimated Revenue Range	Less than 1M	574	1,56	112	10.123,153	574	1,09
	1M to 10 M	1.395	1,74	554	11.859.270	1.395	1,97
	10M to 50 M	189	2,11	122	12.245.808	189	3,87
No. Employes	1-10	732	1,39	50	9.070.685	732	0,32
	11-50	952	1,82	361	9.858.970	952	1,92
	51-100	210	2,07	151	12.680.107	210	3,46
	101-250	210	1,82	361	9.858.970	952	1,92
	251-500	54	1,76	46	17.332.491	54	5,11
Industry	ArtificialIntelligence	147	1,71	61	10.593.852	147	2,95
	Biotech and Healthcare	350	1,58	141	13.664.602	350	1,98
	BusinessServices	411	1,73	104	9.005.015	411	1,21
	ConsumerServices	175	1,79	49	9.037.971	175	1,61
	FinancialServices	261	2,11	122	11.481.253	261	2,84
	Hardware	262	1,52	96	13.655.051	262	1,62
	Other	314	1,75	138	12.492.657	314	2,45
	Software	236	1,62	75	10.538.894	236	1,56
Midas	No	2.001	1,66	688	11.465.316	2.001	1,77
	Yes	157	2,45	100	13.096.802	157	3,45
Number of Investors r1	Q1	697	0,61	163	12.665.499	697	0,86
	Q2	431	1,37	156	11.403.217	431	1,68
	Q3	514	2,04	222	11.657.757	514	2,33
	Q4	516	3,20	247	11.200.070	516	3,04
Money Raised r1	Q1	551	0,89	57	10.429.592	551	0,46
	Q2	529	1,40	128	7.789.844	529	1,16
	Q3	539	2,03	302	11.039.282	539	2,88
	Q4	539	2,57	301	14.193.912	539	3,09

Table 3: summary statistics

Lastly, information on the correlation among independent variables are available in the Table 10.

	(1)	(2)	-3	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Funding Attraction Index (Norm.)																		
Num investments (Norm.)	0.1035																	
Success Rate (IPO/Tot.Inv.)	0.1136	-0.1205																
Midas Investor	0.1024	0.0657	0.1439															
Years From First Investment (Norm.)	0.0208	0.0596	0.2418	0.1774														
Money Raised Seed	0.0062	-0.0417	-0.0127	0.0020	0.0142													
Number of Investors Seed	0.4800	0.1195	0.0761	0.0081	0.0161	0.0220												
Number of Funding Rounds	0.1535	0.0298	0.0589	0.0290	0.0501	-0.0034	0.3398											
Series A	0.1674	-0.0892	0.1406	0.1581	0.0785	0.0231	0.1806	0.4894										
Money Raised Series A	0.0269	-0.0309	-0.0976	0.0470	0.0429	0.0765	-0.0298	0.0174										
Number of Investors Series A	0.1966	-0.0449	0.1081	0.1301	0.0523	0.0203	0.2631	0.3706	0.7205	0.0750								
Time (Months)	-0.0649	-0.0283	-0.0631	-0.0967	-0.0480	-0.0074	-0.0666	-0.1345	0.0613	0.0796	-0.0990							
Topological Overlap Score	0.3904	0.0536	0.0507	0.0115	0.0503	0.0045	0.3629	0.1336	0.0506	0.0027	0.0888	-0.0148						
Company's Network Betweeness	0.1930	0.2312	0.0516	0.0039	0.0138	-0.0098	0.5285	0.2453	0.0863	-0.0519	0.1540	-0.0234	0.2137					
Investors' Network Betweeness	0.1384	0.6700	-0.0528	-0.0073	0.0017	-0.0464	0.2027	0.0108	-0.0775	-0.0357	-0.0170	-0.0030	0.0835	0.3532				
Investors' Eigenvector Centrality	0.2739	0.5859	0.0396	0.1096	0.0386	-0.0496	0.2825	0.0321	0.0123	-0.0247	0.0693	-0.0218	0.1076	0.3375	0.8785			
Company's Eigenvector Centrality	0.1739	0.4947	0.0209	0.0662	0.0327	-0.0372	0.2414	0.0037	-0.0146	-0.0302	0.0387	0.0521	0.0809	0.3483	0.7318	0.8536		
Investors' Degree Centrality	0.2322	0.5757	-0.0487	0.0599	0.0353	-0.0371	0.2277	0.0343	-0.0382	-0.0004	0.0195	-0.0309	0.1239	0.1929	0.8630	0.8793	0.7046	
Company's Degree Centrality	0.2645	0.5493	-0.0405	0.0608	0.0432	-0.0286	0.3279	0.0782	-0.0157	0.0053	0.0444	-0.0224	0.1766	0.2776	0.8152	0.8328	0.7459	0.9605

Figure 10: Correlation of Independent Variables

6 Results

The empirical assignment is straightforward: to test the hypotheses that a company getting funding from high funding attraction investors in an investment round is (1) more likely to raise a the following funding round, but the mechanism fades as the company proceeds in its business life, and (2) more likely to raise higher amounts of capital in the follow-on round. The former hypothesis suggests a greater importance of high attraction investors' involvement in the earliest stages of funding rather than later in business life, albeit the latter advocates for their role not only in moving a company forward in financing, but also in accessing higher amounts of capital. Both moving forward in business life and attracting more capital are positive effects under a company's perspective. The section is therefore organized around empirical tables that demonstrate the relationships in univariate and multivariate settings, followed by robustness checks addressing company's quality, most successful investors' preferential access to funding rounds, and geographical biases

6.1 Moving Forward in Business Life

Table 4 examines the companies likelihood of being matched with investors in a round, and the consequent likelihood of accessing the follow on round. Specification (4-1) shows the positive effect that the seed investors' network of influence has on the happening of the match between the investor and a company in the seed round. Further, taking into account the selection effect, specification (4-2) displays the positive relationship between high FAI investor and the access to the following round of funding of the companies they back. The results are statistically significant at a 1% level. The estimated coefficient in (4-1) implies an increase in likelihood of accessing seed funding equal to 6% for a discrete change in the measure of the FAI, while (4-2) shows an increase of 12% in the likelihood of raising series A taking into account the selection effect through the inverse Mills ratio. The magnitude of the effect is confirmed in the complete form of the model, (5-1), with a statistical significance of 5%.

Notice that the observations in the selection equation (4-1) are numerous since they are all the possible matches between companies raising seed funding and investors investing in seed, while the outcome equation (4-2) works exclusively on the realized seed deals. Indeed, the dependent variable in (4-1) is the *seed dummy* (was the seed rounds realized?), while in (4-2) it is the *series A dummy* (was the series A realized?). Both the equations include the investors' overall experience and the industry experience to account for the likelihood of an investor in matching with the company. Results are controlled for estimated revenue range, number of employees, industry group, year

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Funding Attraction Index (FAI)	0.0653***	0.1216***	0.0533***	0.0700***	0.0128*	-0,0682*
	(0.002)	(0.015)	(0.003)	(0.013)	(0.008)	(0.040)
Inverse Mills Ratio		1.3174		0.5835**		-2.9237
		(0.533)		(0.239)		(1.855)
Total Experience	0.0113***		0.0107***		0.0390***	
	(0.0001)		(0.0001)		(0.0390)	
Experience Industry	-0.0041***		-0.0025***		-0.0223***	
	(0.001)		(0.000)		(0.004)	
Controls:						
Year (Previous Round)	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.3052	-0.6095	-3.3612	-1.3165	-2.9796	-0.0483
	(0.024)	(0.119)	(0.017)	(0.113)	(0.030)	(0.130)
Observation Number	30,267,644	18,791	56,390,336	23,585	5,544,965	7,762
R squared	3.886%	27.98%	3.842%	28.04%	4.647%	20.82%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 4: Selection and Outcome FAI

of former funding, number of funding rounds. Differently, in specification (4-1) the observations are the deals (and not the company-investor match as before). Thus, the inclusion of deal specific controls, as money raised and the number of investors in the seed round, is possible.

Moving forward in business life, specification (4-3) shows the positive effect that the series A investors' network of influence has on the investment in a company in series A funding. Further, taking into account the selection effect, specification (4-4) displays the positive relationship between high FAI investor and the access to the following round of funding of the companies they back. The results are statistically significant at a 1% level. The estimated coefficient in (4-3) implies an increase in likelihood of accessing series A funding equal to 5% for a discrete change in the measure of the FAI, while (4-4) shows an increase of 7% in the likelihood of raising series B taking into account the selection effect through the inverse Mills ratio. The magnitude of the effect is confirmed, while slightly increased, in the complete form of the model, (5-2), with a statistical significance of 5%. Specifics on observations and controls are the same as

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Funding Attraction Index (FAI)	0.1218**	0.0842**	0.0229*
	(0.042)	(0.041)	(0.091)
Inverse Mills Ratio	1.0579	0.9215	-0.5710
	(1.495)	(0.881)	(5.887)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (Previous Round)	Yes	Yes	Yes
Money Raised (Previous Round)	Yes	Yes	Yes
Number of Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	-1.5057	-2.5285	-0.4816
	(0.413)	(0.425)	(0.330)
Observation Number	2,635	2,633	1,358
R squared	38.92%	36.76%	25.84%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 5: Access to the following round: FAI &model

the previous case.

The impact of the FAI decreases as we move to the jump from series B to series C, as outlined in specifications (4-5), (4-6), and (5-3). Indeed, the FAI decreases its impact over the likelihood of getting a company selected in series B with respect to the one it had in seed stage and series A: the effect now equals to 1%. However, the outcome equation effect appears to be negative. In the full model specification (5-3), which controls for the large set of variables related to the company, the investors, and the deal, the effect becomes slightly positive (2%), although the level of significance is slightly decreased (to 10%), with respect to the previous specifications ((5-1) and (5-2)). Indeed, the difference remarks the importance of the controls referring to the previous rounds. Lastly, specifics on observations and controls are the same as in the models regarding seed stage and series A.

Further, we observe that the downwards pattern in the effects of the investor's networks of influence on a company's business life is confirmed. The data support our claim that network dynamics can be useful in filling the lack of clarity typical of the early stage investment landscape. Indeed, the correlation between a company's access

to a round of funding and the network of influence of the investors in its previous round is observed to be positively related. Furthermore, it's noted that the influence effect appears to be more pronounced in the initial stages of deals rather than in the later stages. This finding is rationalized by considering that as the company progresses and matures, its quality and potential become more evident, possibly leading to a diminished reliance on investor networks for subsequent funding rounds. In essence, the early stages of a company's development may heavily rely on investor connections and influence to secure funding, whereas as the company evolves and establishes itself, other factors such as performance metrics and market reputation may play a more prominent role in funding decisions. Appendix 1 confirms that the FAI has a higher explanatory power with respect to the VC firms' classic reputation measure employed in the literature, as success and experience (Hsu, 2005; Nahata, 2008; Nanda et al., 2020); to network analysis metrics, as eigenvector and betweeness centrality, as they are able to capture different aspects of knowledge or information transfer; and to industry notoriety, as being mentioned in Forbes' Midas list.

6.2 Raising Higher Amounts of Capital

Table 6 presents the *Money Raised* multivariate regressions. Specification (6-1) shows that the normalized *funding attraction index* is positively associated with the logarithm of the *money raised in Series A*, at a statistically significant level (1%), and implies an increase in money raised of 6% for a discrete change in the measure of the funding attraction index. The effect descreases to 4% and 3% when the logarithm of the *money raised in Series B* and *Series C* are considered, respectively in specifications (6-2) and (6-3).

The economic significance of the FAI effect in the two models equals 10%. Note that the observations included are the deals for which the amount of money raised was disclosed. Further, the inverse Mills ratios included in the models of table 6 are the ones computed in through the three specifications of 5. Taking as an example case (6-1), the selection effect accounts for the inclusion of the observation in the sample which received series A funding, which was already skimmed from the sample which received seed funding. Indeed, the results confirm that the influence effect does not vanish, but slightly fades as the company becomes more autonomous in its ability of moving forward in business life and it is praised by the market with subsequent conspicuous funding rounds.

	Multivariate Regression		
	(1)	(2)	(3)
Money Raised in:	Series A	Series B	Series C
Funding Attraction Index (FAI)	0.0646***	0.0427*	0.0312**
	(0.024)	(0.014)	(0.034)
Inverse Mills Ratio 2	-0.9569**	0.0886*	0.0180
	(0.395)	(0.092)	(0.806)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Number of Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	16.7608	16.1207	16.8252
	(0.307)	(0.242)	(0.673)
Observation Number	935	935	523
R squared	26.3%	25.9%	35.9%
Adjusted R squared	23.4%	23.0%	31.1%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 6: Raising Capital: FAI

6.3 Robustness to Quality Bias: Company Fixed Effects and Pre-Money Valuation

	Outcome OLS	Outcome OLS
	(1)	(2)
Following Round:	Access	Money Raised
FAI	0.1291***	0.0474***
	(0.004)	(0.008)
FAI x seed	0.0860***	0.0640***
	(0.003)	(0.005)
FAI x series A	0.0860***	0.0640***
	(0.003)	(0.005)
FAI x series B	-0.0494***	0.0380**
	(0.006)	(0.010)
Inverse Mills Ratio	0.1781	-0.0269
	(0.216)	(0.019)
Organization FE	Yes	Yes
Controls:		
Year	Yes	Yes
Industry Group	Yes	Yes
Estimated Revenue Range	Yes	Yes
No. Investors (previous rounds)	Yes	Yes
Money Raised (previous rounds)	Yes	Yes
Number of Rounds	Yes	Yes
Number of Employees	Yes	Yes
Stage of Investment	Yes	Yes
Constant	9.6865	26.3544
	(19.009)	(7.283)
Observation Number	6,624	24,110
R squared	95.5%	50.3%
Adjusted R squared	90.3%	530.%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively		

Table 7: Company Fixed Effects

We introduce several sections to address possible biases that might affect the analyses. We start by further addressing the elusive nature of a company's quality, which poses a formidable challenge owing to its inherent unobservability. The intricacies associated with evaluating the quality of a firm, particularly in its early stages, render it a latent and multifaceted construct that eludes direct measurement. Unlike readily observable financial metrics or market indicators, the underlying attributes that signify

a company's quality, such as the efficacy of its management team, the potential for innovation, or the scalability of its business model, are often concealed from external observers. The opaqueness surrounding these vital dimensions of quality impedes researchers from capturing a comprehensive and accurate representation of a company's true worth. Consequently, the unobservable nature of company quality introduces a layer of complexity into VC research, complicating the establishment of causal relationships and potentially confounding empirical analyses. On top of introducing the Sorenson-Heckmann model, we further tackled the issue in two ways: by introducing company fixed effects (Table 7), which allow us to observe the same company through time, and by introducing a pre-money valuation control to our sample (Table 8).

Table 7 explores the robustness of the Access to Following Funding Rounds and Amount of Capital Raised in the Following Round results (Series A, Series B, and Series C). A parallel specification structure to that used in the previous specifications, first outlined in Tables 5 and 6, is employed, with the exception of company fixed-effects inclusion. Further, the initial underlying datasets at the basis of the analysis exhibited in Section 6.1 are not suited for the inclusion of company fixed-effects as they included one observation per company, making the addition of company fixed-effects unreal. Indeed, the dataset employed in models (5-1) and (6-1) encompassed companies raising exclusively seed or series A funding, while the dataset underlying specification (5-2) and (6-2) consisted of companies reaching series B funding, and in (5-3) and (6-3) firms that reached series C funding. Thus, in order to allow for company fixed effect a new data structure is needed. We built a novel dataframe containing one observation for each round raised by the company, up to series C. For instance, a company raising seed, series A, series B and series C funding will appear in four observations in the dataset. All the variables related to the company, as *estimated revenue range* or *number of employees*, to its investors, and *Number of IPOs*, and to the position in the network and in the transitive fitness model, as *funding attraction score*, are updated at each funding stage according to time, the business evolution, and the network dynamics.

Specification (7-1) shows that the *funding attraction index* is positively associated to the attraction of a following round of capital at a statistically significant level (1%) and implies an increase in likelihood of entering a following round of 12%. The effect is significant, therefore confirms the outputs of Table 5. Further, the association is positive for the interaction between the FAI and the seed stage, and the FAI and series A, but it is confirmed negative for the interaction between the FAI and series B funding, as occurs in specification (5-6). The second model (specification (**??**-2), displays the positive association between the likelihood of raising more capital in the follow on round and the funding attraction index. The association is confirmed and statistically significant at conventional levels. Indeed, the results of the main models

are corroborated against the company specificity bias.

We further address the unobservable quality issue by matching each company to a comparable, fitting industry, size, and pre-money valuation. Unfortunately, only a subsample of our deal-level data discloses Pre-Money Valuation, thus the analysis had to be repeated on a reduced dataset. Nevertheless, we replicate the model employed in the previous section, adding the pre-money valuation to the specifications. Results are confirmed at statistically significant levels, corroborating our hypotheses.

	Selection Probit	Outcome Probit	Outcome Probit	Outcome OLS
	(1)	(2)	(3)	(4)
Access to:	Seed	Series A	Series A	Money raised
Network Score Mean	0.0773***	0.4236***	0.1150**	0.0744*
	(0.000)	(0.001)	(0.108)	(0.043)
Premoney	0.0322***	0.016***	0.1115**	0.5834*
	(0.009)	(0.001)	(0.092)	(0.116)
Inverse Mills Ratio Seed		0.6597	7.7858	
			(10.880)	
Inverse Mills Ratio Series A				1.5178
				(3.896)
Total Experience	0.0107			
	(0.000)			
Experience Industry	0.0027**			
	(0.001)			
Controls:				
Year Seed	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes
No. Investors (previous rounds)	No	No	Yes	Yes
Money Raised (previous rounds)	No	No	Yes	Yes
Funding Rounds	No	No	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes
Constant	-3.7877	-0.5012	-2.1849	3.8716
	(0.179)	(0.130)	(3.911)	(6.319)
Observation Number	1,808,150	1,136	584	341
R squared	4.29%	38.43%	57.78%	76.5%
Adjusted R squared				-4.6%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 8: Pre-Money Valuation Robustness

6.4 Robustness to Successful Investors' Preferential Access

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Network Score Mean	0.0651***	0.1238***	0.0532***	0.2628***	0.0128*	-0.0706***
	(0.002)	(0.015)	(0.003)	(0.018)	(0.008)	(0.040)
Top IPO	0.0342***	0.1834***	0.0371***	0.1049**	0.0203*	0.1299*
	(0.009)	(0.043)	(0.008)	(0.041)	(0.019)	(0.083)
Inverse Mills Ratio		0.9171		0.9109***		2.0305*
		(0.349)		(0.285)		(1.118)
Total Experience	0.0113***		0.0107***		0.0390***	
	(0.000)		(0.000)		(0.000)	
Experience Industry	-0.0045***		-0.0029***		-0.0222***	
	(0.001)		(0.000)		(0.004)	
Controls:						
Year Seed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.3086	-0.5771	-2.8717	-1.0949	-2.9795	-0.0415
	(0.024)	(0.120)	0.039)	(0.084)	(0.030)	(0.130)
Observation Number	30,267,644	18,791	56,390,336	29,442	5,544,965	7,762
R squared	3.89%	28.06%	3.85%	33.33%	4.65%	20.85%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 9: Preferred Access: Selection and Outcome

Most successful investors' preferential access to higher quality startups is a common burden in venture capital literature (Nanda et al. 2020). The deal flow available to each fund is secret and undisclosed, but companies reach out to higher quality and most famous entrepreneurs first, a practice which translated in the possibility for most successful investors to pick the rounds in which they would like to participate before others (Lerner et al., 2022). Similar situation arise in all the environments in which reputation plays a role. Due to the private nature of proposed investments, the need to find a proxy arises. In the table presented in the current section, we proposed as proxy for investors' success a dummy which takes value of 1 if the investor belongs to the *top 20% on the success rate*, defined as number of IPOs over number of total investments (Hsu, 2005; Lerner, 2022).

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Network Score Mean	0.1245***	0.0837***	0.0226**
	(0.042)	(0.035)	(0.092)
Top IPO	0.2901**	0.1252***	0.0036*
	(0.140)	(0.031)	(0.094)
Inverse Mills Ratio	0.3745	0.3997	-0.3410
	(0.992)	(0.241)	(3.255)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	-1.4979	-2.8201	-0.4810
	(0.416)	(0.425)	(0.331)
Observation Number	2,635	2,633	1,358
R squared	39.05%	35.9%	25.84%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 10: Preferred Access: Model

Table 9 delves into the probability of companies being matched with investors in a funding round and the subsequent likelihood of accessing follow-on funding rounds. Specification (9-1) illustrates the positive impact of both seed investors' network influence and of the presence of successful investors on the likelihood of matching investors with companies in the seed round. Furthermore, considering the selection effect, specification (9-2) unveils the positive association between high FAI investors and successful investors with the accessing subsequent funding rounds for the supported companies. These findings hold statistical significance at the 1% level. In (9-1), the estimated coefficient implies a 6% increase in the likelihood of accessing seed funding for each discrete change in the FAI measure and a 3% increase for the precess of successful investors. Similarly, (9-2) indicates a 12% increase in the likelihood of raising Series A funding for a discrete increase in FAI and a 18% increase for the presence of a top-IPO investor, accounting for the selection effect through the inverse Mills ratio. These effects are further confirmed in the comprehensive model (10-1) with a 5% level of statistical significance.

It's worth noting that while the selection equation (9-1) observes numerous matches

	Outcome OLS	Outcome OLS	Outcome OLS
	(1)	(2)	(3)
Money Raised:	Series A	Series B	Series C
Network Score Mean	0.1098**	0.0468**	0.0370**
	(0.044)	(0.058)	(0.090)
Top IPO	0.0386**	0.1244**	0.0799**
	(0.089)	(0.015)	(0.069)
Inverse Mills Ratio 2	-0.8708**	0.0330	-0.3288
	(0.415)	(0.277)	(0.620)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Number of Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	16.7758	16.6585	17.1562
	(0.341)	(0.232)	(0.630)
Observation Number	935	948	523
R squared	26.4%	39.3%	36.0%
Adjusted R squared	23.4%	36.8%	31.1%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 11: Preferred Access: Money Raised

between companies raising seed funding and investors investing in seed, the outcome equation (9-2) focuses exclusively on realized seed deals. The dependent variable in (9-1) is the *seed dummy* (indicating if the seed round was realized), while in (9-2), it is the *Series A dummy* (indicating if the Series A round was realized). Both equations incorporate overall investor experience and industry expertise to gauge the likelihood of investor-company matches. Results are adjusted for estimated revenue range, number of employees, industry group, year of former funding, and number of funding rounds. However, (9-1) allows for deal-specific controls such as money raised and the number of investors in the seed round.

Advancing in the business lifecycle, specification (9-3) highlights the positive impact of Series A investors' network influence and top IPO investors on a company access to Series A funding. Similarly, considering the selection effect, specification (9-4) reveals the favorable relationship between high pref investors and accessing subsequent funding rounds for supported companies. These results maintain statistical significance at the

1% level. The estimated coefficient in (9-3) implies a 5% increase in the likelihood of accessing Series A funding for each discrete change in the FAI measure and 4% for high IPO investors, while (9-4) shows a 26% increase in the likelihood of raising Series B funding for a discrete change in the FAI and a 11% for the presence of top IPO investors, considering the selection effect. These effects are slightly enhanced but still significant in the comprehensive model (10-2) with a 5% level of statistical significance. Observations and controls remain consistent with the previous cases.

However, the impact of the pref diminishes as we transition from Series B to Series C funding rounds, as illustrated in specifications (9-5), (9-6), and (10-3). The FAI's effect decreases on the likelihood of companies being selected in Series B compared to its impact in the seed stage and Series A, now equating to 1%. Similarly, the top-IPO effect decreased to 2%. Nonetheless, the outcome equation effect appears negative. In the comprehensive model specification (10-3), controlling for a wide range of variables related to the company, investors, and the deal, the effect becomes slightly positive (2%), although the level of significance decreases to 10% compared to previous specifications (10-1 and 10-2). This underscores the importance of controls related to previous rounds. Once again, specifics regarding observations and controls mirror those in models concerning the seed stage and Series A.

Table 11 outlines the multivariate regressions for *Money Raised*. In specification (11-1), a positive correlation is observed between the normalized *funding attraction index (FAI)*, the *top IPO dummy* and the logarithm of *money raised in Series A*, with statistical significance at the 5% level. This suggests a 11% increase in money raised for each discrete change in the FAI measure, and a 4% increase in money raised if a top performing investor is involved. However, the FAI effect diminishes respectively to 5% and 4% , while the top IPO effect increases to 12% and 8% when considering the *logarithm of money raised in Series B* and *Series C*, respectively, as seen in specifications (11-2) and (11-3).

The economic significance of the FAI effect in both models remains at 10%. It's important to note that the observations included are deals for which the amount of money raised was disclosed. Furthermore, the inverse Mills ratios included in the models of Table 11 are computed through the three specifications of Table 11. For instance, in case (11-1), the selection effect considers the inclusion of observations in the sample which received Series A funding, having already been excluded from the sample which received seed funding. These results affirm that while the influence effect persists, it slightly diminishes as companies gain autonomy in their ability to progress in the business lifecycle and attract subsequent substantial funding rounds. Thus, results are verified to be robust against the preferential access bias.

6.5 Robustness to Geographical Bias

The limited geographical scope of the analysis, focused on the State of California, might also have led to some sort of distortion in the results. Indeed, the choice landed on it as California, and more specifically Silicon Valley, is the most well developed venture capital market in the world, the promised land of startups and early stage investors. Nevertheless, the solidity of results on another comparable, but different, area arises. In this regard, the analysis was entirely repeated in all its four models (Access to Series A, Access to Series B, Capital Raised in Series A, and Capital Raised in Series B) on a novel dataset emcompassing all the deals funding companies established between 2010 and 2021 in the State of New York. The area is picked as it is considered the second most mature market after California (Chen et al., 2010), notwithstanding the inherent differences among the two States. Embracing the two distinct souls is essential to fully comprehend the difference in the results. The sample was recreated strictly following the same methods and specifics employed in the building of the main dataset used in the current analysis.

Table 12 investigates the probability of companies being matched with investors in a funding round and the subsequent likelihood of accessing follow-on funding rounds. Specification (12-1) demonstrates the positive impact of seed investors' network influence on the likelihood of matching investors with companies in the seed round. Moreover, considering the selection effect, specification (12-2) reveals the positive association between high FAI (Financial Activity Index) investors and accessing subsequent funding rounds for the supported companies. These findings are statistically significant at the 1% level. In (12-1), the estimated coefficient suggests a 4% increase in the likelihood of accessing seed funding for each discrete change in the FAI measure, while (12-2) indicates a 5% increase in the likelihood of raising Series A funding, accounting for the selection effect through the inverse Mills ratio. These effects are further confirmed in the comprehensive model (13-1) with a 5% level of statistical significance.

It's noteworthy that while the selection equation (12-1) observes numerous matches between companies raising seed funding and investors investing in seed, the outcome equation (12-2) focuses exclusively on realized seed deals. The dependent variable in (12-1) is the *seed dummy* (indicating if the seed round was realized), while in (12-2), it is the *series A dummy* (indicating if the Series A round was realized). Both equations incorporate overall investor experience and industry expertise to gauge the likelihood of investor-company matches. Results are adjusted for estimated revenue range, number of employees, industry group, year of former funding, and number of funding rounds. However, (12-1) allows for deal-specific controls such as money raised and the number of investors in the seed round.

Advancing in the business lifecycle, specification (12-3) highlights the positive impact of Series A investors' network influence on company investment in Series A funding. Similarly, considering the selection effect, specification (12-4) reveals the favorable relationship between high FAI investors and accessing subsequent funding rounds for supported companies. These results maintain statistical significance at the 1% level. The estimated coefficient in (12-3) implies a 6% increase in the likelihood of accessing Series A funding for each discrete change in the FAI measure, while (12-4) shows a 11% increase in the likelihood of raising Series B funding, considering the selection effect. These effects are slightly enhanced but still significant in the comprehensive model (13-2) with a 1% level of statistical significance. Observations and controls remain consistent with the previous cases.

However, the impact of the FAI diminishes as we transition from Series B to Series C funding rounds, as illustrated in specifications (12-5), (12-6), and (13-3). The FAI's effect decreases on the likelihood of companies being selected in Series B compared to its impact in the seed stage and Series A, now equating to 5%. Nonetheless, the outcome equation effect appears negative. In the comprehensive model specification (13-3), controlling for a wide range of variables related to the company, investors, and the deal, the effect becomes slightly positive (3%), although the level of significance decreases to 10% compared to previous specifications (13-1 and 13-2). This underscores the importance of controls related to previous rounds. Once again, specifics regarding observations and controls mirror those in models concerning the seed stage and Series A.

Table 14 displays the multivariate regressions for *Money Raised*. Specification (14-1) reveals a positive association between the normalized *funding attraction index* (FAI) and the logarithm of *money raised in Series A*, significant at the 5% level. This implies a 6% increase in money raised for each discrete change in the FAI measure. However, this effect diminishes to 3% and 2% when considering the logarithm of *money raised in Series B* and *Series C*, respectively, in specifications (14-2) and (14-3).

The economic significance of the FAI effect in the two models remains at 10%. It's important to note that the observations included are deals for which the amount of money raised was disclosed. Additionally, the inverse Mills ratios included in the models of Table 14 are computed through the three specifications of Table 13. For instance, in case (14-1), the selection effect accounts for the inclusion of observations in the sample which received Series A funding, already removed from the sample which received seed funding. Indeed, the results affirm that the influence effect persists but slightly diminishes as the company gains autonomy in its ability to progress in the business lifecycle and garners subsequent substantial funding rounds.

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Network Score Max	0.0385***	0.0275**	0.0601***	0.1087***	0.0481***	-0.0311*
	(0.001)	(0.011)	(0.003)	(0.014)	(0.004)	(0.020)
Inverse Mills Ratio		0.6962***		-2.8968***		2.4860*
		(0.209)		(0.962)		(1.430)
Total Experience	0.0052***		0.0518***		0.0851***	
	(0.000)		(0.001)		(0.002)	
Experience Industry	-0.0070***		-0.0260***		-0.0291**	
	(0.000)		(0.006)		(0.012)	
Controls:						
Year Seed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.8380	3.0823	-2.8717	0.2382	-2.6803	0.3683
	(0.263)	(211.815)	0.039)	(0.204)	(0.061)	(0.298)
Observation Number	26,301,350	21,508	2,223,399	6,065	777,952	2,405
R squared	4.81%	29.02%	5.04%	32.76%	5.65%	35.87%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 12: New York: Selection and Outcome.

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Network Score Max	0.0464	0.1168***	0.0273*
	(0.031)	(0.044)	(0.067)
Inverse Mills Ratio	3.5047	-2.3275	3.2776
	(3.036)	(3.103)	(4.593)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	5.2138	-0.7759	-2.8717
	(0.000)	(0.511)	0.039)
Observation Number	1,045	717	430
R squared	37.01%	32.11%	40.21%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 13: New York: Model

	Outcome OLS	Outcome OLS	Outcome OLS
	(1)	(2)	(3)
Money Raised:	Series A	Series B	Series C
Network Score Max	0.0568**	0.0346**	0.0159*
	(0.027)	(0.049)	(0.054)
Inverse Mills Ratio 2	0.7074	1.1844**	-0.8962
	(0.675)	(0.541)	(1.087)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	15.9055	14.1658	18.4553
	(0.468)	(0.728)	(1.605)
Observation Number	332	313	171
R squared	28.2%	45.1%	55.6%
Adjusted R squared	18.9%	38.0%	43.7%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 14: New York: Money Raised

7 Conclusion and Discussion

The paper tested, through the lenses of network theory, the impact of securing funding from a high *funding attraction* investor in the earliest stages of a company's business life. After the formulation the novel *funding attraction score*, possible only as a result of the construction of a continuously updated temporal bipartite network model fed with real-world data and the consequent conception and employment of the innovative transitive fitness model, the *score*'s explanatory power is tested on *access to follow-on funding rounds* and *capital raised during follow-on rounds*. The results clearly indicate an influence of the *funding attraction index* on the access to series A funding, but the impact dissolves as the company proceeds in business life through the raise of further funding rounds, namely series B and C funding. The outcomes of the analysis argue for an increased role of investors' reputation in the earliest phases of a company's establishment. A young venture can make the best use of the reputation of its investors in the seeding phase, when its future is unclear. Indeed, at the seeding stage, a company had yet to prove the resilience of its team, the strengths of its value and production chains, or the quality of its products. Thus, being endorsed by a successful investor is more valuable than in later stages, when the enterprise had the money and means to prove its quality.

Participating through capital is a signal of the trust that the investor puts on the company. The reasoning is confirmed by classical literature measures and industry measures aiming at quantifying venture capitalists' reputation, but the newly developed *funding attraction score*, by virtue of its continuously updating nature, furthers the analysis to one additional step. The *score* is built on the investors' success, explained in terms of raising follow-on funding, of the startups they have previously backed, and changes according to the latest deals executed in the market. Therefore, we argue that the *funding attraction score* better captures the investors' ability to provide non-financial support to the companies they back, in a way which fits the current market and the current moment, than other considered reputation measures - which take success of portfolio companies into account in an accumulative fashion rather than decaying one. As venture capital is aimed at financing the avant-garde of technology and innovation, the feature is crucial.

Once the companies secured access to the following funding rounds, the *FAI* has an influence on the amount of capital raised, increasing it across a company's business life. Interestingly, although the effect of the *funding attraction index* fades in the company seeks to attract later funding, its latent effect persists and is revealed through more conspicuous capital raises. The result argues that, if a company is considered worth to proceed in business life by the investment community (i.e. secures series B and C

funding), then the influence of its investors increase the participation and involvement of other financial players, as they have faith in the future success of the venture. The several results of the models outlined in the study help deepen our understanding of the mechanism ruling how a company elevates from the uncertain ground of seed financing by picking the angels and venture capitalists who can better support them in their seek for the limelight, and the findings are consistent with the view that investors matter beyond their financial power.

Additionally, the innovative transitive fitness model allows us to make the company the protagonist of the ultimate model, while searching for answers looking at its investors. Besides, notwithstanding the impossibility of disentangling other reputation measures from the score that we propose, we have shown that no significant direct correlation is present between said measures and the score. Moreover, as we have proved in our study, the formulated score has a higher impact on the likelihood of raising series A and of attracting higher amount of capital than traditional measures of reputation employed in the literature (experience and success) and in the industry (being nominated in Forbes' Midas list) and in venture capital networks (Appendix 1).

The several robustness checks conducted offer additional insights. Firstly, the inclusion of company fixed-effects helps the model to disentangle the outcomes from the uncapturable specifics of each venture. Besides, the inclusion of a pre-money valuation variable increases the explanatory power of the model. Secondly, results are validated against the eventual preferential access granted to the most famous or successful investors. Interestingly, investors who enjoy the highest success rate seem to increase their effect on the company access to follow on funding as the company proceeds in business life. Lastly, the geographical bias robustness is explored. The outcomes of the main dataset, located in California, are confirmed for the New York dataset.

Several alternative explanations to the empirical pattern have been considered throughout the empirical analysis. Firstly, in the study, the funding attraction mechanism is contingent to the number of investors entering the series A deal, while it could be alternatively computed on the basis of the amount of investments attracted in series A funding or the quality of the investors attracted. Undeniably, at a first glance the latter strategy could be considered as more relevant in purely financial terms. Indeed, the overall funding amount typically has a significant impact on the startup's ability to execute its growth strategies effectively. Having a larger amount of capital gives the startup more financial flexibility, allows for faster scaling, and increases the chances of successfully reaching critical milestones. Nevertheless, the choice of proceeding basing the score on the number of investors was driven by the concept that the measure is a primary indicator of the confidence and support that the investors' community has in the startup's potential. Furthermore, a higher number of investors implies diversify-

ing risk and accessing a broader network of connections and of information, as well as dragging attention towards the company.

A further additional explanation could consider some investors having better skills, without considering connections, than others in nurturing a company in its earliest phases. Naturally, early-stage investors play a crucial role in supporting startups beyond just providing capital. Their skills encompass mentorship, guidance, strategic planning to assist founders in achieving long-term goals, operational expertise to optimize processes and guide startups in building strong teams, and insights into market trends and fundraising strategies. Undeniably, some investors have better skills than others - and unfortunately skills are unobservable. Therefore, by making our score conditional on time and recent closed deals information, and considering that skills are hardly lost over time (at best increased), we could consider that an investor with high skills which continuously invest will keep a high attraction score across time. But, an investor associated with a high attraction score could also have no skills other than a big network of contacts which endorses its investment by backing them at the first available moment. However, it is hard to believe that an untalented investor will be followed by its community, therefore we can comfortably affirm that the attraction score also partially incorporates an investor's skills, pendant on the fact that the investor continues to put the skin in the game. Nevertheless, a highly skilled investors could have a low funding attraction score in peculiar cases in which the companies she backs opt for a nontraditional path, as a direct exit after seed financing, pursuing growth exclusively with internal resources or debt financing, or raising more rounds through the same unique investor. These events occur rarely.

In addition, various potential limitations of the analyses should be noted. First, investors might have undisclosed preferences when selecting startups for funding, leading to the funding of specific types of startups that go unnoticed. We try to account for this aspect by inserting the industry experience variable in our model. Alternatively, it is possible that certain types of entrepreneurs, such as those of higher quality or those pursuing a commercialization strategy, opt to be considered for funding in ways that are not easily observed. Secondly, as investors tend to prefer funding startups that prioritize innovative strategies, it becomes challenging to separate the impact of their presence from pre-existing technological advancements. The empirical analysis could address this issue by controlling for the number of patents in the month and year preceding external funding. Thirdly, the effort to include the most correct and exhaustive information implied setting strict temporal and geographical limits, which overlook the global interconnectedness of the modern early-investments world. Lastly, albeit Crunchbase is among the most complete databases in the field (Retterath and Braun, 2020), it faces two main challenges: survivorship bias and nondisclosure of pri-

vate deals. The survivorship bias could occur when the information related to failed or unsuccessful companies gets deleted. However, it must be noted that the companies included in our dataset all received at least one round of funding, which is a quality indicator, and are located in a primary location. The nondisclosure bias is a result of private deals not being under the requirement of being publicly disclosed. Investors and companies decided to be reserved about their deals for a wide variety of reasons (Dixit and Pindyck, 1998).

Looking to the future, while the current study does not test the ex post performance implications of selecting a particular investor, it would be interesting to do so. For example, have startups which received a investment by a high-attraction investor gone public or have they been acquired? Have they become an unicorn, or acquired their competitors? Have they delivered products to the market more quickly? Have they registered patents? Nevertheless, the findings in this study are consistent with the theory that early investors are crucial in the faith of a company, and the right backing can allow for more opportunities in startup growth, as they have deepen our understanding of early-stage investments dynamics and the impact of reputation and networks in early-stage funding. The innovative funding attraction score and the novel transitive model applications to the other fields ruled by trust, reputation, and continuous performance are yet to be explored.

8 References

Albert, R., and Barabási, A. L. (2002). Statistical mechanics of complex networks. Reviews of modern physics, 74(1), 47.

Amor, S. B., and Kooli, M. (2020). Do M-and-A exits have the same effect on venture capital reputation than IPO exits?. Journal of Banking and Finance, 111, 105704.

Atanasov, V., Ivanov, V., and Litvak, K. (2012). Does reputation limit opportunistic behavior in the VC industry? Evidence from litigation against VCs. The Journal of Finance, 67(6), 2215-2246.

Barabási, A. L., Fraiberger, S., Sinatra, R., Resch, M., and Riedl, C. (2018). Quantifying reputation and success in art. Science. 362. eaau7224. 10.1126/science.aau7224.

Bonaventura, M., Ciotti, V., Panzarasa, P., Liverani, S., Lacasa, L., and Latora, V. (2020). Predicting success in the worldwide start-up network. Sci Rep 10, 345.

Bonini, S., Capizzi, V., and Zocchi, P. (2019). The performance of angel-backed companies. Journal of Banking and Finance, 100, 328-345.

Bottazzi, L., Da Rin, M., and Hellmann, T. (2008). Who are the active investors?: Evidence from venture capital. Journal of Financial Economics, 89(3), 488-512.

Bygrave, W. D. (1988). The structure of the investment networks of venture capital firms. Journal of Business Venturing, 3(2), 137-157.

Caldarelli, G., Capocci, A., De Los Rios, P., and Munoz, M. A. (2002). Scale-free networks from varying vertex intrinsic fitness. Physical review letters, 89(25), 258702.

Capizzi, V., Croce, A., and Tenca, F. (2022). Do Business Angels' Investments Make It Easier to Raise Follow-on Venture Capital Financing? An Analysis of the Relevance of Business Angels' Investment Practices. British journal of management, 33(1), 306-326.

Castellano, C., Fortunato, S., and Loreto, V. (2007). Statistical physics of social dynamics. Reviews of Modern Physics. 81, 591–646.

Chen, H., Gompers, P., Kovner, A., and Lerner, J. (2010). Buy local? the geography of Venture Capital. Journal of Urban Economics, 67(1), 90–102. doi:10.1016/j.jue.2009.09.013.

Croce, A., Ughetto, E., Bonini, S., and Capizzi, V. (2021). Gazelles, ponies, and the impact of business angels' characteristics on firm growth. Journal of Small Business Management, 59(2), 223-248.

Dutta S., and Folta, T. B. (2016). A comparison of the effect of angels and venture capitalists on innovation and value creation. Journal of Business Venturing, Volume 31, Issue 1, 39-54.

Gafni, H., Marom, D., Robb, A., and Sade, O. (2021). Gender dynamics in crowdfunding (Kickstarter): Evidence on entrepreneurs, backers, and taste-based discrimination. Review of Finance, 25(2), 235-274.

Gompers, P. A., Kaplan, S. N., and Mukharlyamov, V. (2015). What Do Private Equity Firms Say They Do? (No. w21133). National Bureau of Economic Research.

Gompers, P. A., Mukharlyamov, V., Weisburst, E., and Xuan, Y. (2014). Gender effects in venture capital. Available at SSRN, 2445497.

Florida, R., and Kenney, M. (1988). Venture Capital and high technology entrepreneurship. Journal of Business Venturing, 3(4), 301–319. doi:10.1016/0883-9026(88)90011-0.

Hellman, T., and Thiele, V. (2015). Friends or foes? The interrelationship between angel and venture capital markets. Journal of Financial Economics, 115:639-53.

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

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation?. The journal of finance, 59(4), 1805-1844.

Hsu, D. H. (2007). Experienced entrepreneurial founders, organizational capital, and venture capital funding. Research policy, 36(5), 722-741.

Inderst, R., and Mueller, H. M. (2009). Early-stage financing and firm growth in new industries. Journal of Financial Economics, 93(2), 276-291.

Kerr, W. R., Lerner, J., and Schoar, A. (2014). The Consequences of Entrepreneurial Finance: Evidence from Angel Financings. The Review of Financial Studies, 27(1), 20-55.

Krishnan, C. N. V., Ivanov, V. I., Masulis, R. W., and Singh, A. K. (2011). Venture capital reputation, post-IPO performance, and corporate governance. Journal of Financial and Quantitative Analysis, 46(5), 1295-1333.

Lerner, J. (1994). The syndication of venture capital investments. Financial management, 16-27.

Lerner, J., and Nanda, R. (2020). Venture capital's role in financing innovation: What we know and how much we still need to learn. Journal of Economic Perspectives, 34(3), 237-61.

Lerner, J., Mao, J., Schoar, A., Zhang, N.R., (2022). Investing outside the box: Evidence from alternative vehicles in private equity, Journal of Financial Economics, Volume 143, Issue 1.

Lindsey, L., 2008, Blurring Firm Boundaries: The Role of Venture Capital in Strategic Alliances, The Journal of Finance, 63(3), 1137–1168.

Manigart et al, 2010

Nahata, R. (2008). Venture capital reputation and investment performance. Journal of financial economics, 90(2), 127-151.

Nanda 2020 The persistent effect of initial success: Evidence from venture capitalR Ramana Nandaa, Sampsa Samilab, Olav Sorensonc,

Parra, Á., and Winter, R. A. (2022). Early-stage venture financing. Journal of Corporate Finance, 77, 102291.

Piacentino, G. (2019). Venture capital and capital allocation. The Journal of Finance, 74(3), 1261-1314.

Retterath, A., and Braun, R. (2020). Benchmarking venture capital databases. Available at SSRN 3706108.

Sridharan, K. (2016, February 1). Key Difference of Silicon Valley Vs. Alley: Big Thoughts. Wharton Magazine.

Sorenson, O., and Stuart, T. E. (2008). Bringing the context back in: Settings and the search for syndicate partners in venture capital investment networks. Administrative Science Quarterly, 53(2), 266-294.

Sorenson, O., and Waguespack, D. M. (2006). Social structure and exchange: Self-confirming dynamics in Hollywood. Administrative Science Quarterly, 51(4), 560-589.

Stross, R. (2012). The Launch Pad: Inside Y Combinator, Silicon Valley's Most Exclusive School for Startups. Penguin, UK.

Timmons J.A. (1990). New venture creation: Entrepreneurship in the 1990s. Homewood, IL: Irwin.

Wang, X., Wu, L., and Hitt, L. M. (2023). Social Media Alleviates Venture Capital Funding Inequality for Women and Less Connected Entrepreneurs. Management Science.

Werth, J., and Böert, P. (2011). Co-Investment Networks of Business Angels and the Performance of Their Start-Up Investments. International Journal of Entrepreneurial Venturing. 5. 10.2139/ssrn.1977970.

Zava, M. and Caselli, S. (2023). Venture Capital Networks. The Palgrave Encyclopedia of Private Equity, Available at SSRN: https://ssrn.com/abstract=4433596

The relevance of broker networks for information diffusion in the stock marketR Marco Di Maggioa,, Francesco Franzonib, Amir Kermanic, Carlo Sommavilla

(Retterath and Braun, 2020)

Dixit and Pindyck, 1998

9 Appendix

9.1 Comparison with Other Metrics

9.1.1 VS literature metrics

	Selection Probit	Outcome Probit	Outcome Probit	Outcome OLS
	(1)	(2)	(3)	(4)
Access to:	Seed	Series A	Series A	Money raised
Success Rate	0.0037***	0.0003***	0.0003***	0.0001***
	(0.0000)	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio Seed		0.2473	-0.2805	
		(0.294)	(0.769)	
Inverse Mills Ratio Series A				-0.6128
				(0.534)
Total Experience	0.0126***			
	(0.000)			
Experience Industry	-0.0016***			
	(0.001)			
Controls:				
Year Seed	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes
No. Investors (previous rounds)	No	No	Yes	Yes
Money Raised (previous rounds)	No	No	Yes	Yes
Funding Rounds	No	No	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes
Constant	-2.7030	-0.5190	-1.3629	16.7625
	(0.026)	(0.130)	(0.417)	(0.453)
Observation Number	30,264,902	13,876	2,280	784
R squared	6.157%	26.36%	38.94%	25.9%
Adjusted R squared				22.5%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 15: Success Rate

The section is dedicated to comparing the results obtained through the model featuring the FAI to the ones obtained in models using the most common measures used in the literature to estimate venture capital reputation: success rate (15) and overall investments (16). Parallel specification structures to that used in the previous tables are employed.

In both cases, the specifications shows that

	Selection Probit	Outcome Probit	Outcome Probit	Outcome OLS
	(1)	(2)	(3)	(4)
Access to:	Seed	Series A	Series A	Money raised
Fund Overall Investments	0.0007***	0.0001***	0.0002**	0.0002**
	(0.000)	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio Seed		0.9877**	0.9708	
		(0.462)	(1.030)	
Inverse Mills Ratio Series A				-0.7770
				(0.508)
Total Experience	0.0121***			
	(0.000)			
Experience Industry	-0.0032***			
	(0.001)			
Controls:				
Year Seed	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes
No. Investors (previous rounds)	No	No	Yes	Yes
Money Raised (previous rounds)	No	No	Yes	Yes
Funding Rounds	No	No	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes
Constant	-3.4087	-0.5562	-1.4091	16.8112
	(0.025)	(0.129)	(0.415)	(0.421)
Observation Number	30,264,902	13,876	2,280	784
R squared	4.066%	26.25%	38.98%	26.2%
Adjusted R squared				22.8%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 16: Fund Lifetime investments

9.1.2 VS network metrics

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Investor's Degree Centrality	-0.0001***	-0.0001***	0.0001***	0.0001***	0.001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio		0.9223***		0.4932***		-1.3350
		(0.304)		(0.170)		(0.955)
Total Experience	0.0114***		0.0108***		0.0379***	
	(0.000)		(0.000)		(0.000)	
Experience Industry	-0.0022***		-0.0019***		-0.0246***	
	(0.000)		(0.000)		(0.003)	
Controls:						
Year (Previous Round)	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.1926	-0.8046	-3.3094	-1.3204	-2.9703	-0.0796
	(0.019)	(0.093)	(0.013)	(0.081)	(0.027)	(0.116)
Observation Number	43,168,591	26,966	89,826,081	46,049	7,342,606	10,620
R squared	35.84%	28.55%	3.85%	28.11%	4.75%	21.95%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 17: Selection and Outcome: Investor's Degree Centrality

Besides, Table **??** explores the effect of traditional network measures on the likelihood of raising series A results. A parallel specification structure to that used in the previous table is employed. The most commonly used network metric, the *centrality degree*, accounting for the number of connections a node made, displays statistically insignificant and very small parameter estimates in both the company's and the investors' specification. Instead, results associated with betweenness centrality, usually considered as a measure of the transit of information among parties (Lindsey, 2008), are controversial. It is important to notice that the information travelling on the univariate nets considered in our experiment is financial information, as the ties are identified as having an investor in common (when company nodes are examined) or having invested in the same company (if investor nodes are under analysis).

While also the *investors betweenness centrality* appears to have no effect, the case is different. Investors central in the information flow have access to better deals, and

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Investor's Degree Centrality	0.0001***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio	1.0665	0.7821	-0.7572
	(0.940)	(0.851)	(3.473)
Controls:			
Year (previous round)	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	-1.7738	-2.4902	-0.3997
	(0.417)	(0.424)	(0.330)
Observation Number	2,635	2,685	1,358
R squared	40.37%	36.61%	25.83%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 18: Access to the following round: Investor's Degree Centrality

are proved to have better portfolio performances (Hochberg et al., 2007), as the knowledge they gain favors them in their decisions. However, it might be the case that the information favors them in making better investment decisions, but does not favor the companies they back in moving forward with their funding life.

	Outcome OLS	Outcome OLS	Outcome OLS
	(1)	(2)	(3)
Money Raised:	Series A	Series B	Series C
Investor's Degree Centrality	0.0001***	0.0001***	0.0001**
	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio 2	-0.6112***	-0.1110**	-0.9810***
	(0.125)	(0.082)	(0.302)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	16.4254	16.1139	17.5539
	(0.066)	(0.044)	(0.249)
Observation Number	935	935	523
R squared	27.2%	30.2%	33.1%
Adjusted R Squared	26.9%	30.1%	32.5%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 19: Raising Capital: Investor's Degree Centrality.

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Investor's Betweeness Centrality	-1.5691***	-2.3977***	-0.3028***	0.2688	0.0646	2.6699***
	(0.114)	(0.577)	(0.070)	(0.344)	(0.190)	(0.858)
Inverse Mills Ratio		0.9234***		0.4827***		-1.3332
		(0.305)		(0.169)		(0.955)
Total Experience	0.0114***		0.0108***		0.0379***	
	(0.000)		(0.000)		(0.000)	
Experience Industry	-0.0027***		-0.0020***		-0.0246***	
	(0.000)		(0.000)		(0.003)	
Controls:						
Year Seed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.2012	-0.8087	-3.3116	-1.3153	-2.9705	-0.0829
	(0.010)	(0.092)	(0.013)	(0.0081)	(0.027)	(0.116)
Observation Number	43,168,591	26,966	89,826,081	46,049	7,342,606	10,620
R squared	35.39%	28.56%	3.85%	28.10%	4.75%	21.96%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 20: Selection and Outcome: Investor's Betweeness Centrality

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Investor's Betweeness Centrality	0.8590	-0.4788	1.9044
	(1.635)	(1.242)	(2.291)
Inverse Mills Ratio	0.8374	0.7526	-0.7539
	(0.954)	(0.850)	(3.476)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	-1.7366	-2.4758	-0.4038
	(0.417)	(0.423)	(0.330)
Observation Number	2,635	2,685	1,358
R squared	40.31%	36.61%	25.86%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 21: Access to the following round: Investor's Betweeness Centrality

	Outcome OLS	Outcome OLS	Outcome OLS
	(1)	(2)	(3)
Money Raised:	Series A	Series B	Series C
Investor's Betweeness Centrality	-1.8784***	-1.1021***	2.6763***
	(0.363)	(0.265)	(0.803)
Inverse Mills Ratio 2	-0.6265***	0.1092***	-1.0495***
	(0.125)	(0.023)	(0.311)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	16.4381	16.1357	17.6066
	(0.066)	(0.065)	(0.256)
Observation Number	935	935	523
R squared	27.3%	29.6%	33.1%
Adjusted R Squared	27.0%	29.4%	32.5%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 22: Raising Money: Investor's Betweeness Centrality

9.1.3 VS company's network metrics

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Company Max Degree	0.0001***	0.0005**	0.0001***	0.0002***	0.0001***	0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio		1.0317***		0.7753***		-1.7648
		(0.353)		(0.229)		(1.108)
Total Experience	0.0113***		0.0106***		0.0390***	
	(0.000)		(0.000)		(0.000)	
Experience Industry	-0.0045***		-0.0060***		-0.0221***	
	(0.001)		(0.000)		(0.004)	
Controls:						
Year Seed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.2762	-0.5684	-3.4159	-1.5104	-2.9721	-0.0540
	(0.024)	(0.119)	(0.017)	(0.112)	(0.030)	(0.129)
Observation Number	30,267,644	18,791	56,390,336	24,045	5,544,965	7,762
R squared	3.68%	27.77%	4.30%	28.62%	4.66%	21.31%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 23: Selection and Outcome: Company's Degree Centrality

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Company Max Degree	0.0003***	0.0002***	0.0001**
	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio	1.2008	1.0221	-0.7161
	(0.964)	(0.809)	(3.333)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	-1.6373	-2.5892	-0.4359
	(0.410)	(0.425)	(0.331)
Observation Number	2,635	2,685	1,358
R squared	39.19%	36.84%	25.77%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 24: Access to the following round: Company's Degree Centrality

	Outcome OLS	Outcome OLS	Outcome OLS
	(1)	(2)	(3)
Company Max Degree	0.0001***	0.0001***	0.0002***
	(0.000)	(0.000)	(0.000)
Inverse Mills Ratio 2	-0.3874***	0.1092	-0.7012**
	(0.145)	(0.103)	(0.344)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	16.3068	16.1357	17.4780
	(0.091)	(0.065)	(0.287)
Observation Number	935	935	523
R squared	25.5%	29.6%	31.8%
Adjusted R Squared	25.1%	29.4%	30.9%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 25: Raising Capital: Company's Degree Centrality

9.1.4 VS midas

	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit	Selection Probit	Outcome Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Round:	Seed	Series A	Series A	Series B	Series B	Series C
Midas Dummy	0.2815***	0.3370***	0.3786***	0.4788***	0.2536***	0.1565***
	(0.018)	(0.083)	(0.012)	(0.072)	(0.014)	(0.051)
Inverse Mills Ratio		0.9851		00.5951***		-0.7015
		(0.352)		(0.243)		(1.155)
Total Experience	0.0113***		0.0106***		0.0349***	
	(0.000)		(0.000)		(0.001)	
Experience Industry	-0.0040***		-0.0025***		-0.0223***	
	(0.001)		(0.000)		(0.004)	
Controls:						
Year Seed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.2625	-0.5202	-3.3540	-1.3176	-2.9895	-0.0968
	(0.024)	(0.118)	(0.017)	(0.111)	(0.030)	(0.129)
Observation Number	30,267,644	18,791	56,390,336	24,045	5,544,965	7,762
R squared	3.73%	27.81%	3.91%	28.36%	4.90%	20.86%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 26: Selection and Outcome: Midas Investors

	Outcome Probit	Outcome Probit	Outcome Probit
	(1)	(2)	(3)
Access to Round:	Series A	Series B	Series C
Midas Dummy	0.5330***	0.7497***	0.1517
	(0.184)	(0.193)	(0.121)
Inverse Mills Ratio	0.6093	0.9741	2.5050
	(0.971)	(0.887)	(3.191)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	-1.4660	-2.4813	1,358
	(0.414)	(0.424)	(0.331)
Observation Number	2,635	2,685	1,358
R squared	38.92%	37.05%	25.86%
* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively			

Table 27: Access to the following round: Midas Investors

	Outcome OLS	Outcome OLS	Outcome OLS
	(1)	(2)	(3)
Money Raised:	Series A	Series B	Series C
Midas Dummy	0.1333**	0.1682***	0.1655**
	(0.062)	(0.112)	(0.068)
Inverse Mills Ratio 2	-0.3718**	-0.0668	-1.1493***
	(0.147)	(0.112)	(0.428)
Controls:			
Year Seed	Yes	Yes	Yes
Industry Group	Yes	Yes	Yes
Estimated Revenue Range	Yes	Yes	Yes
No. Investors (previous rounds)	Yes	Yes	Yes
Money Raised (previous rounds)	Yes	Yes	Yes
Funding Rounds	Yes	Yes	Yes
Number of Employees	Yes	Yes	Yes
Constant	16.3592	16.1232	17.7528
	(0.097)	(0.061)	(0.337)
Observation Number	935	935	523
R squared	25.2%	29.8%	31.7%
Adjusted R Squared	24.8%	29.5%	30.9%

* or ** or *** indicate statistical significance at the 10%, 5% or 1% level, respectively

Table 28: Raising Capital: Midas Investors

9.2 Synthetic Network

With the goal of testing the funding attraction mechanism introduced in the novel *funding attraction index (M_i)* we proposed, we enacted it on a synthetic network to observe if it would reproduce the real-network dynamics. Synthetic networks are an important tool in network science, as they allow researchers to explore the properties of networks under controlled conditions and validate the quality of their intuition. In complex network theory, the occurrence of a hypothesized mechanism in the real world can be verified and studied by applying that very same mechanism to an artificial net, resembling in its characteristics the real world (before the occurrence of the phenomenon) as closely as possible. Consequent to the application of the mechanism, the synthetic network will display other characteristics. If the latter are similar to the real-world specifics after the occurrence of the phenomenon, then we can confidently state that the mechanism occurs in the understood (and re-enacted) fashion. Indeed, the common approach to synthetic networks generation is to use mathematical models that capture key features of real-world networks, such as degree distribution, clustering, and community structure.

After seeding a bipartite network and the related ties, we build a growth model, capable of introducing new companies at different points in time and adjust the direction of new links based on the funding attraction mechanism. Naturally, the synthetic network tries to resemble the original financial world as closely as possible, aiming to test and enact its key features, as described in this paragraph. It is important to stress that artificial nets aim at explaining mechanisms that we witness in the real world in the simplest way possible, thus limitations are tangible. The time steps are set to 10, and the initial number of investors, companies, and funding ties are 11,000, 5,500, and 6,050 respectively. The initial ties connecting investors to companies are arranged in a stochastic fashion to obtain a density of 0.06%, as in the Californian sample. The number of investors is stable during the lifespan considered in the model. Differently, at each step in time, 200 new companies are inserted in the synthetic net and are open to the first round of funding. Investors entering the seed round are assigned stochastically. Contemporarily, other 750 companies are seeking to raise the second funding round. Each company is open to receive funding in two moments in time (first and second funding rounds) but, while the first is certain as it is a prerequisite to be inserted in the network, the second is not. The probability that a company receives funding during the second round depends on its initial investors, and can be depicted by the formula:

$$k_j^{[2]} = g_V * max_c(M_i)$$

where g_V is set to keep the density of the overall network equal to 0.06%, as extracted

from the Californian dataset. M_i is the funding attraction score assigned to investors and evolves at each step in time as described in the methodological section. Following the enactment of the attraction mechanism over time, the final number of companies and edges results in 7,500 and 36,300 respectively, likewise the dataset we use for the analysis.

To test whether the funding attraction mechanism occurs in the real world and it is captured by the funding attraction index, it is necessary to compare the indegrees, outdegrees and score distribution attributed to companies and investors in the artificially created network (displayed in 11) with the specifics of the real-world dynamics (shown in 12). As shown in Figure 11, the indegree distribution of targets is similar to the dataset analysed: most companies have between 1 and 10 investors, with a few outliers involving several tens. As investors and funds are considered, the median degree value is 7, meaning that the investor is connected to 7 different companies, while the range spans from 2 to 14. Results are consistent across the synthetic and the real network. Considering the attraction scores attributed to investors and companies and their respective logarithmic distributions, the two different settings develop nets with corresponding scale-free degree distribution, characterized by a few highly connected hub nodes and many nodes with low degree. The presence of hub nodes in the artificial network is a key feature that distinguishes them from random networks, where the degree distribution follows a Poisson distribution and there are no hub nodes, further supporting the claim of the occurrence of the funding attraction dynamics we test in the current analysis.

9.3 Panel C Statistics

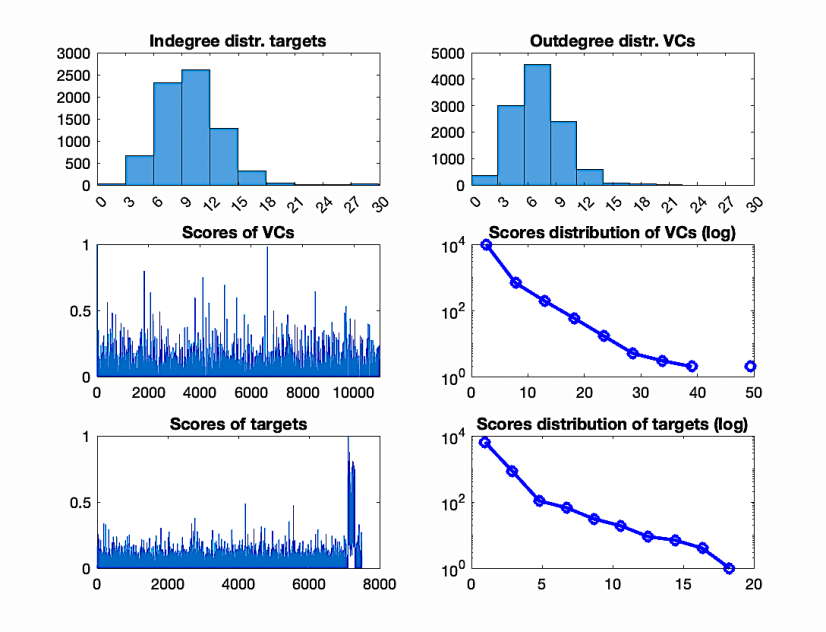


Figure 11: Synthetic Network Statics

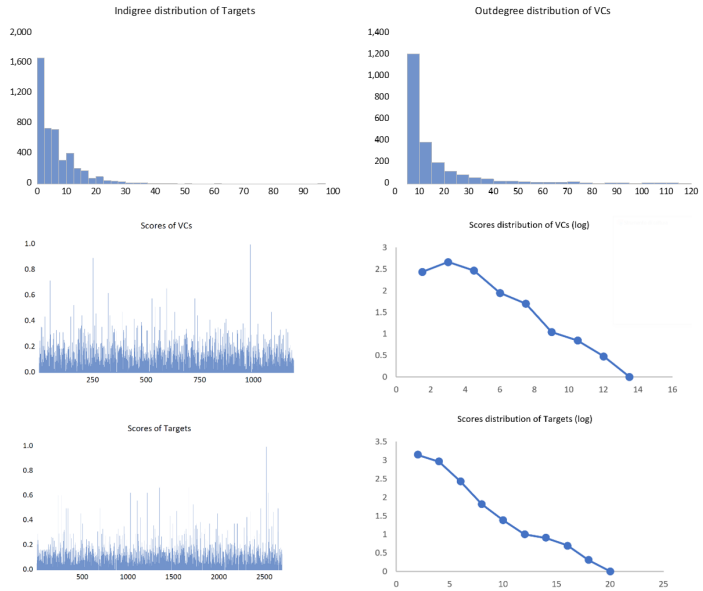


Figure 12: Real-World Network Statistics

Descriptive Statistics Panel C: by Series B

Independent Variables		Full Sample (0&1)				
N. obs:	1665					
	Average	Min	1st Quartile	Median	3rd Quartile	Max
Funding Attraction Index (Norm.)	1.26	0.00	0.00	1.05	2.00	14.50
Topological Overlap Score	0.50	0.00	0.13	0.50	1.00	1.00
Investors' Network Betweeness	0.02	0.00	0.00	0.01	0.02	0.10
Success Rate (IPO/Tot.Inv.)	0.19	0.01	0.13	0.18	0.24	1.00
Years From First Investment (Norm.)	24.58	3.00	11.00	16.00	28.00	212.00
Company's Network Betweeness	0.00	0.00	0.00	0.00	0.00	0.03
No. Investors Series A	9	1	3	7	13	99
Money Raised Series A	14,850,031	30,000	7,000,000	10,500,000	15,550,000	1,200,000,000
No. Investments	917	0	117	361	1,059	11,516

Independent Variables		Not Raised Series A (0)				
N. obs:	822					
	Average	Min	1st Quartile	Median	3rd Quartile	Max
Funding Attraction Index (Norm.)	1.24	0.00	0.00	0.97	1.98	14.50
Topological Overlap Score	0.51	0.00	0.13	0.50	1.00	1.00
Investors' Network Betweeness	0.02	0.00	0.00	0.01	0.02	0.09
Success Rate (IPO/Tot.Inv.)	0.19	0.01	0.13	0.18	0.24	1.00
Years From First Investment (Norm.)	25.26	3.00	11.00	16.00	28.25	212.00
Company's Network Betweeness	0.00	0.00	0.00	0.00	0.00	0.03
No. Investors Series A	8	1	3	7	12	46
Money Raised Series A	12,569,613	100,000	5,737,500	9,525,000	14,000,000	455,500,000
No. Investments	870	0	100	338	906	8,864

Independent Variables		Raised Series A (1)				
N. obs:	843					
	Average	Min	1st Quartile	Median	3rd Quartile	Max
Funding Attraction Index (Norm.)	1.28	0.00	0.00	1.14	2.07	9.67
Topological Overlap Score	0.48	0.00	0.13	0.33	1.00	1.00
Investors' Network Betweeness	0.02	0.00	0.00	0.01	0.02	0.10
Success Rate (IPO/Tot.Inv.)	0.19	0.02	0.13	0.18	0.24	0.72
Years From First Investment (Norm.)	23.91	3.00	11.00	17.00	27.00	209.00
Company's Network Betweeness	0.00	0.00	0.00	0.00	0.00	0.02
No. Investors Series A	10	1	3	7	13	99
Money Raised Series A	17,073,641	30,000	8,171,593	12,000,000	17,450,000	1,200,000,000
No. Investments	963	0	134	397	1,178	11,516

Figure 13: Descriptive Statistics Panel C: by Series B

Descriptive Statistics Panel C: Outcome Variables

		Funding Attraction Index (Norm.)		Money Raised series B	
		N. obs	Average	N. obs	Average
Estimated Revenue Range	Less than $1M	225	1.09	75	274,443,039
	$1M to $10M	1,143	1.30	565	33,622,276
	$10M to $50M	297	1.26	203	47,906,952
No. Employees	1-10	134	0.91	134	28,509,442
	11-50	782	1.22	782	43,772,460
	51-100	309	1.22	309	25,305,199
	101-250	344	1.48	344	36,516,221
	251-500	94	1.44	94	42,636,510
Industry	Artificial Intelligence	139	1.19	58	22,768,364
	Biotech & Healthcare	336	1.21	195	54,117,783
	Business Services	246	1.26	111	24,525,124
	Consumer Services	108	1.40	45	36,346,553
	Financial Services	186	1.39	106	30,620,049
	Hardware	236	1.11	120	26,719,426
	Software	185	1.30	82	24,279,012
	Other	229	1.34	126	48,658,152
Midas	No	1,342	1.15	641	37,921,676
	Yes	323	1.70	202	32,190,201
Number of investors r1	Q1	458	0.60	214	31,131,259
	Q2	428	1.11	220	40,559,544
	Q3	408	1.50	201	31,709,977
	Q4	371	1.99	208	42,409,523
Money raised r1	Q1	421	0.77	152	17,905,083
	Q2	415	1.12	201	20,275,753
	Q3	413	1.60	234	28,061,159
	Q4	416	1.56	256	67,664,455

Figure 14: Descriptive Statistics Panel C: by Outcome Variable