Friends with Benefits:

Social Connection and Venture Capital Investment Efficiency*

Fanqi Meng† Linxiang Ma‡

This version: January 2024
First draft: July 2022

Abstract

This study investigates the impact of social connectedness on venture capital (VC) investments and entrepreneurial performance using innovative data. Our findings reveal that VC investors exhibit a higher propensity to invest in startups located in regions where they have strong social connections, even when these regions are geographically distant. The results hold true in both the U.S. private market and the global private market. Additionally, we observe that VC-entrepreneur matches characterized by stronger social ties positively influence subsequent performance, benefiting both VC investors and entrepreneurial firms. Furthermore, we attribute this positive correlation to the reduction in agency costs and the ability to select higher-quality entrepreneurial firms during the pitching and screening process. Our research underscores the pivotal role of social networks in shaping VC-entrepreneur interactions and post-investment outcomes, providing valuable insights into VC decision-making and performance.

JEL Classifications: G15, G23, G24, G30, G32

Keywords: Social Networks, Social Connectedness Index, Venture Capital; Private Equity; Matching

*For helpful comments, we thank Nick Crain, Garry Twite, Maurice Mccourt, and the comments received from the 2023 Chinese Economists Society Annual Conference and 2023 China Finance Review International & China International Risk Forum Joint Conference

†Department of Finance, University of Melbourne, 198 Berkeley Street, VIC 3053, Australia. Correspondence: fanqim@student.unimelb.edu.au.

‡Department of Accounting and Finance, University of Strathclyde Business School, 199 Cathedral Street, Glasgow, G4 0QU, UK. Correspondence: linxiang.ma@strath.ac.uk

1 Introduction

Venture capital (VC) plays a crucial role in the economy, with VC-backed firms accounting for approximately 50% of initial public offerings (IPOs) in the US (Janeway et al., 2021). These firms also contribute to over three-quarters of the US market capitalization and nearly 90% of reported R&D expenditures among listed firms (Lerner and Nanda, 2020). Despite its significance, the process of VC selection for start-up firms has not received sufficient attention. Existing research suggests that VC investors exhibit a preference for start-ups located in close proximity (Sorenson and Stuart, 2001; Tian, 2011), in countries they trust (Bottazzi et al., 2016), and led by entrepreneurs with similar ethnic backgrounds (Bengtsson and Hsu, 2010; Hegde and Tumlinson, 2014).

This study aims to investigate the role of geographic structure of social connections in facilitating deal formation and enhancing the success of VC investments. Drawing from the social finance literature (Kuchler et al., 2021; Duggan et al., 2016; Bailey et al., 2018; Rehbein and Rother, 2020), we employ the Facebook Social Connectedness Index (SCI) as a proxy to measure the geographic structure of social connections between VC firms and start-up firms at both the county and/or country levels to conduct empirical analysis. Our findings demonstrate that higher social connections between VC firms and start-ups significantly increase the likelihood of deal formation between the two entities. Furthermore, investing in start-ups with higher geographic structure of social connections to VC investors leads to higher investment returns.

The importance and motivation behind this topic stem from notable examples in the industry and several of fieldwork examples. For instance, the fundraising history of Airbnb illustrates how social connections can facilitate the convergence of entrepreneurs and VC investors. When Airbnb was accepted into Y Combinator, Paul Graham, the co-founder of Y Combinator, played a pivotal role in securing venture funding from Sequoia and Y Ventures in the same year. Similarly, Micah Officer, a professor at Loyola Marymount University, suggests that if a private equity firm invests in one portfolio company within a specific industry, it is highly likely they have another portfolio company in the same industry. This suggests a guaranteed synergy spillover, implying that if two portfolio companies operate in the same industry and have top executives who are acquainted with each other, the one with VC financing may assist in introducing VC investors to the other company without VC financing. This matching can be attributed to the power of social connections.

Our impetus for this study is rooted in the social finance literature (Kuchler et al., 2022), which investigates the correlation between institutional investor asset allocation and social connections. Similarly, Nguyen et al. (2023) have observed analogous outcomes, indicating a substantial influence of social connections on asset allocation by VC firms. This prompts our curiosity regarding the potential role of the geographic structure of social connections in shaping the process of matching and screening within private market investments, motivating our exploration of this subject matter.

Our study introduces a novel framework for examining VC investment decision-making, centering on the geographic structure of social connections as facilitators in the VC screening process. Drawing on insights from organizational theory, as discussed by Scott and Cable (2002), the concept of utilizing indirect social ties to overcome information asymmetry has gained prominence. Accordingly, we posit that enhanced indirect social ties—signified by a higher level of geographic structure of social connections—between VC investors and entrepreneurs are likely to contribute to the formation of VC deals. In essence, stronger indirect social connections fostered by a robust geographic structure of social connections positively influence the likelihood of VC deal formation.

The VC-entrepreneur matching process encompasses several pivotal stages, including pitching, screening, contract negotiations, monitoring, and post-investment interactions. These steps are inherently resource-intensive and time-consuming. Given the significant information asymmetry and limited performance history of entrepreneurial ventures, VC investors undertake thorough due diligence to ascertain the genuine value and investment potential of such firms (Gompers, 1995; Sahlman, 1990; Stuart and Sorenson, 2005; Gompers and Metrick, 2001). Our inquiry centers on the potential means to establish indirect social ties between VC investors and potential entrepreneurial firms, which could offer insights into entrepreneurs prior to embarking on the due diligence process. By doing so, information asymmetry can be mitigated, leading to decreased processing costs.

Indeed, this avenue holds promise. Through leveraging social media and networking platforms, information asymmetry between parties can be markedly diminished even before a deal is formalized. For instance, if a VC investor shares acquaintances, mutual friends, or encounters information about an entrepreneur through platforms like Facebook, the information asymmetry lessens. This personal connection empowers VC investors to glean deeper insights into the actual value of entrepreneurial firms ahead of contract finalization (Bottazzi et al., 2016). This proposition is substantiated by survey findings from fieldwork, as exemplified by Scott and Cable (2002). A VC investor corroborates this notion:

> was a controversial figure at the time he approached us for money and we needed to know if he was any good. The trick to finding an answer is to get off the reference list and get to people we know who also intimately know the person in question so that we will get an unbiased reference. We did that with [PERSON G]. There were few people in [PERSON G's] position. He was on [COMPANY H's] executive compensation committee for the board and really knew what was happening on the inside. Since [PERSON G] was also on another company's board with one of my partners we knew him and could ask him things about [ENTREPRENEUR F] that other people could not ask. We made our investment largely because of [PERSON G]. We figured that he had better information than us on [ENTREPRENEUR F] and if he believed in [ENTREPRENEUR F], then we should too. (ENTREPRENEUR F)

In essence, heightened social connections between parties, even when channeled through indirect ties, mitigate the presence of private information, streamline the selection and matching process, and potentially augment the probability of a successful match.

Therefore, we propose our first hypothesis: a higher social connectedness index between the locations of the VC investor and entrepreneur headquarters increases the propensity for a successful match. Such connections simplify and reduce the costs of matching in VC investments by alleviating information asymmetry.

The results provide support for our matching hypothesis, indicating that the level of social connections between venture capital (VC) firms and entrepreneurial firms positively influences the likelihood of a deal being formed between the two entities. Our results demonstrate that an 10% increase in the social connectedness index leads to a 17.8% higher probability of a potential VC-entrepreneur match in the financial market in the baseline regression. This relationship holds not only at the U.S. county level but also extends to the global country level. Furthermore, the linkages between social connections and deal formation are more pronounced when the distance between the VC and the portfolio company is greater.

Given the diminished information asymmetry between the parties, leading to reduced agency costs in the post-investment period, a natural analysis follows the impact on the post-investment performance of entrepreneurial firms. Consequently, we delve into the influence of social connections on the post-investment performance of these firms, building upon our initial hypothesis. Our second hypothesis posits a positive correlation between the post-investment performance of entrepreneurial firms and social connections (Stuart and Sorenson, 2005; Hochberg et al., 2007). Several factors may have led to this hypothesis. Building upon the agency model, our hypothesis leverages the potential of increased social connections between VC investors and entrepreneurs to mitigate agency costs due to decreased information asymmetry (Stuart and Sorenson, 2005; Gompers, 1995), resulting in improved subsequent performance. Secondly, the power of social media and networks increases the costs of deviating from commitments. Within shared social networks, bad news spreads fast compared to conventional communication methods, especially within constrained communities. Increased social connections between investors and entrepreneurs notably reduce the opacity of asymmetric information, subsequently mitigating window-dressing problems and potentially lowering post-investment monitoring costs, resulting in improved performance.

However, Shane and Cable (2002) assert that while social obligations arise from social ties, investors could exploit these relationships to identify superior investments rather than being constrained solely by social obligations. This underscores that investment decisions are underpinned by the intrinsic quality of entrepreneurial firms rather than the pure intensity of social connections. Hence, the heightened post-investment performance of entrepreneurial firms predominantly stems from the firms' inherent qualities upon investment. Furthermore, extant literature establishes that VC investors offer value-added services to their portfolio companies. These services include enhancing corporate governance (Hochberg, 2008), offering professional management services, shaping investment structures, and expanding professional networks (Hellmann and Puri, 2002; Hsu, 2006). Entrepreneurs with high social connections could receive better value-added services from their connected VC investors, contributing to increased synergies, and resulting in better firm performance in the post-investment period.

The outcomes are consistent with our hypothesis, revealing a positive correlation between the post-investment performance of entrepreneurial firms and the intensity of indirect social ties developed between venture capitalists and these firms. Further investigation indicates that while the enhanced performance of entrepreneurial firms doesn't align with reduced monitoring costs, it does correspond to underlying qualities during the investment process. To evaluate post-investment performance, we employ two proxies. Firstly, we estimate the performance of the portfolio companies by analyzing the actual return achieved by each VC investor, aiming to determine if higher social connections result in higher exit returns. Secondly, we employ the average internal rate of return (IRR) reported by each VC investor to measure the relationship between portfolio company performance and social connections. In both cases, we find a positive association between portfolio company performance and the level of social connections between VC and entrepreneurial firms. Specifically, a 1-unit increase in social connectedness is associated with a 12.7% increase in the exit return of the entrepreneurial firm and an approximate 0.371% increase in the average IRR of VC investors. These results remain robust and statistically significant across various control variables and specifications. The stability of our estimations reduces the likelihood of omitted variable bias and potential endogeneity issues. Notably, the coefficients estimated for the monitoring measurement, geographic distance, do not align with the hypothesis that increased performance is solely attributable to decreased post-investment monitoring costs. Consequently, we conclude that the improved performance of entrepreneurial firms is primarily driven by reduced post-investment window dressing costs, diminished agency costs, and enhanced quality during the screening and matching process.

Drawing on the social finance literature (Kuchler et al., 2021), we employ the Facebook Social Connectedness Index (SCI) as our measure of social connectedness. The SCI offers a comprehensive depiction of the global social network structure, given the vast scale and representativeness of Facebook's user base. Specifically, we use the probability of two Facebook users residing in the headquarters-counties of a venture capital (VC) firm and a start-up being connected through a Facebook friendship link to proxy their connectedness.

Compared to other proxies utilized in previous literature, the social connectedness index effectively mitigates the endogeneity issue. This is attributed to its measurement of social connections between locations rather than specific individuals. As a result, the entrepreneur cannot endogenously select the optimal location from potential investors before commencing external funding, thus overcoming the limitations associated with geographic distance measurements. Additionally, the measure of social connectedness between VC and entrepreneur leverages real-world friendships and acquaintances on Facebook, the largest online social networking service globally.

Moreover, even without any additional controls, the social connectedness index incorporates essential demographic characteristics such as education level, wealth, life expectancy, migration patterns, and patent citation information (Bailey et al., 2017). In summary, this measurement framework plays a pivotal role in capturing the significance of social networks in facilitating economic and social interactions.

Nonetheless, it remains plausible to contend that this established relationship between social connections and deal formation could contribute to omitted variable scenarios, potentially engendering endogeneity issues. To counteract these potential concerns, we adopt an instrumental variable (IV) methodology. Specifically, we leverage two instrumental variables: market access, quantified through historical transportation costs dating back to 1920, and a constant transportation cost that remains invariant over time. The IV estimates concur with the inferred causal impact of the geographic structure of social connections on the formation of deals.

Our study makes notable contributions to the venture capital literature from various angles. Firstly, we empirically establish the significance of social connections in facilitating deal formation. Existing research, though limited, has indicated that venture capitalists exhibit preferences for investing in firms led by entrepreneurs of the same ethnicity (Bengtsson and Hus, 2010; Hegde and Tumlinson, 2014) or firms headquartered in countries they have higher trust in (Bottazzi et al., 2016). Building upon these findings, our study complements the literature by highlighting the importance of friendship in the matching process between venture capitalists and entrepreneurs.

Secondly, our work addresses the literature exploring factors influencing venture capital performance (Bottazzi et al., 2016; Gompers et al., 2016) [ADD MORE CITATIONS]. We demonstrate that, in addition to startup human capital, business models, and the degree of similarity and trust between investors and entrepreneurs, social proximity between the parties can significantly impact performance outcomes.

Furthermore, our study enriches the broader social finance literature that investigates the role of social interactions in financial decision-making (Kuchler and Stroebel, 2021). Previous studies have emphasized the critical role of social interactions in shaping households' decisions related to retirement savings plans (Dulfo and Saez, 2002), property investments (Bailey et al., 2018; Bayer et al., 2021), mortgages (Bailey et al., 2019; Maturana and Nickerson, 2019), as well as portfolio allocation decisions for retail and professional investors (Ivkovic and Weisbenner, 2007; Kuchler et al., 2022; Maturana and Nickerson, 2019). By documenting the role of social interactions in facilitating private market investments, our study complements these existing investigations.

The remainder of the paper is organized as follows: Section 2 introduces our sample, outlines variable construction procedures, and provides descriptive statistics. Section 3 presents the results of the selection and matching processes between venture capitalists and entrepreneurs and discusses their implications. Finally, Section 4 concludes the paper.

2 Data and Measure

2.1 VC Data

Our analysis draws upon VC investment data sourced from the Preqin dataset, encompassing venture capital investments made by investors into entrepreneurial firms spanning the period from 1969 to 2022. Preqin offers distinct advantages when contradictory with the VentureXpert and Venture Source datasets. Notably, the former two datasets exhibit issues of incompleteness and inconsistencies vis-à-vis recorded SEC filings. Furthermore, data quality within these datasets has demonstrated diminished accuracy over the preceding years. Additionally, both datasets manifest limited coverage, particularly in capturing new investments. Noteworthy gaps also exist in performance and fund-level data within the former two datasets. A salient strength of the Preqin dataset lies in its ability to transparently identify General Partners (GPs) by fund name, thereby ensuring verifiability, correctness, and comprehensive coverage (Kaplan and Lerner, 2015).

Our primary analysis centers on investments transpiring within the United States, while international deals are subjected to robustness testing. Our sample encompasses 131,198 unique deals spanning the timeframe from June 1st, 1969, to July 28th, 2022, across the United States, encompassing 1,075 counties. Furthermore, our scope extends globally, comprising 340,760 unique VC investments ranging from June 1st, 1900, to July 28th, 2022, spanning 187 countries and encompassing 8,127 cities. All pertinent attributes concerning Venture Capitalists (VCs) and entrepreneurial firms are discernible within the Preqin datasets. Additionally, industry-level accounting information for entrepreneurial firms is sourced from Compustat. information from Compustat.

Figure 1 illustrates the primary geographical distribution of VC investment deals across the United States spanning the years 1969 to 2022. Meanwhile, Figure 2 showcases the principal locations of both entrepreneurial and VC firms' headquarters within the U.S. during the same period. Descriptive statistics pertaining to the geographic allocation of entrepreneurial firms and VC firms across states are presented in Table 1. Panel A presents the ten states with the highest concentration of entrepreneurial firms, while Panel B displays the corresponding top ten states for VC firms. Panel C elucidates the leading ten states in terms of the concentration of VC investments across the U.S., and Panel D details the top ten counties where these deals transpired. Both figures and tables collectively reveal that over 30% of investments are concentrated in California, New York, and Texas, with more than 50% of investments clustering within the top five states. Notably, VC firms headquartered in California, New York, and Texas contributed the most substantial investment sums, accounting for a combined total of $11.5 trillion or 65.77% of the total VC investments amounting to $17.5 trillion made between 1969 and 2022.

Table 2 provides a comprehensive overview of key summary statistics encompassing the social connectedness index, distances at the county-to-county, county-to-country, and country-to-country levels, as well as VC firm characteristics, entrepreneurial firm characteristics, and VC investment characteristics. The primary social connectedness index quantifies the number of Facebook links between the county or country locations of firms' headquarters and those of VC firms, scaled by the product of the populations in these two locations, multiplied by 10^{12}. Geographic distances are quantified using associated coordinates. Notably, both key variables and control variables exhibit a high degree of skewness, either left or right. To address the vast differences in magnitudes, logarithms are applied to these variables, rendering their distributions more meaningful from an economic perspective. This transformation aligns with the distribution of social connectedness, which mirrors that of distance. Existing literature (Tian, 2011) has already underscored the positive correlation between VC investment and the geographic distance between entrepreneurial and VC firms. However, this literature has yet to contend with the endogeneity issues inherent in this relationship. Capitalizing on the analogous data structure, our study furnishes the initial evidence supporting our first hypothesis: that geographic distance may not hold primacy in influencing the matching and selection process between entrepreneurial and VC firms.

Furthermore, Table 2 also presents an array of summary statistics concerning VC and entrepreneurial characteristics. Notably, entrepreneurial firms, on average, have a founding year of 2.48 and undergo approximately 3.58 rounds of investment from 4.65 distinct VC firms. The average funding received amounts to approximately \$288.55 million, with the initial round averaging around \$61.48 million.

2.2 Social Connection Data

To gauge the social connectedness across different spatial units, including U.S. counties, U.S. counties to countries, and country-to-country interactions, we employ the Social Connectedness Index (SCI) initially introduced by Bailey et al. (2018). This index is established through the utilization of anonymized data concerning user activity on social media platforms, whereby each active user is associated with specific geographic locations based on their registered information and prominent interactions within the online sphere. Notably, research by Duggan et al. (2016) and Kuchler et al. (2021) substantiates the efficacy of this index as a reliable proxy for quantifying genuine social connections between two designated locations. This validation stems from the predominant use of platforms such as Facebook by individuals who share real-life connections, as well as the noted similarity in demographic attributes among individuals who form connections on these platforms.

Consequently, the *Social Connectedness Index* between county i and j is formulated in the ensuing manner:

$$\textit{Social Connectedness Index}_{i,j} = \frac{FB\_Connections_{i,j}}{FB\_Users_i \times FB\_Users_j} \tag{1}$$

where $FB_{Users\,i}$ and $FB_{Users}j$ signify the counts of social media users within counties i and j, respectively, while $FB_{Connections}i, j$ represents the aggregate count of Facebook friendship connections established between individuals situated in the two specified locations. Importantly, the *Social Connectedness Index* [1]employed in our investigation has undergone normalization by population size. Consequently, this index encapsulates the relative likelihood of forming connections between users from divergent locales.

To illuminate the intricate interplay of social connectedness in the VC-entrepreneurial firm matching process, we employ visual aids in the form of heatmaps, illustrating our social connections metric within the context of VC investments specifically within the United States. The heatmaps Figure 3 and Figure 4 show the cross-sectional VC investments made in 2020, where varying intensities of color indicate varying degrees of connection strength between the featured locations.

It is pertinent to acknowledge the unique characteristics of venture capitalists (VCs) as relatively modest financial entities. This observation aligns with findings by Kaplan and Lerner (2009), demonstrating that roughly 50% of firm initial public offerings (IPOs) involve venture capital backing. However, in light of the fact that only a mere 0.2% of all firms secure venture funding, the investments of VCs exhibit pronounced concentration within select counties. This phenomenon is particularly pronounced in substantial states, including California, New York, and Texas, irrespective of considering cumulative or year-specific investments. Furthermore, our empirical scrutiny underscores that the phenomenon of social connections cannot be solely attributed to geographical proximity. For instance, King County demonstrates robust connections not only with Cook County, Brewster County, and New York

[1] The Social Connectedness Index data is available at http://data.humdata.org/dataset/social-connectedness-index

County, but these connections persist even as the geographic distances between these locales vary. This empirical finding stands as secondary evidence lending support to our primary hypothesis, asserting the pivotal role of social connections in shaping the matching dynamics inherent in VC investments.

3 Empirical Analysis

This section presents the empirical findings concerning the intricate dynamics of matching and selection processes between entrepreneurial firms and venture capital partners. Our analysis delves into the impact of social connectedness on various facets of this interaction, encompassing the alignment of VC partners with startup founders, the subsequent decisions and behaviors surrounding VC investments, and ultimately, the outcomes derived from these investment endeavors.

3.1 Baseline results on the matching and selecting between VC and startup firms in the United States

Construction of VC Investments Matches Based on our matching hypothesis, we posit that a stronger level of geographic structure of social connection between entrepreneurial firms and venture capitalists enhances the feasibility of a successful match between the two entities. Holding all other factors constant, it becomes more probable for a VC firm to invest in an entrepreneurial firm with which it shares robust social ties. This is due to the resultant reduction in information asymmetry between the two parties, decreased monitoring costs, and streamlined pitching and screening processes. To operationalize the likelihood of VC firms matching with entrepreneurs, we analyze an extensive dataset encompassing 477,982 actual investments made between entrepreneurs and VC firms globally within the time span of 2011 to 2021. Additionally, we consider 592,225,186 hypothetical investments, where we artificially link a VC firm that made an actual investment but with a different entrepreneur in the same year-month to an entrepreneur firm that received funding from another VC firm. In constructing these hypothetical matches, it is imperative to have precise geographic information for each VC firm. Consequently, our sample only includes the known VC firms, enabling the construction of hypothetical matches.

The identification of these hypothetical matches involves a fusion of straightforward, unstructured methods and a ”case-cohort” sampling approach (Stuart and Sorenson, 2001; Bengtsson and Hus, 2010). This methodology, though rigorous, poses computational challenges, time constraints, and potential issues with excessive scale in empirical analysis. On the other hand, it's important to note that the geographic structure of social connections, as derived from Facebook data, remains time-invariant, and the Social Connectedness Index (SCI) is computed using social mapping data from 2021. Despite the demonstrated stability of this social connectedness index, as asserted by the author (Bailey et al., 2020), we limit our panel analysis to the most recent 10 years of data to mitigate potential bias. Therefore, our analysis focuses on the time frame from 2011 to 2021. Our sample can be segregated into two primary categories: actual matches and hypothetical matches. The latter is established by including all VC firms that invested in an entrepreneur within the same year-month but excluding the actual investment in question. Throughout the sampling process, no restrictions were imposed on either VC firms or entrepreneur firms, such as geographical proximity, VC attributes, or firm characteristics. This design results in an actual-to-hypothetical match ratio of approximately 1:1239. The detail of construction on the hypothetical match is shown in **??**

The supplementary online appendix presents univariate mean-comparison tests for both matched VC-entrepreneur pairs and hypothetical VC-entrepreneur pairs. The mean differences between these groups are reported, along with their corresponding t-statistics. Panel A outlines the findings from the main variable mean-comparison tests, while Panel B encompasses the results for significant control variables, VC characteristics, and entrepreneur firm characteristics. Remarkably, the outcomes reveal that the social connectedness index for actual matched pairs substantially surpasses that of hypothetical pairs. Specifically, even without logarithmic transformation, the U.S. county-based social connectedness index for actual matched pairs is 11.18 times higher than that of hypothetical pairs. Even for the weakest social ties among country measures, the social connectedness index for actual matched pairs remains 5.85 times higher than that of hypothetical pairs. The effect size diminishes after applying the logarithmic transformation, but the underlying pattern remains evident. Moreover, we observe that actual matched pairs exhibit a moderate reduction in the geographical distance between VC investors and entrepreneurs, albeit with a lesser magnitude. Given the statistical significance and meaningful economic differences in all mean comparisons, discerning the precise factors contributing to match likelihood while accounting for controls and fixed effects necessitates multivariate regression analysis.

In our analysis, we systematically investigate the impact of social connectedness on the likelihood of matching between VC investors and entrepreneurial firms. To accomplish this, we employ a combination of univariate and multivariate regression models, which encompass key variables and control factors. Our modeling techniques incorporate logit and linear probability models, which are complemented by various cluster error terms determined by our selected fixed effects.

In this context, Table 4 exclusively presents the results of U.S. county-level cross-sectional regression. The dependent variable in this regression is a binary indicator, where an actual match takes a value of 1, and an apocryphal match is represented by a value of 0. It's important to note that these regression results specifically pertain to cross-sectional data from the year 2020. Subsequently, in Table 5, we expand our analysis to encompass panel-level regression, considering VC investments spanning the years 2011 to 2021.

We provide the cross-sectional results initially due to data availability constraints. The Social Connectedness Index (SCI) data, upon which our analysis heavily relies, is not presented in a continuous time-series format; rather, it is available in a static version. The data used in this study was updated until October 2021. Consequently, analyzing VC investments from 2020 is the most appropriate approach, ensuring alignment with the available SCI data.

Our primary independent variable of interest is the Social Connectedness Index (SCI). In addition to this core variable, we introduce Log(1 + Distance) into each regression column to control for the potential influence of geographic distance on the matching between VC and entrepreneurial firms. Furthermore, we incorporate controls for both VC and entrepreneurial firm characteristics, as well as average industry-level factors, while also including various fixed effects. Heteroskedasticity-robust standard errors are presented in parentheses.

The results consistently reveal that the coefficient associated with the Social Connectedness Index is not only positive but also statistically significant. This outcome aligns with our hypothesis, suggesting that a higher degree of social connection between VC investors and entrepreneurs increases the likelihood of a successful match.

3.2 Heterogeneity analysis on the matching and selecting between VC and startup firms in the United States

As shown in Table 4, our analysis begins with univariate regression results in the first column, which excludes any additional controls or fixed effects. These results are statistically significant and economically meaningful, indicating that a one-unit increase in social connections corresponds to a 16.4% higher likelihood of a match between VC and entrepreneurial firms.

To address concerns about omitted variables, we proceed to the second column, incorporating controls for industry, VC, and entrepreneurial characteristics. These results remain consistent, underscoring that our findings are robust and not driven by other potential factors.

In an effort to further explore potential explanations, we introduce additional fixed effects in subsequent analyses. Prior literature has suggested the presence of "home bias," where VC investors tend to favor local investments to reduce monitoring costs (Lerner, 2009; Bernstein et al., 2016). Entrepreneurs may also relocate their businesses closer to potential VC investors (Dahl and Sorenson, 2012), potentially leading to a positive coefficient in our results. However, we account for this by introducing firm and industry fixed effects, capturing location preferences and industry clusters. Encouragingly, these results, as presented in columns 5, 7, and 8, continue to show a positive and statistically significant relationship between social connectedness and match likelihood. This suggests that our findings are not influenced by strategic relocation decisions or industry clustering.

To discern the effect of geographic distance more clearly, we introduce natural logarithms of geographic distance as a control variable. Consistent with the distance literature, our analysis shows a negative correlation between match likelihood and geographic distance. However, it's important to note that the magnitude of this coefficient is substantially lower than that of our primary variable of interest. Additionally, when accounting for county fixed effects, as displayed in column 4, the coefficient for distance becomes statistically insignificant. This implies that the likelihood of a match is not primarily driven by physical proximity. Our regression results substantiate the patterns observed in the heat map, further confirming that social connectedness serves as a superior proxy for assessing the likelihood of matching between VC investors and entrepreneurial firms.

Critics may raise concerns about potential endogeneity in our results, questioning whether the observed outcomes are driven by shared preferences among counties and individuals. For instance, Bengtsson and Hsu (2010) highlight that personal similarities between founders and investors, such as ethnicity or educational background, can heighten the likelihood of a match. To address this concern, the untabulated supplementary analyses that included demographic information, encompassing population, age, employment, income, and education, for each county as additional control variables. Notably, these additional controls did not alter our results, reaffirming the robustness of our findings.

Moreover, Bailey et al. (2017) assert that the Social Connectedness Index data incorporates demographic factors, such as population, educational background, and life expectancy, among the measured locations. Furthermore, the data integrates information on cross-county migration and patent citation. Taken together, our analyses indicate that the relationship between social connection and the likelihood of matching remains unaffected by other potential drivers, as suggested by existing literature.

Furthermore, we explore the potential influence of investor pitching behaviors on our results. A body of literature suggests that successful pitching by VC investors plays a pivotal role in fostering matches (Sorensen, 2007; Cumming and Dai, 2009). Specifically, these studies propose that more experienced VCs have a tendency to invest in superior startups, as their influence and selection capacity enables them to choose promising companies. In light of this, we introduced controls for both VC and entrepreneur characteristics to account for this effect. Remarkably, our results remained consistent and unaffected by these additional variables, reinforcing the robustness of our findings.

Our cross-sectional results, whether with or without supplementary controls, exhibit both statistical significance and economic meaningful. This stability in estimation reduces the likelihood of omitted variable scenarios and potential endogeneity concerns. In summary, we assert that a higher level of social connection indeed enhances the likelihood of a match between VC investors and entrepreneurs, aligning with our initial hypothesis.

Matching and selecting between VC and startup firms in the United States using panel analysis

Furthermore, we have undertaken additional analyses to verify the robustness of our findings. Table 5 presents the results of U.S. county-level panel regressions, encompassing all investments made between 2011 and 2021. This panel analysis is justified due to the considerable stability of the social connectedness index over time. As demonstrated in other studies, Bailey et al. (2021) illustrate that contemporary measurements of social connectedness can predict trade flows in both the 1980s and the present day, while Kuchler et al. (2021) suggest that today's measurements can forecast mutual fund investments in both the 2000s and the present. In our investigation, we have exclusively employed VC investment data from the past decade under the assumption that the social connectedness index displays minimal time-series variation during this period. To account for potential time-varying unobservable effects that might correlate with social connectedness or match likelihood, we have included time-fixed effects, time-interacted fixed effects, and other relevant controls. Notably, the coefficients related to social connectedness exhibit consistent qualitative patterns, displaying even greater statistical significance across the diverse regression results. This collectively strengthens our overall argument that higher levels of social ties correspond to an increased likelihood of a match between VC investors and entrepreneurs. In addition, we have reported cross-sectional regression results for the years 2019 and 2021 in the online appendix, which, as anticipated, demonstrate that the coefficient associated with social connectedness remains positive and statistically significant, further reinforcing our findings.

3.3 Matching and selecting and heterogeneity analysis between VC and startup firms in the global market

The majority of studies examining VC-entrepreneur matching have predominantly concentrated on the United States. Notably, extant literature, including works by Bertoni et al. (2019) and Dai et al. (2011), underscores the disadvantage faced by remote VC investors in the context of matching with entrepreneurs. This disadvantage arises from the prevalence of local bias and the relative "thinness" of VC markets in such remote areas. Thus, we extend our investigation to assess whether social connections continue to hold significance in the VC-entrepreneur matching process at the county-to-country and country-to-country levels. In particular, the country-to-country level analysis excludes VC-entrepreneur matching within the United States.

The cross-sectional results of our county-to-country analysis are presented in Table 6, where, in addition to U.S. county-to-county data, we include VC deals occurring outside U.S. territory. This encompasses cases where VC firms are headquartered outside the United States, and entrepreneur companies have their headquarters in foreign countries. Conversely, Table 7 exclusively employs county-to-country data while excluding county-to-county information. Existing research has already established that VC-entrepreneur matching is more challenging when investors and startup companies are geographically dispersed compared to scenarios where they are concentrated. This is largely attributed to information-related challenges. Greater physical distance leads to increased difficulties in conducting effective pre-selection screening and post-investment monitoring, resulting in higher screening, due diligence, and monitoring costs for long-distance investments and potentially lower VC-entrepreneur matching rates. In our analysis, we posit that heightened social ties increase the likelihood of VC-entrepreneur matching. This is due to the hypothesis that social connectedness enhances information transparency, reduces the likelihood of a window-dressing effect, and increases the cost of deception, particularly in an era characterized by extensive social media coverage. Consequently, we hypothesize that social connectedness assumes an even more pivotal role in facilitating deal formation when substantial geographic distances separate VC investors and entrepreneurs.

The findings presented in Table 6 closely align with our regression results within the U.S. Specifically, the coefficient associated with social connectedness remains positive, statistically significant, and economically meaningful. Notably, while the coefficients pertaining to the primary variable exhibit quantitative and qualitative similarity to those in the baseline regression, the coefficients related to distance appear quantitatively larger and possess stronger statistical significance. It's worth reiterating that our results remain robust even when subjected to a battery of fixed effects tests, including examinations for home bias and potential location endogeneity issues. Importantly, our panel tests, which are detailed in the online appendix, yield results that mirror those of the regression using U.S. data.

Moreover, in order to delve deeper into the association between matching and social connectedness over longer distances, we curate a sample that excludes instances where both VC investors and portfolio companies are situated within the United States. The regression results, as presented in Table 7, affirm the resilience of our hypothesis. These outcomes lend support to our conjecture that social connectedness assumes an even more significant role in VC deal formation when confronted with greater geographic distances and heightened information opacity.

To delve deeper into the association between matching and social connectedness over extended distances, we eliminate samples where both VC investors and portfolio companies are situated within the United States. In Table 7, we present the regression outcomes, demonstrating their robustness to our hypothesis. The findings substantiate our conjecture that social connectedness assumes an even more pivotal role in VC deal formation when confronted with greater distances and potentially heightened information opacity. Specifically, coefficients relating to social connectedness are notably larger and more significant than those derived from the baseline regression results.

For instance, with an increased distance between VC investors and entrepreneur firms, a one-unit rise in social connectedness amplifies the likelihood of matching between VC-entrepreneur by 39.6%, nearly 2.5 times higher than the corresponding magnitude in the regression using U.S. data. The results are also in line with the location bias literature, illustrating that as the distance between investors and start-up companies lengthens, the probability of a match diminishes. To be precise, a one-unit increase in distance between VC-entrepreneur reduces the likelihood of matching by roughly 19.9%. This trend persists across various fixed effects and controls, particularly in columns 5 and 7, each encapsulating the strategic relocation purpose and industry clustering effect. This implies that our findings are not driven by location preferences at the firm level. However, our results do not align with the hypothesis posited in earlier literature (Bengtsson and Ravid, 2009; Tian, 2011; Bengtsson and Hsu, 2010), which suggests that the reduction in matching probability over greater distances is attributable to increased monitoring costs. This discrepancy arises from the fact that coefficients relating to industry-level control, VC characteristics, and financing rounds exhibit inconsistencies across various fixed effects. Notably, the panel test, also available in the online appendix, produces similar outcomes.

To enhance the robustness of our findings, our subsequent analysis delves into whether the results persist when examining data samples involving cross-border connections, which inherently involve greater distances between VC investors and entrepreneurial firms. In Table 8, we present the results of the cross-sectional, country-to-country regression for the year 2020. These findings consistently demonstrate that the coefficients related to social connectedness remain robust across all specifications, exhibiting larger effect sizes and stronger statistical significance compared to the various sample measurements previously presented in preceding tables. In summary, our results, spanning the realms of univariate, multivariate, cross-sectional, and panel analyses, consistently affirm our hypothesis that a higher likelihood of matching exists when VC investors and startup companies share a more substantial social tie, even across extended geographical distances.

3.4 Social connectedness and investment outcomes

In the previous section, we established a compelling relationship between VC investors and entrepreneurs, demonstrating that VC investors exhibit a strong inclination to invest in entrepreneurial firms in which they have strong social ties with. In this section, we delve into an examination of the impact of the geographical structure of social connections on the post-investment performance of these entrepreneurial firms. As we elucidated previously, the heightened geographical structure of social connections tends to mitigate information asymmetry between the two parties, consequently diminishing agency costs during the post-investment phase and ultimately enhancing performance in the aftermath of investment.

Our primary analytical framework centers on the agency model, and we hypothesize that a positive correlation exists between the post-investment performance of entrepreneurial firms and the degree of social connections, which in line with previous literature (Start and Sorenson, 2005; Hochberg et al., 2007). However, the specific channels through which this positive correlation manifests remain undetermined. We posit that these channels might be attributed to one of the following mechanisms:

Firstly, the increase in social connections between VCs and entrepreneurs, which reduces asymmetric information, lowers the likelihood of window dressing issues. This is particularly significant as bad news travels rapidly within shared networks, especially within smaller communities, i.e. private market community. Consequently, the overall effect of an augmented geographical structure of social networks potentially leads to reduced post-investment monitoring costs, thus improving performance during the post-investment phase.

Secondly, the decreased asymmetric information between VCs and entrepreneurs due to increased social connections might enable VC investors to identify and invest in higher-quality entrepreneurial firms instead of being confined to social obligations to choose only connected entrepreneurs (Shane and Cable, 2002). Therefore, heightened social connections may lead to improved post-investment performance due to the intrinsic quality of the entrepreneurial firms, representing another channel contributing to firm performance.

Additionally, well-connected entrepreneurial firms are more likely to benefit from value-adding services provided by their connected VC investors. Prior literature (Hochberg, 2008; Hellmann and Puri, 2002; Hsu, 2006) has established that one of the key contributions VC investors make to their portfolio companies is the provision of value-adding services. These services encompass aspects such as enhanced corporate governance, professional management services, investment structures, and professional networks, all of which can contribute to increased synergies for entrepreneurs, ultimately resulting in better firm performance during the post-investment period.

To discern which of these channels ultimately underlies the positive correlation between social connections and post-investment performance, we conduct an analysis from both VC investors' and entrepreneurial firms' perspectives. Specifically, we investigate how social connectedness influences VC investors' actual returns and how it impacts VC's average internal rate of return. This analysis offers valuable insights into the mechanisms guiding VC investor behavior and contributes to the existing body of literature on how VC investors navigate the competitive landscape of the founder's market.

3.4.1 Performance measure using VC investors' actual return

Preqin's dataset has several notable advantages, one of them being its extensive coverage of venture capital (VC) investments. This dataset provides detailed information at the deal level, encompassing crucial aspects such as the location details of both VC investors and entrepreneur firms, deal dates, deal statuses, deal types, and industry classifications. Moreover, approximately 30% of the recorded deals also furnish exit-related information, which includes the exit date, exit type, and the exact amount denominated in the relevant currency. Leveraging this data availability, we can construct a variable to track the actual returns from VC investments, utilizing the precise investment amounts made by VC investors in target companies and the exit values realized upon the exit of the investment.

A noteworthy advantage of this dataset lies in its comprehensive reporting of exit values not only for successful exits but also for those that ended unsuccessfully, such as write-offs or trade sales. It is imperative to note that our analysis exclusively employs panel regression for two principal reasons. Firstly, this choice is driven by data limitations; as previously mentioned, merely 30% of the VC deals within Preqin's dataset contain exit-related information. Secondly, the characteristics of VC investments entail that these deals can span a duration of up to 10 years, rendering cross-sectional point estimates of social connectedness less suitable for our analysis. Furthermore, the underlying construct captured by our chosen measurement, the geographic structure of networks between regions, exhibits substantial stability over time. In essence, the social connectedness index employed in our analysis is well-suited for gauging social connections and interactions over time, i.e. the panel analysis.

In Table 9, we present the regression results for entrepreneurial firm performance, with returns serving as the dependent variable. The primary independent variable of interest is the social connectedness index, reflecting the degree of social connection between the headquarters locations of two firms. If our hypothesis holds true, suggesting a positive correlation between social connectedness and firm performance, we would expect the coefficient of this variable to be positive. The table indeed demonstrates that the coefficient for social connectedness is not only positive but also statistically significant, achieving significance levels of at least 10%. These results lend support to the notion that the post-investment performance of entrepreneurial firms tends to improve when there are strong social ties between investors and receivers. To provide more precise figures, in the univariate regression model, a 1-unit increase in social connectedness corresponds to a 12.7% increase in VC post-investment returns. After incorporating the control variables, an average 1-unit increase in social connectedness results in a 3.2% increase in returns.

Furthermore, to discern whether the positive correlation could be attributed to the impact of decreased monitoring costs, we introduced geographic distance as an additional independent variable and incorporated investment characteristics as control variables into our analysis, shown in Table 9. If the distance hypothesis holds (Bengtsson and Ravid, 2009; Tian, 2011; Bengtsson and Hsu, 2010), it would suggest that shorter distances between VC investors and entrepreneur firms enable more effective post-investment monitoring, subsequently leading to reduced monitoring costs and improved post-investment performance for the entrepreneur firm. If the monitoring cost claim holds, one would anticipate observing negative and statistically significant coefficients on either the distance variable or the investment characteristics variable.

However, our findings do not align with these expectations. Under the univariate regression tests, the coefficient estimate on distance is positive and statistically insignificant. Furthermore, when examining results across various fixed effects in multivariate tests, the coefficients on investment characteristics exhibit a contrary pattern to what the location bias hypothesis implies. Specifically, firms with a greater number of financing rounds tend to experience higher post-investment returns, while a larger initial rounding size is associated with a decrease in returns. This evidence is consistent with our second claims, which posits that stronger social ties between VC investors and entrepreneurs decreases the asymmetric information allowing VC investors pitching high-quality entrepreneurial firm during the screening process.

In summary, our analysis demonstrates that social connectedness not only enhances the likelihood of matching between VC investors and entrepreneurial firms but also exerts influence over post-investment performance. This phenomenon elucidates why VC investors exhibit a propensity to invest in firms with which they share strong social ties.

3.4.2 Performance measure using VC investors' average internal rate of return

In this section, we employ an alternative measure of VC investors' returns to explore the influence of social connectedness on the outcomes of VC investments. To address concerns related to data limitations and accuracy, we introduce another variable provided directly by Preqin - the VC investors' internal rate of return, hereafter referred to as IRR. IRR is a crucial industry-standard metric for assessing VC investment performance, given the unique characteristics of VC investments, such as their lengthy lifecycles and the privacy-related nature of investment data. One might question why we opt for IRR as a supplementary measure of VC investment returns, given its complex calculation.

Firstly, according to Preqin's data description, most IRR values are reported either directly by the general partners of VC investors or by limited partners, who are investors in VC funds. Moreover, IRR calculations rely on dependable cash flow data generated by the respective VC investors. This reliability factor sets IRR apart from the measure employed in the previous section.

Secondly, IRR serves as a common method for gauging the success of a VC fund. Analogous to other forms of investment, IRR can be interpreted as the annualized return that a VC fund has generated or expects to generate over the course of an investment. Additionally, historical IRR often plays a role in the decision-making process of limited partners when selecting general partners for investment. Given our objective of assessing the relationship between social connectedness and post-investment returns from the perspective of VC investors, IRR emerges as one of the most suitable variables to consider.

Furthermore, recent filings with the SEC highlight the potential for substantial year-to-year variations in IRR. For instance, the IRR reported by TPG, a leveraged buyout firm, fluctuated between 6% and 36% over a 13-year [2] period. Consequently, we utilize each VC investor's average IRR as a metric for post-investment returns.

Table 10 presents the regression results, with the dependent variable being the average IRR for each investor, and the primary independent variable remains the social connectedness between VC investors and entrepreneurs. The outcomes align

[2] The example could be found at https://www.titan.com/articles/venture-capital-irr#:~:text=IRR%20comes%20in.-,What%20is%20venture%20capital%20IRR%3F,typically%20eight%20years%20or%20so.

with our hypothesis, indicating that stronger social ties between investors and entrepreneurs correspond to higher post-investment returns, even when assessed from the perspective of VC investors. Most of the coefficient estimates are not only positive but also statistically significant at the 1% level. For instance, the univariate analysis reveals that a 1-unit increase in the social connection between VC investors and entrepreneurs leads to an approximately 0.371% increase in average IRR.

To further mitigate potential alternative explanations, we conduct multivariate regressions with various fixed effects. Notably, the relationship between the primary variable and VC average IRR remains robust in the presence of control variables and fixed effects. This robustness suggests that our findings are not contingent on external factors. Furthermore, the adjusted R-squared value notably increases from 0.4% to 36.2% after accounting for VC and entrepreneur characteristics, as well as fixed effects. This increase underscores the substantial explanatory power of these additional controls, while the coefficient estimates for the primary variable remain consistent, both quantitatively and qualitatively. These results reinforce the notion that the social connectedness between investors and entrepreneurs is a pivotal factor in explaining VC investment returns.

Likewise, we explore the impact of physical distance on VC investment returns. In line with our previous study, we do not discern consistent patterns in the relationship between geographic distance and VC investment return. The coefficient estimates associated with geographic distance vary and lack statistical significance. This observation contrasts with the local bias literature, which supports the notion that geographic distance negatively correlates with VC post-investment return, even under univariate analysis. Moreover, the coefficient estimates related to VC characteristics do not align with the decreased monitoring cost claim.

In summary, our findings indicate that, on average, VCs can systematically benefit from investing in socially connected entrepreneurial firms. Strong social ties between VCs and entrepreneurs lead to enhanced post-investment performance, due to the intrinsic better quality of the entrepreneurial firms during the pitching process, whether assessed from the entrepreneur's perspective or the VC's viewpoint. Contrary to the "homophily" literature, we find no systematic correlation between geographic distance and post-investment VC return. Furthermore, our results fail to support the hypothesis that VCs can reduce pre-investment screening, pitching, or post-investment monitoring costs through proxies such as geographic distance between VC investors and entrepreneurs. Consequently, our results substantiate our hypothesis that VC investors are inclined to invest in firms where they possess an informational advantage, ultimately resulting in improved post-investment returns due to reduced management costs, better intrinsic quality, and the establishment of trust between investors and entrepreneurs.

4 Conclusion

A growing literature explores how VCs make a selection to invest in entrepreneurs. We contribute to this literature by empirically investigating how the social network, measured using geographic structure, influences the process matching between VCs and entrepreneurs along with the outcomes. Our results show that VC investors are more likely to invest in firms that have strong social ties, which are measured based on firm locations. In addition, given the information advantage gained through social connection and trust built based on the connection, VC and entrepreneur paired in regions with higher social connectedness have high post-investment returns from both VCs and entrepreneur firms' perspectives. Our results are not only applied in U.S. territory but are robust to worldwide VC investments.

Our results in deal formation are consistent with the existing local bias literature that the shorter the distance the higher the probability of a match, but the regression results are qualitatively and quantitatively lower compared to that of social connectedness. However, the subsequent outcome is incompatible with them. Specifically, we find no systematic results on the relationship between VC post-investment return and the proxy based on the geographic location of an entrepreneurial firm and the distance between the VC investor and the firm. Our results suggest that pre-investment pitching and screening costs and post-investment costs are independent of the geographic distance between VCs and entrepreneur firms. Our results indicate that social connectedness is a better proxy for measuring the relationship between VC and entrepreneur than that physical distance.

Overall, our results are consistent with our hypothesis that higher social connectedness between VC-entrepreneur increases the likelihood of a deal match and subsequently increases the post-investment returns.

References

Bailey, M., Cao, R., Kuchler, T., and Stroebel, J. (2018a). The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126(6):2224–2276.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018b). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–280.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., and Stroebel, J. (2021). International trade and social connectedness. *Journal of International Economics*, 129.

Bailey, M. C., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018c). Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*, 32(3):259–80.

Bengtsson, O. and Hsu, D. H. (2012). How Do Venture Capital Partners Match with Startup Founders? *SSRN Electronic Journal*.

Bertoni, F., D'Adda, D., and Grilli, L. (2019). Self-selection of entrepreneurial firms in thin venture capital markets: Theory and empirical evidence. *Strategic Entrepreneurship Journal*, 13(1):47–74.

Bhagwat, V. (2012). Manager Networks and Investment Syndication: Evidence from Venture Capital. *SSRN Electronic Journal*.

Boocock, G. and Woods, M. (1997). The evaluation criteria used by Venture capitalists: Evidence from a UK Venture fund. *International Small Business Journal*, 16(1):36–57.

Bottazzi, L., Da Rin, M., and Hellmann, T. (2016). The Importance of Trust for Investment: Evidence from Venture Capital. *Review of Financial Studies*, 29(9):2283–2318.

Cumming, D. and Dai, N. (2013a). Why do entrepreneurs switch lead venture capitalists? *Entrepreneurship: Theory and Practice*, 37(5):999–1017.

Cumming, D. and Dai, N. (2013b). Why do entrepreneurs switch lead venture capitalists? *Entrepreneurship: Theory and Practice*, 37(5):999–1017.

Cumming, D. J. and Zambelli, S. (2016). Due Diligence and Investee Performance. *European Financial Management*, 23(2):211–253.

Dai, N., Jo, H., and Kassicieh, S. (2012). Cross-border venture capital investments in Asia: Selection and exit performance. *Journal of Business Venturing*, 27(6):666–684.

Engel, D. and Keilbach, M. (2007). Firm-level implications of early stage venture capital investment - An empirical investigation. *Journal of Empirical Finance*, 14(2):150–167.

Fried, V. H. and Hisrich, R. D. (1994). Toward a Model of Venture Capital Investment Decision Making. *Financial Management*, 23(3):28.

Giudici, G., Guerini, M., and Rossi-Lamastra, C. (2020a). Elective affinities: exploring the matching between entrepreneurs and investors in equity crowdfunding. *Baltic Journal of Management*, 15(2):183–198.

Giudici, G., Guerini, M., and Rossi-Lamastra, C. (2020b). Elective affinities: exploring the matching between entrepreneurs and investors in equity crowdfunding. *Baltic Journal of Management*, 15(2):183–198.

Glaeser, E. L., Laibson, D., and Sacerdote, B. (2002). An economic approach to social capital. *Economic Journal*, 112(483):F437–F458.

Gompers, P. and Lerner, J. (2001). The venture capital revolution.

Gompers, P. A., Mukharlyamov, V., and Xuan, Y. (2016). The Cost of Friendship. *Journal of Financial economics*, 119(3):626–644.

Greif, A. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. Technical Report 3.

Gu, L., Huang, R., Mao, Y., and Tian, X. (2022). How Does Human Capital Matter? Evidence from Venture Capital. *Journal of Financial and Quantitative Analysis*, 57(6):2063–2094.

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

Hellmann, T. F. and Puri, M. (2000). The Interaction between Product Market and Financing Strategy: The Role of Venture Capital. *The Review of Financial Studies*, 13(4):959–984.

Hirshleifer, D. A., Peng, L., and Wang, Q. (2023). Social Networks and Market Reactions to Earnings News. *NATIONAL BUREAU OF ECONOMIC RESEARCH*.

Kaplan, S. N. and Lerner, J. (2015). Venture capital data : Opportunities and challenges. NBER Working Paper 22500. *National Bureau of Economic Research*.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022a). Social Proximity to Capital: Implications for Investors and Firms. *Review of Financial Studies*, 35(6):2743–2789.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022b). Social Proximity to Capital: Implications for Investors and Firms. *Review of Financial Studies*, 35(6):2743–2789.

Kuchler, T. and Stroebel, J. (2021). Social Finance.

Lee, H. S. (2017). Peer networks in venture capital. *Journal of Empirical Finance*, 41:19–30.

Nahata, R., Hazarika, S., and Tandon, K. (2014). Success in global venture capital investing: Do institutional and cultural differences matter? *Journal of Financial and Quantitative Analysis*, 49(4):1039–1070.

Neher, D. V. (1999). Staged financing: An agency perspective.

Peneder, M. (2010). The impact of venture capital on innovation behavior and firm growth. *An International Journal of Entrepreneurial Finance*, 12(2):83–107.

SØrensen, M. (2007). How smart is smart money? A two-sided matching model of venture capital. *Journal of Finance*, 62(6):2725–2762.

Sorenson, O. and Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology*, 10(6):1546–1588.

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

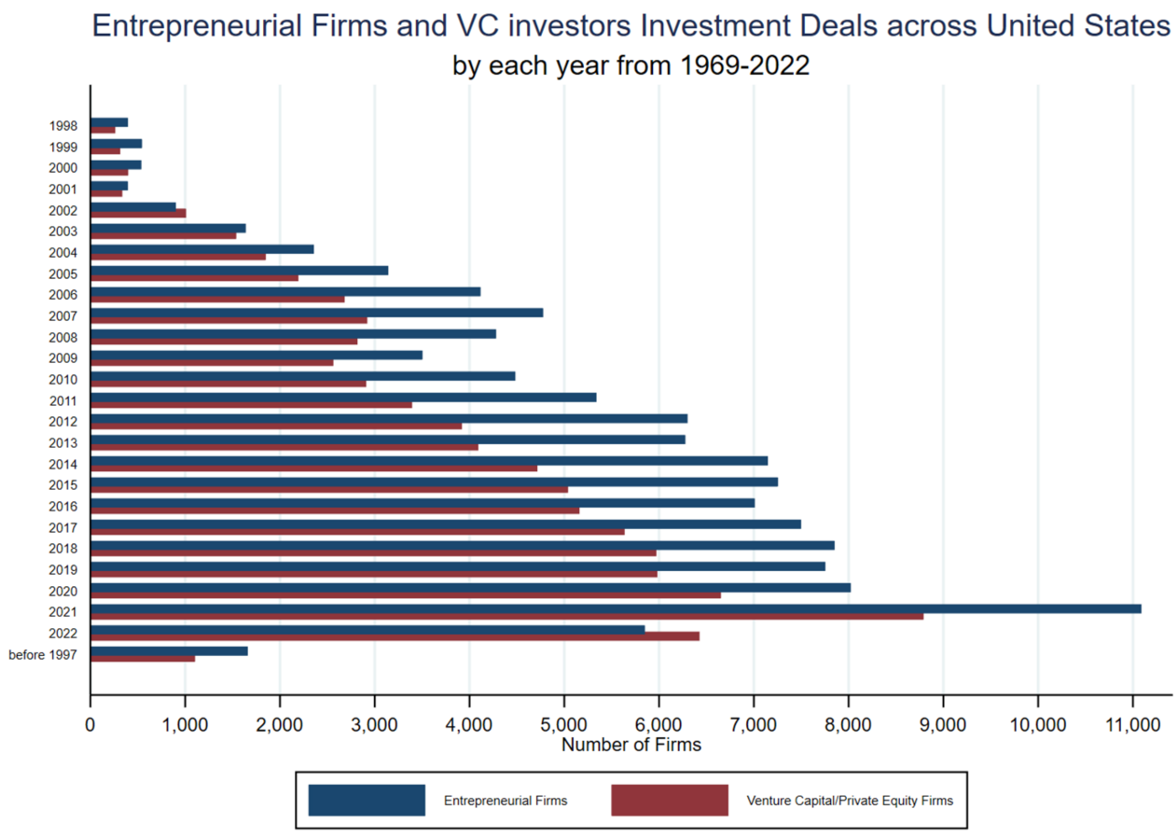
Tian, X. (2012). The role of venture capital syndication in value creation for entrepreneurial firms. *Review of Finance*, 16(1):245–283.

Wang, J., Wang, Y., Wu, C., Yang, X., and Zhao, L. (2021). Social Proximity, Information, and Incentives in Local Bank Lending. *The Review of Corporate Finance Studies*.

Figures

Figure 1: Figure — Number of VC investments across United States.

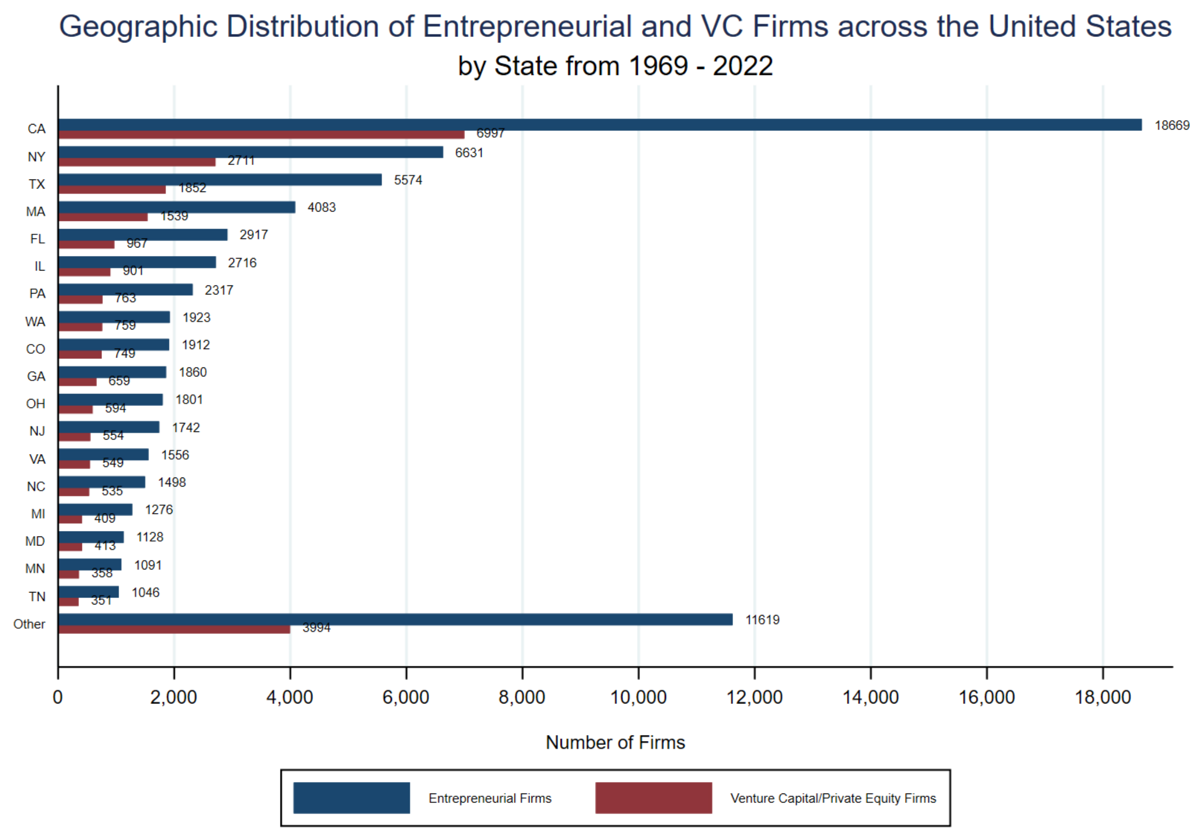
The figure below shows the number of entrepreneurial firms and VC firms involved in Venture Capital, Private Equity, and Private Debt deals each year from 1969 to 2022 across the United States. The blue bar represents the total number of entrepreneurial firms in VC and private equity investment deals each year from 1969 to 2022 across the United States; the red bar represents the total number of VC/PE firms in VC and private equity investment deals each year from 1969 to 2022 across the United States. Data about VC/PE investment deals are obtained from the Preqin database.



.

Figure 2: Figure — Geographic distribution of VC investments across United States.

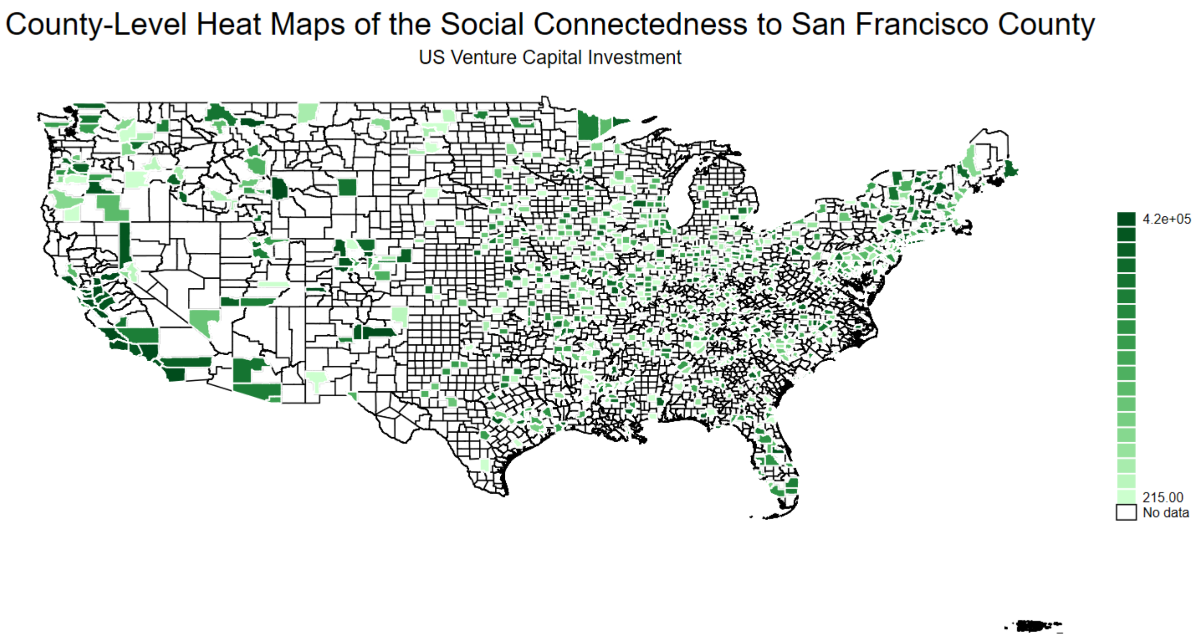
The figure below shows the geographic distribution of the number of entrepreneurial firms and the VC/PE firms who participated in Venture Capital, Private Equity, and Private Debt investments from 1969 to 2022 across the United States, respectively. The blue bar represents the geographic distribution of the total number of entrepreneurial firms in VC and private equity investments from 1969 to 2022 across the United States; the red bar represents the geographic distribution of the total number of VC/PE firms in VC/PE investments from 1969 to 2022 across the United States. Data about entrepreneurial firms and VC firms are obtained from the Preqin database.



.

Figure 3: Figure — Heat map of the social connectedness to San Francisco County..

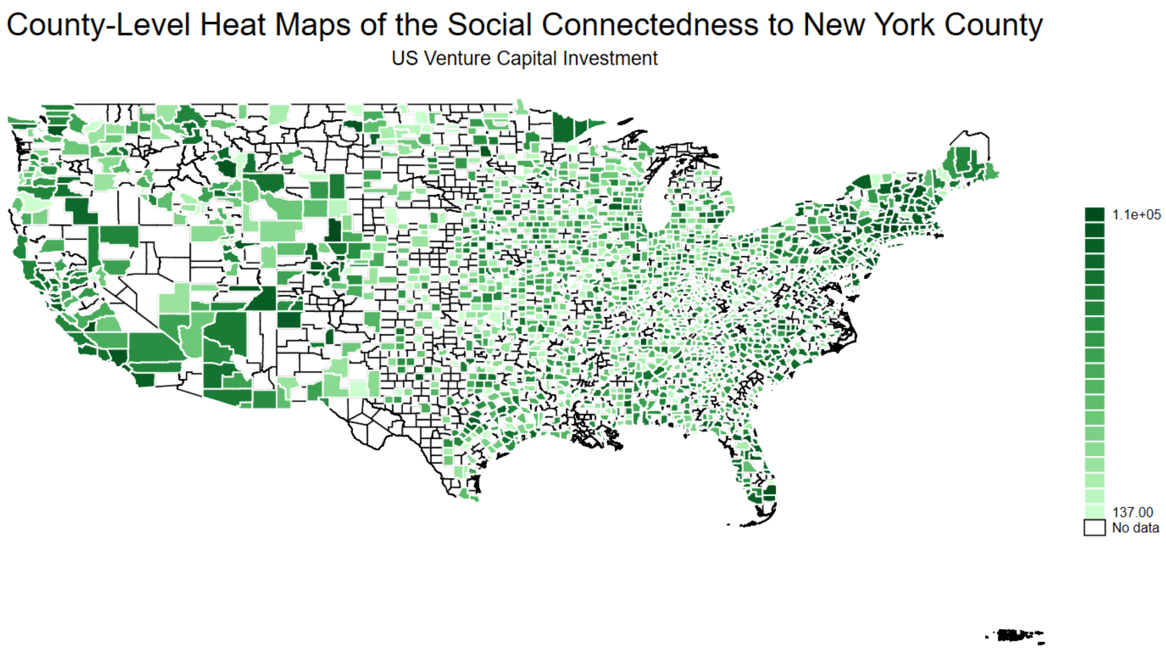
The figure below shows the heat map of the social connectedness of the other county where a VC/PE investment is made with San Francisco County, CA. The blank means no investment was ever made between that county and San Francisco County, CA, from 1969 to 2022. Darker colors represent higher social connectedness between that county to San Francisco County, CA. The map excludes Hawaii and Alaska states, as well as the islands where longitude is above 0 or below -130.



.

Figure 4: Figure — Heat map of the social connectedness to New York County..

The figure below shows the heat map of the social connectedness of the other county where a VC/PE investment is made with New York County, NY. The blank means no investment was ever made between that county and New York County, NY, from 1969 to 2022. Darker colors represent higher social connectedness between that county to New York County, NY. The map excludes Hawaii and Alaska states, as well as the islands where longitude is above 0 or below -130.



Tables

Table 1: Geographic Distribution Summary Statistics for Entrepreneurial Firms and VC Investors across the United States

The table below reports the summary statistics for the sample of entrepreneurial firms and VC/PE firms. Panel A displays the summary statistics for the geographic distribution of entrepreneurial firms across the United States by state from 1969 to 2022. Panel B displays the geographic distribution of VC/PE investors across the United States by state from 1969 to 2022. Panel C displays the geographic distribution of VC/PE investment deals across the United States by state from 1969 to 2022. Panel D displays the geographic distribution of VC/PE investment deals across the United States by County from 1969 to 2022.

Panel A: Geographic Distribution of Entrepreneurial Firms Across the United States by States from 1969 to 2022

State	Number	Percentage
California	18669	26.162%
New York	6631	9.292%
Texas	5574	7.811%
Massachusetts	4083	5.722%
Florida	2917	4.088%
Illinois	2716	3.806%
Pennsylvania	2317	3.247%
Washington	1923	2.695%
Colorado	1912	2.679%
Georgia	1860	2.607%
Other	22757	31.891%
Total	**71359**	**100.000%**

Panel B: Geographic Distribution of VC Investors Across the United States by States from 1969 to 2022

State	Number	Percentage
California	5972	22.946%
New York	3833	14.728%
Texas	2016	7.746%
Massachusetts	1409	5.414%
Illinois	1291	4.960%
Florida	1056	4.057%
Pennsylvania	764	2.936%
Georgia	657	2.524%
Colorado	623	2.394%
Ohio	604	2.321%
Other	7801	29.974%
Total	**26026**	**1**

(*continued*)

Panel C: Geographic Distribution of VC Investment Deals Across the United States by States from 1969 to 2022

State	Number	Percentage
California	41934	32.173%
New York	13038	10.003%
Massachusetts	9276	7.117%
Texas	8729	6.697%
Illinois	4411	3.384%
Florida	4075	3.126%
Pennsylvania	3884	2.980%
Washington	3706	2.843%
Colorado	3374	2.589%
Georgia	3021	2.318%
Other	34892	26.770%
Total	**130340**	**1**

Panel D: Geographic Distribution of VC Investment Deals Across the United States by County from 1969 to 2022

County	Number	Percentage
San Francisco County	12239	10.052%
New York County	10384	8.529%
Santa Clara County	8686	7.134%
Los Angeles County	5384	4.422%
Middlesex County	5243	4.306%
San Mateo County	4435	3.643%
King County	3081	2.531%
Suffolk County	2901	2.383%
Cook County	2839	2.332%
Alameda County	2763	2.269%
Other	63796	52.399%
Total	**121751**	**1**

Table 2: Summary Statistics

The table below reports the summary statistics for all the variables used in this paper, including social connectedness measurement, distance variables, industry level variables, and VC characteristic variables. SCI is defined as the number of Facebook links between an entrepreneurial firm's headquarters' county and a VC/PE's headquarters' county, scaled by the product of the populations in these two counties (multiplied by 10^{12}). County-level SCI is defined as the natural logarithm of the SCI variable. Distance is the distance in miles between an entrepreneurial firm's headquarters county and a VC/PE headquarters' county. Distance is the distance in miles between an entrepreneurial firm's headquarters' county and a VC/PE's headquarters' county coordinates. County-level distance is defined as the natural logarithm of (1 + distance). The same logic applies to county-to-country level SCI measurement and distance measurement. Industrial level variables are calculated based on Compustats SIC code. And VC characteristics are calculated based on the Preqin dataset. Panel A reports the summary statistics used to test for matching tests. Panel B reports the summary statistics used to test for the performance tests.

Panel A: Summary Statistics for Matching Test

Variable	Mean	S.D.	Minimum	25%	Median	75%	Maximum
Log of County Level SCI	8.44	2.04	0	7.15	8.03	9.1	19.95
Log of County to Country Level SCI	10.48	2.8	0	8.17	10.1	13.22	19.95
Log of Country Level SCI	11.28	2.35	4.38	8.74	12.27	12.27	19.04
Log of County Level Distance	7.7	2.28	0	7.61	8.51	8.84	9.43
Log of Country Level Distance	7.57	2.67	0	7.77	8.58	8.94	9.43
Industry Asset Tangibility	0.21	0.15	0	0.09	0.16	0.29	0.94
Industry Market/Book Ratio	1.13	57.22	-1208	0.03	0.2	0.7	4116.13
Industry R&D/Asset	0.21	1.02	0	0.01	0.07	0.23	30.06
Firm Asset	19207.76	78770.62	0	148.11	1094.7	6632.98	5.00E+06
Number of the VC investors	3.55	3.39	1	1	2	5	49
VC age	18.1	9.95	0	10	18	24	75
Number of Financing Rounds	2.94	2.38	1	1	2	4	30
Investment Amount at Round One	412.84	2165.88	0	7.97	25	100	67000
Total Deal Raised by Firm	902.71	3681.13	0.01	15	68.73	316.9	78663.8
Investment Return	2.56	1.76	-8.9	1.39	2.66	3.73	11.31
Average IRR	17.84	17.24	-100	10.84	17.36	24.45	344.56

Panel B: Summary Statistics for the Return Test

Variable	Mean	S.D.	Minimum	25%	Median	75%	Maximum
County Level SCI	8.26	1.69	0	7.17	8.14	8.97	19.95
County to Country Level SCI	9.67	2.16	0	8.08	9.51	11.11	19.95
Country Level SCI	9.04	2.12	3.87	7.75	8.14	9.59	19.77
County Level Distance	8.07	1.84	0	8.11	8.54	8.84	9.43
Country Level Distance	8.45	1.25	0	8.3	8.71	9.01	9.43
Industry Asset Tangibility	0.19	0.14	0	0.09	0.13	0.26	0.93
Industry Market/Book Ratio	1.59	8.43	-20.53	0.06	0.3	0.81	91.11
Industry R&D/Asset	0.14	0.27	0	0.01	0.07	0.24	2.34

(*continued*)

Firm Asset	14249.61	1.70E+05	0	90.27	495.93	2454.92	8.50E+06
Number of the VC investors	2.56	2.38	1	1	2	3	48
VC age	11.15	8.57	2	5	8	16	122
Number of Financing Rounds	3.01	2.47	1	1	2	4	30
Investment Amount at Round One	43.62	390.21	0	0.96	2.5	8	20000
Total Deal Raised by Firm	192.79	1357.14	0.01	4.6	18.54	77.1	1.40E+05

Table 3: Logic to create VC_Entrepreneur_Dummy

The table below illustrates the methodology for generating counterfactual investments between venture capitalists (VCs) and entrepreneurs. A counterfactual match is established by associating a VC investor who made an investment in the same year month, but with a different entrepreneur, to an entrepreneur who did not receive investment from this particular VC. The specifics of our construction process are detailed below.

Actual Investment:

Venture Capitalist	**Portfolio Companies**	**Deal Date**
A	1	2000
A	4	2000
B	2	2001
C	3	2002
B	2	2002
C	2	2002

Creating Hypothetical Investment Based on Actual Investment:

Venture Capitalist	**Portfolio Companies**	**Deal Date**	**VC_Entrepreneur_Dummy**
A	1	2000	Y(1)
B	1	2000	N(0)
C	1	2000	N(0)
A	4	2000	Y(1)
B	4	2000	N(0)
C	4	2000	N(0)
B	2	2001	Y(1)
A	2	2001	N(0)
C	2	2001	N(0)
C	3	2002	Y(1)
B	3	2002	N(0)
A	3	2002	N(0)
B	2	2002	Y(1)
A	2	2002	N(0)
C	2	2002	Y(1)

Table 4: Cross-Sectional County-to-County Matching Test

The table below reports the baseline regression of the matching test. The sample contains 19756 matched investment deals of entrepreneurial firms and VC/PE firms in 2020 across the United States and 11021034 created investment deals in 2020 across the United States. The first two regression estimation applies the logit model, while the rest of the models apply Poisson Pseudo Maximum Likelihood Regressions (PPML). The dependent variable is a dummy variable, where the actual investment takes the value of 1, and the apocryphal investment pair takes the value of 0, which represents a comparable investment made by a VC/PE firm that investments in a different entrepreneurial firm in the same month. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the county level and the natural logarithm of the distance at the county level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	VC_Entrepreneur_Dummy							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
County Level SCI	0.164	0.162	0.141	0.163	0.145	0.162	0.134	0.165
	(35.21)***	(26.75)***	(11.54)***	(8.31)***	(12.67)***	(9.52)***	(15.40)***	(11.42)***
County Level Distance	-0.0823	-0.0731	-0.0984	-0.0868	-0.116	-0.0745	-0.0941	-0.0898
	(-24.83)***	(-17.45)***	(-12.91)***	(-0.67)	(-9.70)***	(-9.07)***	(-15.97)***	(-3.27)***
Industry Asset Tangibility		-0.125	-0.126	-0.128	-0.129	1.942	1.246	1.972
		(-1.46)	(-1.31)	(-6.70)***	(-1.34)	(1.41)	(1.28)	(0.78)
Industry Market/Book Ratio		0.00186	0.00165	0.00186	0.00163	-0.0482	0.00852	-0.0491
		(1.29)	(1.11)	(1.16)	(1.12)	(-0.99)	(1.16)	(-0.86)
Industry R&D/Asset		-0.00818	-0.0150	-0.00551	-0.0138	-0.887	-0.650	-0.865
		(-0.16)	(-0.19)	(-0.07)	(-0.18)	(-0.93)	(-1.36)	(-1.03)
Firm Asset		0.000000209	0	0.000000207	0	0.000000207	0.00000122	0.000000207
		(11.01)***	(.)	(21.45)***	(.)	(13.26)***	(2.70)***	(13.37)***

							Continued	
				VC_Entrepreneur_Dummy				
Number of the VC investors		0.160	0.160	0.159	0.160	0.166	0.168	0.166
		(77.57)***	(68.78)***	(22.25)***	(70.78)***	(28.52)***	(74.94)***	(38.01)***
VC age		0.0126	0	0.0125	0	0.0124	0.000243	0.0124
		(13.69)***	(.)	(4.83)***	(.)	(7.85)***	(0.01)	(7.56)***
Number of Financing Rounds		-0.0186	-0.0181	-0.0186	-0.0182	-0.0210	-0.0183	-0.0211
		(-4.62)***	(-3.42)***	(-2.77)***	(-3.65)***	(-2.91)***	(-3.86)***	(-3.72)***
Initial Rounding Amount		0.00000308	0.00000378	0.00000307	0.00000434	-0.00000340	-0.0000348	-0.00000297
		(0.08)	(0.09)	(0.40)	(0.11)	(-0.09)	(-0.78)	(-0.10)
Total Deal Raised by Firm		0.000000865	0.00000139	0.000000835	0.00000126	0.000000451	0.0000167	0.000000317
		(0.05)	(0.06)	(0.05)	(0.06)	(0.02)	(0.88)	(0.02)
IPO Dummy		0.0618	0.0875	0.0650	0.0892	0.0437	0.0342	0.0440
		(0.33)	(0.43)	(0.22)	(0.44)	(0.44)	(0.18)	(0.23)
Acquisition Dummy		-0.106	-0.1000	-0.109	-0.0989	-0.0909	-0.127	-0.0905
		(-0.58)	(-0.52)	(-0.43)	(-0.51)	(-0.94)	(-0.68)	(-0.45)
Write-Off Dummy		0.0337	0.0732	0.0344	0.0739	0.0238	0.0285	0.0238
		(0.19)	(0.39)	(0.18)	(0.39)	(0.30)	(0.16)	(0.14)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	No	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	No	No	Yes
Industry Fixed Effect	No	No	No	No	No	Yes	No	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes	No
Observations	11040790	5626416	5626408	5626416	5626408	5626416	5424744	5626416

Table 5: Panel County-to-County Matching Test

The table below reports the baseline regression of the matching test. The sample contains # matched investment deals of entrepreneurial firms and VC/PE firms from 2016 to 2021 across the United States and # created investment deals from 2016 to 2021 across the United States. The first two regression estimation applies the logit model, while the rest of the models apply Poisson Pseudo Maximum Likelihood Regressions (PPML). The dependent variable is a dummy variable, where the actual investment takes the value of 1, and the apocryphal investment pair takes the value of 0, which represents a comparable investment made by a VC/PE firm that investments in a different entrepreneurial firm in the same month. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the county level and the natural logarithm of the distance at the county level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	VC_Entrepreneur_Dummy						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
County Level SCI	0.192	0.198	0.155	0.157	0.174	0.174	0.154
	(113.84)***	(89.48)***	(11.81)***	(15.51)***	(8.36)***	(8.61)***	(12.07)***
County Level Distance	-0.0725	-0.0603	-0.0944	-0.101	-0.0732	-0.0738	-0.103
	(-58.96)***	(-38.49)***	(-7.52)***	(-1.90)*	(-0.61)	(-0.69)	(-2.22)**
Industry Asset Tangibility		-0.0375	-0.137	-0.0716	0.715	0.203	0.220
		(-1.38)	(-0.73)	(-2.98)***	(1.78)*	(0.89)	(1.01)
Industry Market/Book Ratio		0.0000628	0.0000539	0.0000598	0.000234	0.000129	0.000112
		(0.61)	(0.36)	(0.69)	(1.87)*	(0.80)	(0.71)
Industry R&D/Asset		0.0356	-0.0231	0.0280	0.0732	-0.199	-0.192
		(4.53)***	(-0.53)	(0.73)	(1.13)	(-2.19)**	(-2.07)**
Firm Asset		0.000000192	0	0	0.000000204	0.000000205	0
		(20.36)***	(.)	(.)	(12.47)***	(6.57)***	(.)
Number of the VC investors		0.0976	0.0863	0.0860	0.0874	0.0887	0.0897

						Continued	
				VC_Entrepreneur_Dummy			
		(191.29)***	(8.89)***	(15.53)***	(7.87)***	(7.48)***	(7.92)***
VC age		0.0181	0	0	0.0135	0.0123	0
		(55.87)***	(.)	(.)	(4.50)***	(3.72)***	(.)
Number of Financing Rounds		0.0317	0.0172	0.0206	0.0218	0.0159	0.0160
		(23.02)***	(2.53)**	(1.66)*	(1.94)*	(1.21)	(1.22)
Initial Rounding Amount		0.0000417	0.0000100	0.0000168	0.0000194	0.0000104	0.0000105
		(5.79)***	(0.38)	(1.80)*	(1.31)	(0.50)	(0.49)
Total Deal Raised by Firm		-0.0000190	-0.00000973	-0.0000136	-0.0000114	-0.00000578	-0.00000578
		(-4.91)***	(-0.69)	(-2.51)**	(-1.12)	(-0.47)	(-0.50)
IPO Dummy		0.123	0.222	0.190	0.0738	0.120	0.121
		(2.36)**	(2.95)***	(1.12)	(0.52)	(1.21)	(0.85)
Acquisition Dummy		0.0491	0.0916	0.0575	-0.0284	0.0296	0.0310
		(0.97)	(1.21)	(0.74)	(-0.32)	(0.48)	(0.39)
Write-Off Dummy		-0.110	0.0298	-0.0132	-0.0778	-0.00419	0.00262
		(-2.23)**	(0.45)	(-0.78)	(-2.42)**	(-0.07)	(0.05)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes
Distance Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	No	No	Yes	No	No	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	Yes	Yes
Industry Fixed Effect	No	No	No	No	Yes	Yes	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes
Observation	75541636	37729977	25351356	25351356	27430058	27430058	23377213

Table 6: Cross-Sectional County-to-Country Matching Test

The table below reports the baseline regression of the matching test. The sample contains 20083 matched investment deals of entrepreneurial firms and VC/PE firms in 2020 across the World and 11257908 created investment deals in 2020 across the World. The first two regression estimation applies the logit model, while the rest of the models apply Poisson Pseudo Maximum Likelihood Regressions (PPML). The dependent variable is a dummy variable, where the actual investment takes the value of 1, and the apocryphal investment pair takes the value of 0, which represents a comparable investment made by a VC/PE firm that investments in a different entrepreneurial firm in the same month. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the county-to-country level and the natural logarithm of the distance at the county-to-country level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	VC_Entrepreneur_Dummy							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
County to Country Level SCI	0.112	0.117	0.120	0.117	0.122	0.118	0.116	0.119
	(29.12)***	(22.64)***	(11.87)***	(359.68)***	(16.74)***	(9.85)***	(15.32)***	(15.55)***
County Level Distance	-0.112	-0.0985	-0.109	-0.105	-0.123	-0.0994	-0.103	-0.108
	(-38.33)***	(-26.44)***	(-16.04)***	(-0.61)	(-1.77)*	(-16.15)***	(-18.96)***	(-0.66)
Industry Asset Tangibility		-0.119	-0.124	-0.121	-0.126	1.813	1.231	1.826
		(-1.40)	(-1.30)	(-48.68)***	(-3.87)***	(1.31)	(1.30)	(0.58)
Industry Market/Book Ratio		0.00187	0.00163	0.00186	0.00162	-0.0493	0.00843	-0.0498
		(1.29)	(1.09)	(1.11)	(1.27)	(-1.02)	(1.19)	(-0.60)
Industry R&D/Asset		0.00146	-0.0113	0.00347	-0.0105	-0.919	-0.674	-0.910
		(0.03)	(-0.15)	(0.04)	(-0.15)	(-1.00)	(-1.42)	(-1.06)
Firm Asset		0.000000204	0	0.000000203	0	0.000000203	0.00000118	0.000000203
		(10.95)***	(.)	(30.36)***	(.)	(12.69)***	(2.63)***	(12.78)***

Continued

		VC_Entrepreneur_Dummy						
Number of the VC investors		0.161	0.161	0.160	0.160	0.166	0.168	0.166
		(78.55)***	(69.16)***	(20.83)***	(24.30)***	(27.73)***	(75.33)***	(19.19)***
VC age		0.0126	0	0.0126	0	0.0125	-0.00256	0.0125
		(13.86)***	(.)	(4.84)***	(.)	(7.78)***	(-0.14)	(4.28)***
Number of Financing Rounds		-0.0170	-0.0172	-0.0169	-0.0173	-0.0195	-0.0173	-0.0195
		(-4.23)***	(-3.27)***	(-2.15)**	(-1.88)*	(-2.62)***	(-3.69)***	(-1.68)*
Initial Rounding Amount		-0.00000372	-0.00000153	-0.00000385	-0.00000120	-0.00000884	-0.0000375	-0.00000868
		(-0.10)	(-0.04)	(-0.35)	(-0.12)	(-0.26)	(-0.87)	(-0.29)
Total Deal Raised by Firm		0.00000329	0.00000442	0.00000326	0.00000434	0.00000290	0.0000193	0.00000284
		(0.20)	(0.21)	(0.18)	(0.31)	(0.15)	(1.03)	(0.10)
IPO Dummy		0.0454	0.0810	0.0484	0.0820	0.0311	0.0307	0.0312
		(0.25)	(0.40)	(0.17)	(0.26)	(0.32)	(0.16)	(0.11)
Acquisition Dummy		-0.130	-0.113	-0.133	-0.112	-0.111	-0.134	-0.111
		(-0.71)	(-0.59)	(-0.55)	(-0.42)	(-1.16)	(-0.72)	(-0.46)
Write-Off Dummy		0.0177	0.0668	0.0185	0.0671	0.00922	0.0263	0.00915
		(0.10)	(0.36)	(0.10)	(0.33)	(0.12)	(0.14)	(0.05)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	No	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	No	No	Yes
Industry Fixed Effect	No	No	No	No	No	Yes	No	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes	No
Observation	11277991	5747322	4299757	5747322	4299757	5742135	970953	5742135

Table 7: Cross-sectional county-to-country without county-to-county matching test

The table below reports the baseline regression of the matching test. The sample contains 50605 matched investment deals of entrepreneurial firms and VC/PE firms in 2020 across the World and 61129857 created investment deals in 2020 across the World. The first two regression estimation applies the logit model, while the rest of the models apply Poisson Pseudo Maximum Likelihood Regressions (PPML). The dependent variable is a dummy variable, where the actual investment takes the value of 1, and the apocryphal investment pair takes the value of 0, which represents a comparable investment made by a VC/PE firm that investments in a different entrepreneurial firm in the same month. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the county-to-country level and the natural logarithm of the distance at the county-to-country level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	VC_Entrepreneur_Dummy							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
County to Country Level SCI	0.396	0.359	0.545	0.426	0.219	0.361	0.485	0.0967
	(165.11)***	(104.22)***	(39.30)***	(31.97)***	(3.46)***	(25.82)***	(65.19)***	(1.01)
County to Country Level Distance	-0.199	-0.203	-0.173	-0.0828	-0.140	-0.202	-0.168	-0.175
	(-144.61)***	(-106.61)***	(-41.44)***	(-6.53)***	(-6.76)***	(-49.75)***	(-54.57)***	(-40.07)***
Industry Asset Tangibility		0.166	0.148	-0.00880	0.0794	0.217	0.850	0.162
		(3.16)***	(2.39)**	(-0.05)	(1.68)*	(0.17)	(2.31)**	(0.12)
Industry Market/Book Ratio		0.00145	0.00103	0.000726	0.000638	-0.0828	0.00404	-0.167
		(1.62)	(1.00)	(1.42)	(0.39)	(-1.53)	(0.66)	(-6.35)***
Industry R&D/Asset		0.0460	0.0174	0.0515	-0.00470	-1.025	0.256	-0.314
		(1.60)	(0.42)	(2.02)**	(-0.14)	(-1.62)	(1.07)	(-0.78)
Firm Asset		0.000000205	0	0.000000230	0	0.000000203	0.000000935	0.000000190
		(13.86)***	(.)	(12.05)***	(.)	(12.13)***	(2.78)***	(2.17)**

								Continued
				VC_Entrepreneur_Dummy				
Number of the VC investors		0.148	0.150	0.149	0.153	0.156	0.162	0.154
		(113.24)***	(94.17)***	(15.28)***	(15.33)***	(24.78)***	(102.19)***	(15.62)***
VC age		0.00950	0	0.0125	0	0.00949	-0.00610	0.00881
		(14.82)***	(.)	(1.44)	(.)	(5.87)***	(-0.76)	(1.01)
Number of Financing Rounds		-0.00424	-0.00264	0.000163	-0.0000788	-0.00753	-0.00292	0.000659
		(-1.67)*	(-0.65)	(0.01)	(-0.00)	(-0.85)	(-0.93)	(0.06)
Initial Rounding Amount		0.0000192	0.0000218	0.0000398	0.0000103	-0.00000646	-0.0000282	-0.0000231
		(0.85)	(1.02)	(1.12)	(0.35)	(-0.23)	(-1.11)	(-0.69)
Total Deal Raised by Firm		-0.00000307	-0.00000222	-0.0000211	-0.00000253	0.00000291	0.0000148	-0.0000869
		(-0.44)	(-0.22)	(-4.05)***	(-0.17)	(0.25)	(2.37)**	(-0.24)
IPO Dummy		-0.0457	0.00249	-0.0903	-0.0419	-0.0397	0.0249	-0.0203
		(-0.38)	(0.02)	(-0.24)	(-0.97)	(-0.37)	(0.18)	(-0.16)
Acquisition Dummy		-0.249	-0.205	-0.203	-0.242	-0.195	-0.0906	-0.191
		(-2.07)**	(-1.65)*	(-0.72)	(-3.86)***	(-1.67)*	(-0.68)	(-3.11)***
Write-Off Dummy		0.00448	0.0548	0.0151	0.0219	0.0271	0.107	-0.00201
		(0.04)	(0.46)	(0.10)	(0.55)	(0.27)	(0.82)	(-0.03)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	No	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	No	No	Yes
Industry Fixed Effect	No	No	No	No	No	Yes	No	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes	No
Observation	61180462	28149058	21682951	28149058	21682951	28149058	3980584	38701643

Table 8: Cross-sectional Country to country matching test

The table below reports the baseline regression of the matching test. The sample contains 50605 matched investment deals of entrepreneurial firms and VC/PE firms in 2020 across the World and 61129857 created investment deals in 2020 across the World. The first two regression estimation applies the logit model, while the rest of the models apply Poisson Pseudo Maximum Likelihood Regressions (PPML). The dependent variable is a dummy variable, where the actual investment takes the value of 1, and the apocryphal investment pair takes the value of 0, which represents a comparable investment made by a VC/PE firm that investments in a different entrepreneurial firm in the same month. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the country-to-country level and the natural logarithm of the distance at the country-to-country level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	VC_Entrepreneur_Dummy							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Country Level SCI	0.411	0.425	0.514	0.418	0.634	0.427	0.478	0.480
	(210.20)***	(138.56)***	(48.87)***	(9.41)***	(10.00)***	(39.84)***	(82.61)***	(40.05)***
Country Level Distance	-0.154	-0.155	-0.127	-0.134	-0.128	-0.152	-0.122	-0.152
	(-110.04)***	(-79.88)***	(-29.31)***	(-10.21)***	(-24.28)***	(-53.16)***	(-38.10)***	(-28.12)***
Industry Asset Tangibility		0.0169	0.00569	-0.0403	-0.000987	-0.229	0.904	-0.285
		(0.32)	(0.09)	(-0.33)	(-0.02)	(-0.17)	(2.47)**	(-0.48)
Industry Market/Book Ratio		0.000931	0.000499	0.000612	0.000533	-0.0701	0.00430	-0.0658
		(1.05)	(0.49)	(2.01)**	(0.31)	(-1.24)	(0.67)	(-2.51)**
Industry R&D/Asset		0.0117	-0.00421	0.0188	-0.000186	-0.970	0.383	-0.966
		(0.41)	(-0.11)	(0.59)	(-0.01)	(-1.55)	(1.48)	(-0.88)
Firm Asset		0.000000208	0	0.000000217	0	0.000000205	0.000000979	0.000000209
		(14.42)***	(.)	(20.06)***	(.)	(12.46)***	(2.97)***	(2.86)***

							Continued	
				VC_Entrepreneur_Dummy				
Number of the VC investors		0.155	0.154	0.153	0.154	0.163	0.170	0.164
		(118.24)***	(90.02)***	(25.89)***	(17.37)***	(26.42)***	(106.32)***	(19.43)***
VC age		0.0116	0	0.0126	0	0.0114	-0.00462	0.0119
		(18.12)***	(.)	(1.68)*	(.)	(6.57)***	(-0.60)	(1.63)
Number of Financing Rounds		0.00143	0.00668	0.00395	0.00850	-0.00381	0.00420	-0.00313
		(0.55)	(1.58)	(0.38)	(0.33)	(-0.46)	(1.35)	(-0.11)
Initial Rounding Amount		0.0000223	0.0000448	0.0000436	0.0000583	-0.0000144	-0.0000327	-0.00000975
		(0.92)	(1.69)*	(2.19)**	(1.54)	(-0.42)	(-1.22)	(-0.28)
Total Deal Raised by Firm		-0.0000191	-0.0000234	-0.0000347	-0.0000318	-7.22e-08	0.0000133	-0.00000208
		(-2.03)**	(-1.02)	(-1.30)	(-1.58)	(-0.00)	(2.03)**	(-0.06)
IPO Dummy		-0.144	-0.0404	-0.140	-0.0328	-0.114	-0.0336	-0.120
		(-1.20)	(-0.31)	(-0.35)	(-0.93)	(-1.66)*	(-0.24)	(-0.75)
Acquisition Dummy		-0.257	-0.249	-0.247	-0.246	-0.183	-0.124	-0.169
		(-2.14)**	(-1.99)**	(-0.94)	(-5.12)***	(-2.41)**	(-0.91)	(-2.14)**
Write-Off Dummy		-0.0235	0.0202	-0.0184	0.0236	-0.000522	0.0514	0.00137
		(-0.20)	(0.17)	(-0.12)	(1.18)	(-0.01)	(0.39)	(0.01)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	No	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	No	No	Yes
Industry Fixed Effect	No	No	No	No	No	Yes	No	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes	No
Observation	61180462	28149058	21682951	28149058	21682951	28149058	3980584	28149058

Table 9: Panel regression Return test/ test for entrepreneur performance

The table below reports the regression in measuring the entrepreneurial firm's performance. The dependent variable is the investment return of the VC/PE firms' investment, calculated using the deal value and exit value reported from the Preqin dataset. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the county-to-county level and the natural logarithm of the distance at the county-to-county level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

				Log_Return			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
County Level SCI	0.127	0.0464	0.0187	0.0251	0.0411	0.0372	0.0206
	(16.10)***	(7.47)***	(1.61)	(2.08)**	(3.26)***	(3.35)***	(1.79)*
Country Level Distance	0.00958	0.00414	0.00348	-0.00397	0.00366	0.00461	-0.00238
	(1.25)	(0.69)	(0.36)	(-0.31)	(0.33)	(0.42)	(-0.19)
Number of the VC investors		0.164	0.152	0.140	0.148	0.162	0.146
		(58.56)***	(30.01)***	(20.74)***	(5.73)***	(15.20)***	(29.50)***
Initial Rounding Amount		-0.0000439	-0.0000589	-0.0000535	-0.0000474	-0.0000528	-0.0000598
		(-8.62)***	(-4.53)***	(-4.39)***	(-2.83)***	(-2.93)***	(-4.89)***
Number of Financing Rounds		0.0730	0.0402	0.0447	0.0626	0.0582	0.0388
		(18.90)***	(5.83)***	(6.52)***	(6.37)***	(5.53)***	(5.49)***
IPO Dummy		0.962	0.697	0.673	0.893	0.925	0.653
		(29.40)***	(18.35)***	(14.11)***	(5.02)***	(11.65)***	(17.99)***
Acquisition Dummy		1.863	1.472	1.499	1.757	1.739	1.408
		(63.69)***	(37.96)***	(34.28)***	(14.09)***	(24.85)***	(36.68)***
Write-Off Dummy		0.707	0.405	0.434	0.665	0.639	0.398

							Continued
			Log_Return				
		(12.76)***	(6.70)***	(6.44)***	(5.72)***	(5.10)***	(7.18)***
Industry Asset Tangibility		-0.796	-0.579	-0.580	-0.704	0.111	-0.217
		(-9.77)***	(-4.40)***	(-3.92)***	(-0.71)	(0.13)	(-0.45)
Industry Market/Book Ratio		0.000642	0.000328	0.000332	0.000535	0.000521	0.000300
		(1.73)*	(1.06)	(1.08)	(2.59)**	(1.99)**	(1.30)
Industry R&D/Asset		0.0299	-0.0187	0.00954	0.0208	-0.00765	-0.0191
		(1.88)*	(-1.22)	(0.35)	(0.34)	(-0.35)	(-1.20)
Firm Asset		-0.00000143	-0.000000489	-0.000000314	-0.00000126	-0.00000141	-0.000000361
		(-8.35)***	(-0.54)	(-0.60)	(-4.68)***	(-4.40)***	(-0.47)
VC age		0.00172	-0.00135	0.00152	0.00142	0.000520	0.00298
		(1.71)*	(-0.22)	(0.21)	(0.92)	(0.37)	(0.36)
Total Deal Raised by Firm		-0.00000572	0.0000140	0.0000128	0.000000512	0.00000178	0.0000130
		(-1.62)	(1.68)*	(1.48)	(0.04)	(0.16)	(1.55)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes
Distance Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	No	No	Yes	No	No	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	Yes	Yes
Industry Fixed Effect	No	No	No	No	Yes	Yes	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes
Observation	17308	14694	14771	14772	14682	14680	14754
R^2	**0.020**	0.467	0.571	0.557	0.503	0.519	0.596

Table 10: Panel regression Return test/ test for VC average IRR

The table below reports the regression in measuring the entrepreneurial firm's performance. The dependent variable is the average IRR for VC/PE firms, calculated using the available IRR of each VC/PE firm reported from the Preqin dataset. The main independent variables are the natural logarithm of the social connectedness index (SCI) at the county-to-county level and the natural logarithm of the distance at the county-to-county level. Industry-level control variables include average industry asset tangibility, average industry market-to-book ratio, and average industry R&D ratio. The VC characteristics variables include the VC's asset under management, VC age, and success rate. The firm characteristics include the number of VCs invested in, the number of financing round the firm received initial rounding amount, and the total funds raised by the firm. The fixed effects for each regression include VC firm fixed effect, county fixed effect, industry fixed effect, and VC firm and industry fixed effect. Data about entrepreneurial firms and VC/PE firms are obtained from the Preqin database, and the data about the industry average are obtained from Compustat. Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	Average_IRR						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
County Level SCI	0.371	0.369	0.294	0.375	0.375	0.372	0.381
	(9.54)***	(7.00)***	(2.22)**	(3.00)***	(4.91)***	(5.57)***	(3.06)***
Country Level Distance	0.194	0.160	0.0934	-0.144	-0.00622	0.0251	-0.132
	(5.23)***	(3.07)***	(0.66)	(-0.87)	(-0.04)	(0.26)	(-0.80)
Number of the VC investors		-0.00770	-0.0723	-0.0311	-0.00891	-0.0782	-0.0709
		(-0.24)	(-0.53)	(-0.24)	(-0.29)	(-1.60)	(-0.57)
Initial Rounding Amount		0.000163	-0.0000187	0.0000248	0.000191	0.000197	-0.0000158
		(3.12)***	(-0.15)	(0.39)	(2.24)**	(2.33)**	(-0.12)
IPO Dummy		1.794	0.238	0.193	2.024	1.691	0.235
		(5.65)***	(2.15)**	(1.26)	(5.77)***	(5.42)***	(2.00)**
Acquisition Dummy		1.126	0.340	0.254	1.259	1.138	0.338
		(4.78)***	(2.20)**	(1.06)	(5.21)***	(4.30)***	(2.02)**
Write-Off Dummy		-0.0556	0.0788	0.104	0.193	0.179	0.0753
		(-0.18)	(0.61)	(0.54)	(0.57)	(0.57)	(0.61)

						Continued	
				Average_IRR			
Industry Asset Tangibility		-2.539	1.553	0.775	1.126	-2.311	1.593
		(-3.61)***	(0.30)	(0.09)	(0.28)	(-0.63)	(0.31)
Industry Market/Book Ratio		-0.0000681	0.00235	-0.000419	-0.000203	-0.000290	0.00235
		(-0.05)	(1.58)	(-0.54)	(-0.19)	(-0.26)	(1.41)
Industry R&D/Asset		0.0255	0.288	0.0386	-0.0762	-0.121	0.306
		(0.17)	(0.72)	(0.10)	(-1.12)	(-1.40)	(0.72)
Firm Asset		0.0000138	0.0000145	0.0000148	0.0000145	0.0000130	0.0000145
		(7.54)***	(2.15)**	(2.24)**	(4.03)***	(4.48)***	(2.15)**
VC age		-0.209	-0.272	-0.269	-0.205	-0.187	-0.273
		(-19.71)***	(-8.35)***	(-8.46)***	(-6.14)***	(-10.08)***	(-8.83)***
Number of Financing Rounds		0.277	0	0	0.243	0.155	0
		(6.81)***	(.)	(.)	(2.44)**	(2.70)***	(.)
Total Deal Raised by Firm		-0.0000837	0	0	-0.0000948	-0.0000853	0
		(-2.51)**	(.)	(.)	(-3.87)***	(-2.81)***	(.)
Control	No	Yes	Yes	Yes	Yes	Yes	Yes
Distance Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	No	No	Yes	No	No	Yes	Yes
VC Firm Fixed Effect	No	No	Yes	Yes	No	No	No
County Fixed Effect	No	No	No	Yes	Yes	Yes	Yes
Industry Fixed Effect	No	No	No	No	Yes	Yes	Yes
VC Firm * Industry Fixed Effect	No	No	No	No	No	No	Yes
Observation	21379	11462	8229	8229	11436	11434	8229
R^2	0.004	0.052	0.362	0.362	0.076	0.091	0.362