Investment Proximity and Venture Capital Returns

CNV Krishnan
Case Western Reserve University

Daniel Nguyen
Case Western Reserve University

Using a large dataset of 82,818 investments in start-up firms from December 1965 through August 2019, we document that the biggest group of investors are venture capital firms (about 79% of all observations). The majority of in-state investments are found clustered in California (with the San Francisco VC Hub), Massachusetts (with the Boston VC Hub) and New York. Early round investors get a higher return than late stage investors, and IPO exits entail the highest return. We find that geographical proximity is significantly and positively associated with cumulative returns or annualized returns, when we use actual distance (in km) between the VC investment firm HQ and portfolio firm HQ, or a cross-region indicator variable when the portfolio firm HQ is out-of-state as compared to VC investment firm HQ. The relation between investment proximity and returns continues to be robust even after controlling for selection bias, and in a matched-sample analysis where we pair each proximate investment with a matched distant investment. We find that VC Director and other outside directors as a proportion of all Directors on the boards on portfolio companies is significantly higher, on average, for in-state and shorter distance investments as compared to out-of-state and longer distance investments, indicating higher involvement of VCs in their more proximate investments.

Keywords: VC investments, returns, cumulative returns, annualized returns, geographical proximity, distance, cross-state investments, VC directors, VC hub, VC favored industries, matched-sample analysis, selection bias control

INTRODUCTION

The extant literature has argued that investors have a bias towards geographically proximate investments (Coval & Moskowitz, 2002; and Cumming & Dai, 2010). Berstein & Townsend (2016) argue that venture capital (VC) investors prefer on-site involvement, which can be achieved by investing in proximate portfolio companies or in regions with direct, short flights. The preference to invest in relatively nearby companies may help VCs obtain superior returns (Barry et al., 1990; and Ellis et al., 2020). Economic rationale for this may include agglomeration effects such as cost savings or economic spillovers. Chen et al. (2010) documented that 49% of VCs in the United States are located in only three regions: San Francisco, New York, and Boston, while 81% of start-ups are located in the top 12 cities in the United States. With such a high concentration, it is reasonable to deduce that both investors and entrepreneurs can

benefit from knowledge-spillovers and cost-savings. Chen et al. (2010) argues that a one standard deviation increase in the number of VC firms within a region will increase VC investments in the area by 49.7%.

However, there are evidence that suggest otherwise. Chen et al. (2010), for example, documents that for VCs located in the three VC hubs of San Francisco, New York, and Boston, most of their outsized returns come from investment outside of their HQ or even branch-office location, in part because venture capital firms may lower the threshold for investment quality in areas where they invest multiple times, whereas they would keep a higher threshold for the more distant investments. Other papers have suggested that the amount of cross-border investments has been growing as local venture capitalists may face disadvantages due to their lack of experience in VC investments (see, for example, Chemmanur, Hull, & Krishnan, 2016). Home markets are increasingly saturated with investors, VCs need to broaden their geographical scope and focus more on international investment opportunities. Indeed, other state/country investment proportion is more than half of all venture capital investments, in our sample. Thus, the link between investment proximity and venture capital investment returns, to our mind, is still an open question. Even more important is the explanation for any such finding. That is our main goal in this paper, using a long time-series of data.

We, however, start by documenting detailed descriptive statistics on all types of venture capital investments, using a large dataset of 82,818 transactions, representing 7435 different venture capital investors in startups, spanning a 54-year period from 1965 through 2019. We document that the biggest group of investors are, by far, venture capital firms (about 79% of all observations) as compared to investments by Corporations, Universities, Angel Investors etc., while the two presumably most successful exits – IPOs and acquisitions - comprise about 26% and 29% of all exits respectively. We classify late-stage investments as those happening after 9th round (after the seed-capital stage, start-up capital stage, early state, first stage, second stage, expansion stage, third stage, mezzanine stages) and these are relative few in number (only about 1% of all observations), but shows a slightly increasing trend in recent years. This is in line with statistics that show late stage funding rounds are much less in number compared to the early-stage funding rounds (GlobalData, a leading data and analytics company). The number of venture investments were few in the 60's and 70's, but peaked in the 90's because of the tech boom. Cross-region (across state/country) investment proportion shows a declining trend, but still remains high at around 60% of all investments. The majority of in-state investments are found clustered in California (with San Francisco VC Hub), Massachusetts (with Boston VC Hub) and New York. Out-of-state investments are more widely spread-out over the map, suggesting that regardless of the clusters, VCs make investments over a wide range of region in the United States. The trend of returns generated from early round investing to late round decreases generally: early investors do get a higher return than late stage investors, and IPO exits entail the highest return for the investor both in terms of cumulative returns and annualized returns, justifying the conclusion that IPO is the most sought after exit. University-fund investments yield the highest return both in terms of cumulative returns and annualized returns, while venture capital firm investments (which, as discussed above, comprise of the dominant portion of all investments in our dataset), the second highest return.

We next come to our main research question: the relationship between proximity of investments and returns. Our two proxies for geographical distance, are an indicator variable "cross-region" when the portfolio firm headquarters (HQ) is out-of-state as compared to VC investment firm HQ, and the actual distance (in km) between the portfolio company HQ and the VC investment firm HQ. We use two alternative measures of investment returns – cumulative returns from financing round to exit, and annualized returns from financing round to exit. We analyze associations between returns and geographic proximity for all data, as well as for data segregated (i) by exit types – all exits and only successful exits via IPOs and acquisitions, (ii) by investor type – Venture Capital firms, Corporations, Universities, Angel Investors etc., (iii) by entry stage – early-stage investment or late-stage investment, (iv) by whether the VC firm HQ location is in a previously-documented VC hub or not (San Francisco, New York Boston, Palo Alto and Menlo Park (see Chen et al., 2010)), and (v) by whether the investment is to a previously-documented favored VC industry (Communications and Networking, Biopharmaceuticals, Media and Content, Software, Healthcare Services), or not, as robustness checks.

Using both univariate and multivariate tests, that controls for other possible determinants of returns such as annual GDP Growth, annual industrial production growth, and annual consumer sentiment growth, as well as controlling for industry and exit-year fixed effects, we find that the cross-region variable is strongly associated with returns, being negative and statistically significant in the full sample, and across all subsamples discussed above. The distance variable is also negative and statistically significant in the full sample and in many subsamples.

The relation between investment proximity and returns continues to be robust even when we control for selection bias (using Instrumental Variables and Limited Information Maximum Likelihood estimation procedure), and in a matched-sample analysis where we pair each proximate investment with a matched distant investment.

Examining reasons for why there is a significant positive association between investment proximity and returns, we find that VC director and other outside directors as a proportion of all directors on the boards on portfolio companies is significantly higher, on average, for in-state and short distance investments as compared to out-of-state or longer distance investments, indicating better involvement by VCs in their more proximate investments. We find that the percentage of VC-appointed and other outside directors in a portfolio company is also significantly and positively associated with both cumulative and annualized returns, indicating better governance, and better monitoring by VCs in their more proximate investments, in line with extant literature that argues that VCs can also take an active role in monitoring and controlling their portfolios (Gompers, 1995; Gorman & Sahlman, 1989; Hellmann & Puri, 2002), and Lerner (1995), who postulates that VCs will more likely to sit on the board of their portfolio companies if the geographical distance between the investor and investee is less.

DATA AND VARIABLES

Our data comes from *VentureSource*, and contains 172,093 financing rounds involving start-ups that are backed by venture capital firms, over a 54-year period from December 1965 through August 2019. From the original dataset, we exclude all observations that do not include information on the startups' valuations at different times, which we use to calculate investors' return, from the financing round to exit. We also get information on the type of investors that invest in each financing round. Since there can be multiple investors in each financing round, the final dataset contains 82,818 transactions, representing 7435 different venture capital investors.

Return, Distance, and Control Variables

Cumulative returns is defined in terms of investment exit from the last financing round. It's noteworthy that the majority of these "defined" exits are actual exits via mergers & acquisitions, initial public offerings, or failure. Only a small percentage of the exits are growth-round or some other types of financing rounds (about 14% of our full sample of investments). Observing that the cumulative return in our data has a wide variation, especially for earlier-round investments, we normalize the returns by computing the natural logarithmic of valuation at exit divided by valuation at the financing round. For robustness, we also use annualized log return to address the concern that different investments are made over different periods.

The last financing round available in the data is our proxy for investor firm's exit. There are 48 different types of exits in our dataset. We categorized these exits into 5 main types: Mergers & Acquisitions (M&A), Initial Public Offerings (IPO), Growth, Failure, and "Others" that comprise private investments in public equity (PIPE), public to private (PTP), recapitalizations (recaps), restarts, and financing rounds post recaps. Categories are defined in the Appendix.

To study the geographical proximity effects on private investment returns, we use of an indicator variable: "cross-region" flag. Cross-region takes the value of one, when the start-ups' main location's (headquarters, HQ) state is different from that of the investor, and zero otherwise. Similarly, cross-country flag will have a value of one, when the start-up's country location is different from that of the investor. In addition, we will also use actual distance (in kilometers) between the investor and the portfolio company,

calculated as the distance between the two cities in which these the investor firm and the portfolio company HQs are located.

As control variables in our regression specifications, we use the US Annual GDP growth rate, and the US Annual Industrial Production growth rate, as these variables can affect the returns earned by venture capital investors. Moreover, the US Annual Consumer Sentiment Growth can also affect returns by affecting investor sentiment in the market. Additionally, the industry of the portfolio firm, as well as the year of the financing exit from the portfolio firm can both affect returns earned, hence we include startups' industry-segment fixed effects, and the exit-year time fixed effects. Finally, we collect information on Investor capital under management, Investor industry preference, and Investor investment-stage preference, which are used as Instrumental Variables (IV) in one of our robustness checks, all collected from *VentureXpert* database. Descriptions of these variables is in the Appendix.

Descriptive Statistics

Table 1 shows the number of investment observations for each investor type – Venture Capital firms, Angel investors, Corporations, Investment Banks, Leasing Companies, Limited Partners, Mezzanine/Buyout Funds, Incubators, Public Sector Organizations, Small Business Investment Companies (SBICs), University funds, and Others, while some are unidentified (we denote all investors as "venture capital" investors through the paper).

The total number of investment observations, as mentioned before, are 82,818, of which the two biggest group of investors are Venture Capital firms (which are dominant with about 79% of all observations) and corporations (about 9% of all observations). The table documents the exits – the two most successful ones – IPOs and acquisitions, comprise about 26% and 29% of all exits respectively, while failures comprise around 32%. Also reported are number of observations in each investment round. The very early round investments, in rounds 1, 2, and 3 are around 20% each. Late-stage investments are classified as those happening after 9th round, and these are relative few in number (about 1% of all observations).

TABLE 1
DESCRIPTIVE STATISTICS

This table shows the number of observations for each Investor Type – Venture capital firms Angel investors, Corporations, Investment Banks etc.; by exit type – IPO, M&A, etc., and by investment round type – 1st round investment through 9th round, and late stage (9+ round). The total number of observations in 82,818 taken from *VentureSource* database spanning the period 1965-2019.

Exit type	Number
Failure	26,135
Growth	7,353
IPO	21,480
MA	23,862
Other	3,988
Total	82,818

Investor type	Number
Venture Capital firms	65,196
Angel investors	1,445
Corporations	7,115
Investment Banks	1,468

Investment Round	Number
1	17,011
2	19,073
3	17,101
4	12,299
5	7,605
6	4,300
7	2,359
8	1,249
9	812
Late Stage (9+)	1,009
Total	82,818

Leasing	254
Limited Partner	429
Mezzanine/Buyout Fund	2,861
Public Sector Organizations	105
SBIC	1034
University	317
Other	1,927
Non-identified	667
Total	82,818

Table 2 shows the time-series descriptive statistics of these venture capital investments. The number of venture investments were few in the 60's and 70's, but peaked in the 90's presumably because of the tech boom. The bulk of these investments were made by venture capital firms, through all years. Late stage investments, though few, shows a slightly increasing trend in recent years. Investments from the VC Hubs remain at less than 50% of the total sample though all years, while investments to favored VC industries peaked in the 90's because of the tech boom. Cross-region investment proportion shows a declining trend, but remains at around 60% of all investments. This is the main variable of analysis in our paper.

Table 3 shows descriptive statistics of Return, our performance measure, defined as Valuation at Exit round divided by Valuation at financing Entry round. Logarithmic return is natural log of this return, and annualized return is this return divided by investment period in years.

TABLE 2
TIME SERIES DESCRIPTIVE STATISTICS OF VENTURE CAPITAL INVESTMENTS

This table shows the descriptive statistics of all venture investments in portfolio companies in our final dataset, period by period – the 60's, the 70's, the 80's etc. through 2019. Of the total number of investments observations shown in the first row, the percentages of investments made the main investor types, by financing round – early or late stage, by investors from the "VC Hub" locations or not, to favored VC industries or not, and whether the investment is cross-region or not, are shown thereafter. Cross-region investments are those when the state of the investor's main office location is different than that of the portfolio company. All variables are defined in the Appendix.

	1965-1970	1971-1980	1981-1990	1991-2000	2001-2010	2011-2019
Investment Observations	15	7	8052	37700	24645	12399
% by Venture Capital Firms	53%	71%	85%	81%	77%	71%
% by Corporations	7%	14%	4%	8%	9%	13%
% by Angel Investors	0%	0%	1%	1%	3%	4%
% Early Stage (1-9 round)	100%	100%	100%	100%	98%	96%
% Late Stage (9+ round)	0%	0%	0%	0%	2%	4%

% from Favored VC location (VC hub)	47%	14%	43%	39%	33%	42%
% to Favored VC industries	33%	0%	41%	57%	50%	42%
% Cross Region	87%	86%	59%	58%	60%	59%
% Not Cross Region (In State)	13%	14%	41%	42%	40%	41%

TABLE 3
DESCRIPTIVE STATISTICS OF THE MAIN DEPENDENT AND EXPLANATORY VARIABLES

This table gives descriptive statistics of the variables that will be used for analyses in this paper. It contains number of observations (N), mean, minimum, maximum and standard deviation (SD) of the variables described. Return is defined as Valuation at Exit Round divided by Valuation at Entry Round. Logarithmic return is natural log of this return, and annualized return is this return divided by investment period in years. Distance is calculated as the distance between the city of the investor's main office location (HQ) and that of portfolio company (HQ). Cross-region is defined those if the state of the investor's main office location is different than that of the portfolio company. All variables are described in Appendix.

	N	**Mean**	**Min**	**Max**	**SD**
Return	82,818	5.29	-1	14,273	100.23
Log Return	82,818	0.56	-13.3	9.56	1.30
Annualized Log Return	82,818	0.23	-4.98	6.61	0.59
Investment Round	82,818	3.19	1	23	2.02
Distance (km)	82,818	1,448	0	12,666	1606
			Cross-region Investments		
	N		**No**		**Yes**
	82,818		33,925		48,893

Distance is calculated as the distance between the city of the investor's main office (HQ) location and that of the portfolio company. *Cross-region* is a dummy variable that is equals 1 if the state of the investor's main office (HQ) location is in different state/country than that of the portfolio company. Table 3 shows that the raw return (cumulative) has a skewed distribution with high standard deviation. So, it may be appropriate to use log returns, which is what we do. We report all result using annualized log returns also, to control for differences in investment horizons. Cross-region investments are more than half, about 59% all observations. The mean distance of investment is about 1450 km.

Figure 1 plots the areas of cross-region investments and in-state investments on the US map, over all the years of our dataset. Black circles represent in-state investments, and red circles out-of-state investments. Bigger circles represent a higher number of investments made in the areas. The majority of in-state investments are found clustered in California (with San Francisco VC Hub), Massachusetts (with Boston VC Hub) and New York. Outside these three main locations, other areas have relatively little in-state investments, showing high clustering of VC investments. Out-of-state investments are more widely spread-out over the map. Thus, although out-of-state investments occur more in the in-state investment areas, there are also investments in other states.

FIGURE 1
GEOGRAPHIC LOCATIONS OF IN-STATE AND OUT-OF-STATE (CROSS REGION) INVESTMENTS

The top panel shows the map of in-state investment, defined as those when the portfolio company HQ is in the same state as the investor firm HQ, and the bottom panel shows the map of investments for cross-region investments, when the portfolio company HQ is not in the same state as the investor firm HQ. Larger circles show greater number of investments. International investor and portfolio companies not shown. The total number of observations in 82,818 taken from *VentureSource* database spanning the period 1965-2019, of which 5256 observations are international.

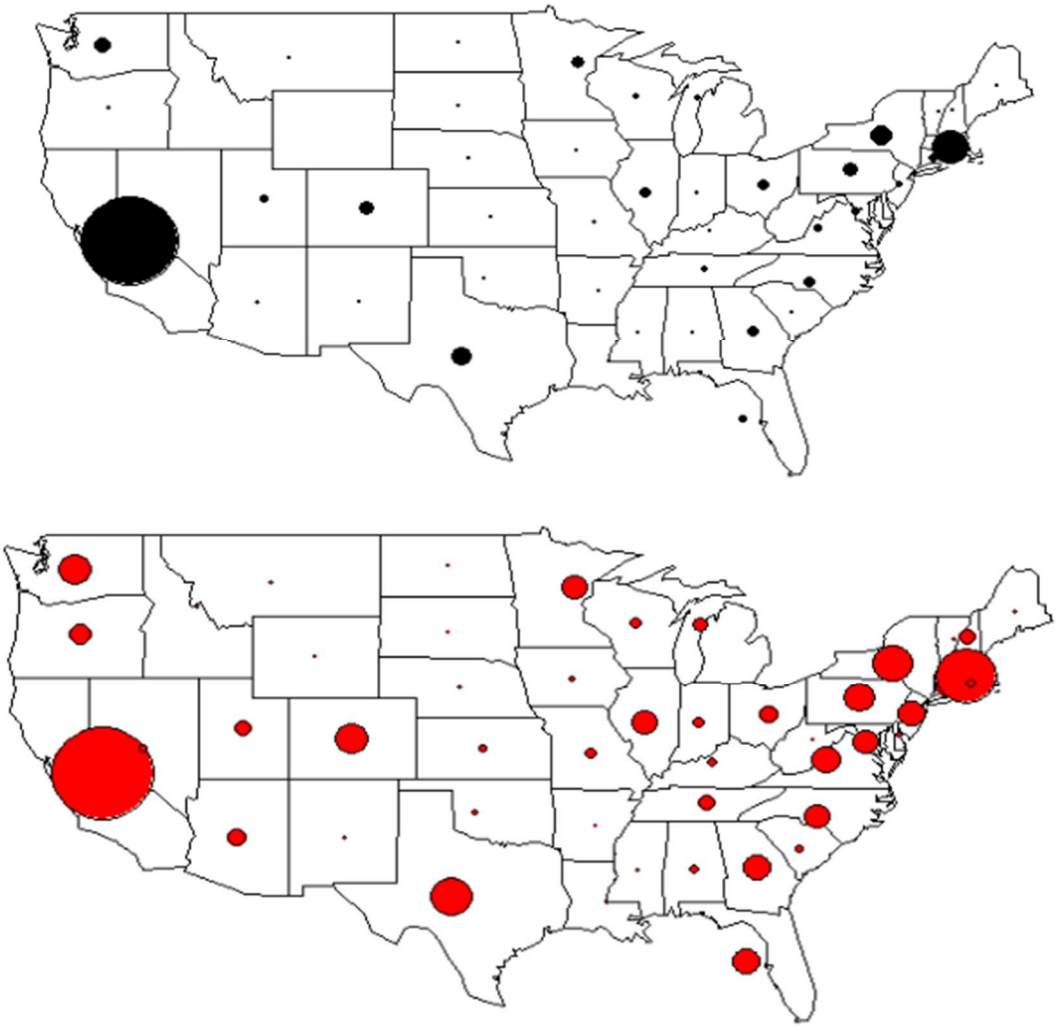
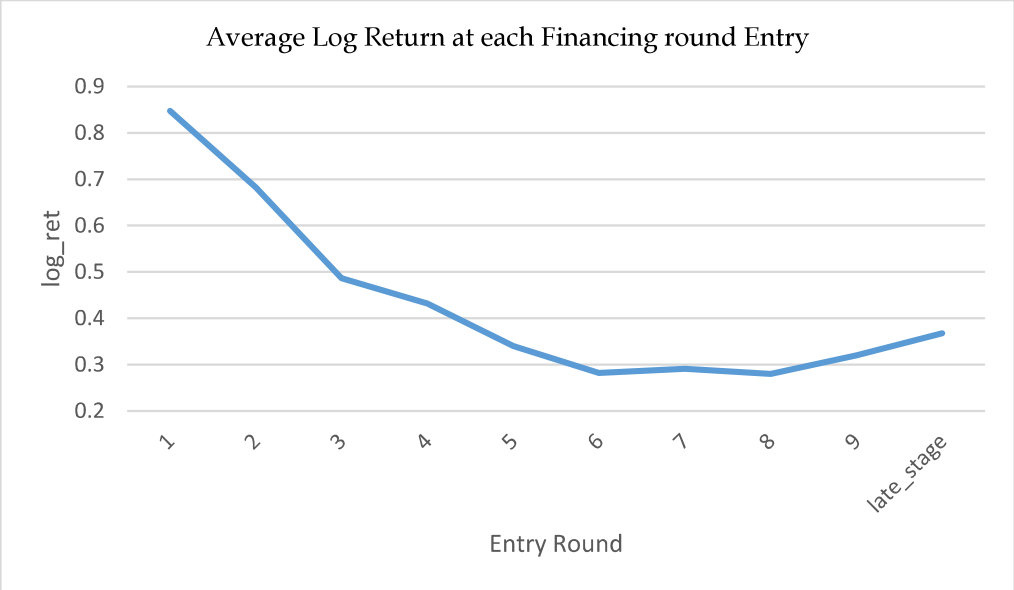


Figure 2 shows the trend of returns generated from early round investing to late round, by plotting the average log return by each entry round. As seen in the chart, there's a general downward trend in the data, signaling that early investors get a higher return than late stage investors, both in terms of log returns and annualized log returns.

FIGURE 2
RETURN BY ENTRY ROUND

The top plot shows cumulative log return by financing entry round, and the bottom plot the annualized cumulative return by financing entry round. The total number of investment observations in 82,818 taken from *VentureSource* database spanning the period 1965-2019. Variables are defined in the Appendix.



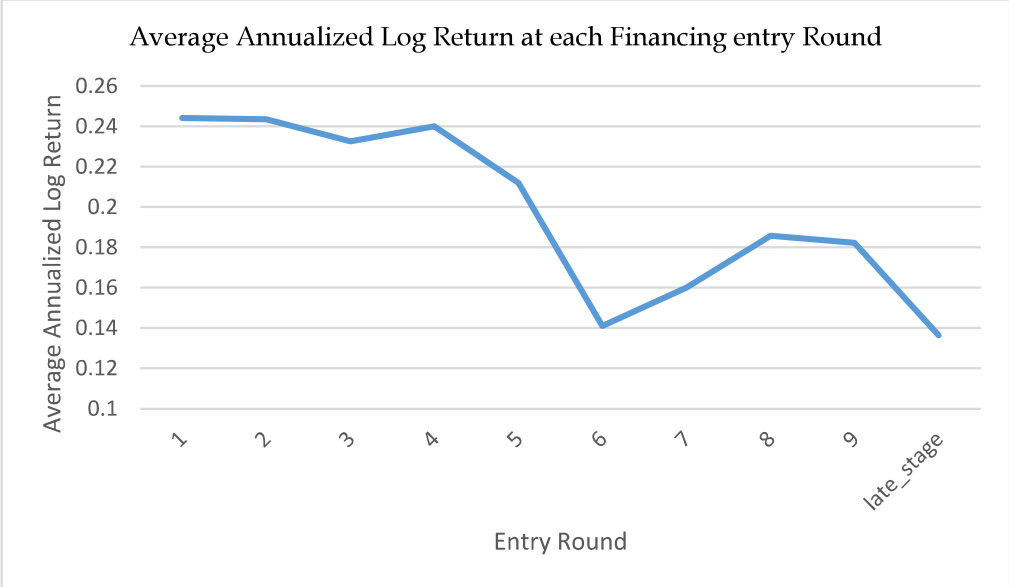
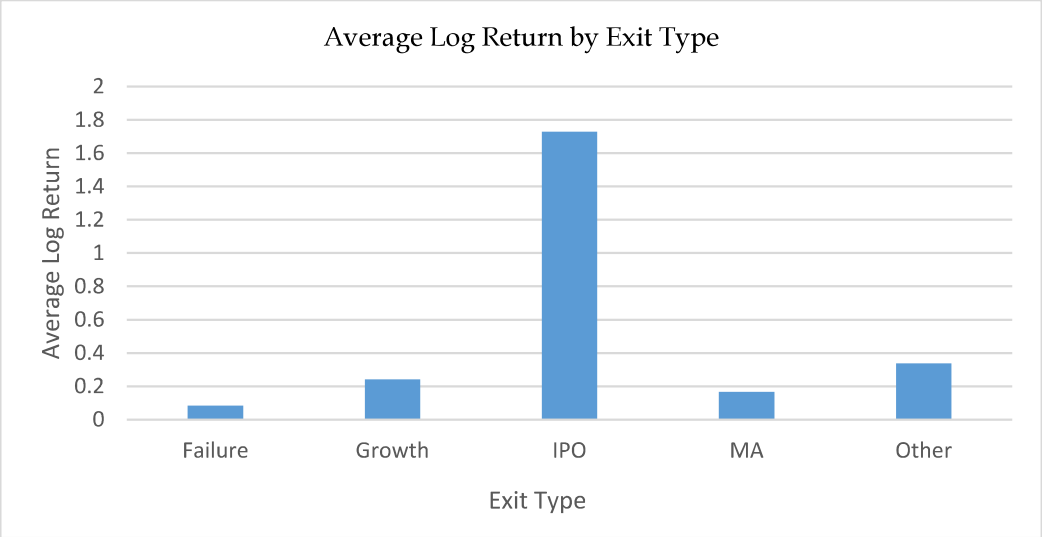


Figure 3 plots the mean returns by exit type. The figures show that IPO exits entail the highest return for the investor both in terms of cumulative log returns and annualized log returns, justifying the conclusion that IPO would be the most sought after exit. Acquisitions are also a desired exit, although the returns are much lower compared to that of IPOs. Failures, as expected, yield the lowest return on investment. Figure 4 plots returns by investor type. University fund investments yield the highest return both in terms of log returns and annualized log returns, while Venture Capital firm investments (which, as discussed above, comprise of the dominant portion of all investments in our dataset), the second highest return.

FIGURE 3
RETURN BY EXIT TYPE

The top plot shows cumulative log return by exit type, and the bottom plot the annualized cumulative return by exit type. The total number of investment observations in 82,818 taken from *VentureSource* database spanning the period 1965-2019. Variables are defined in the Appendix.



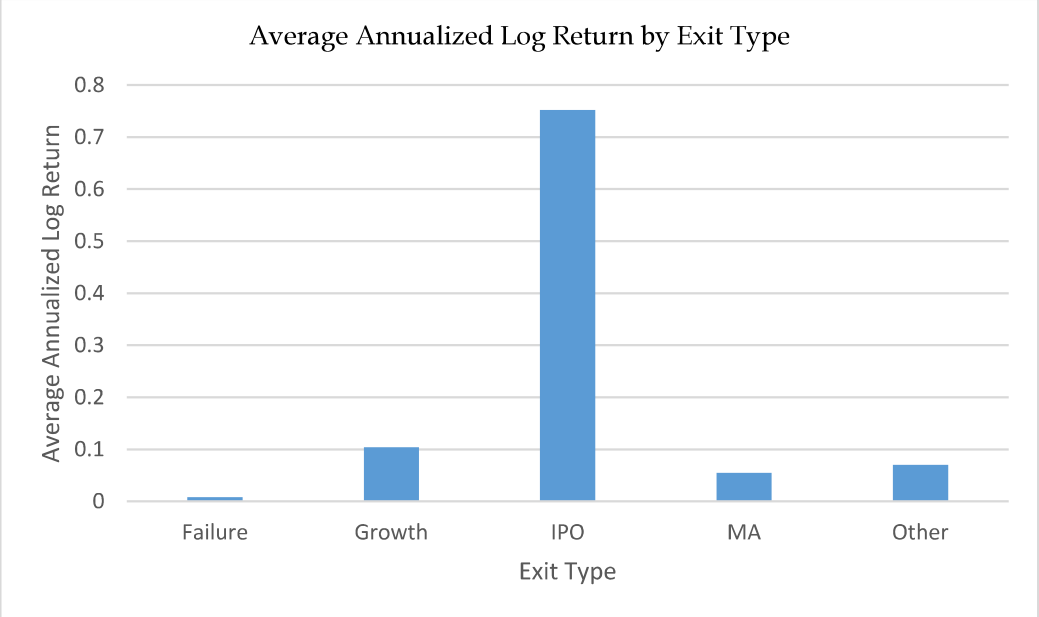
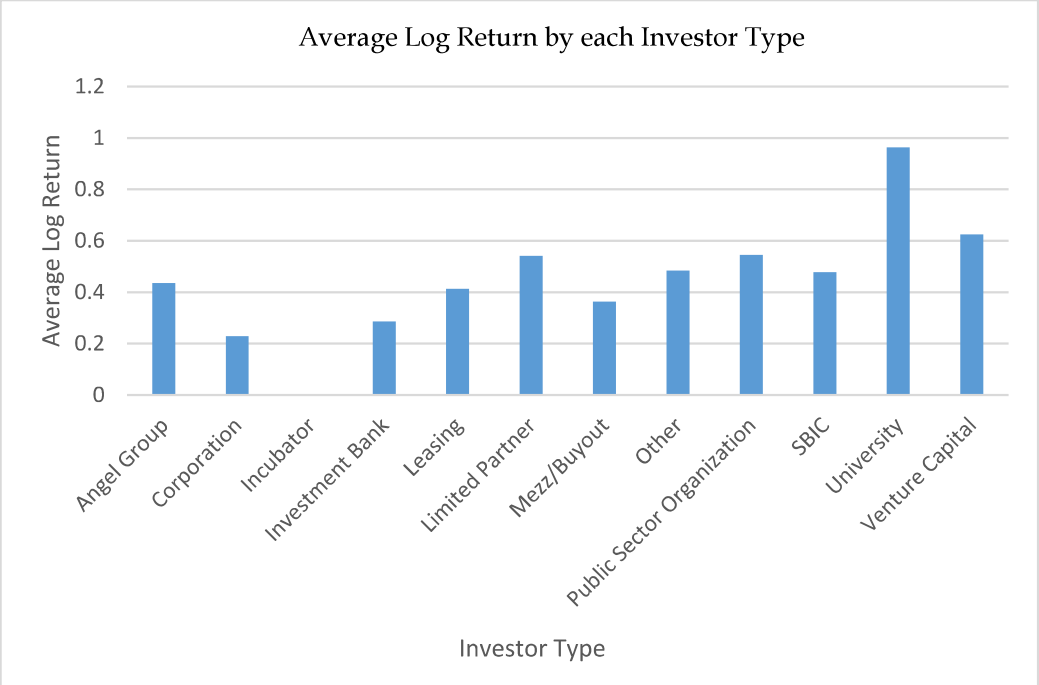
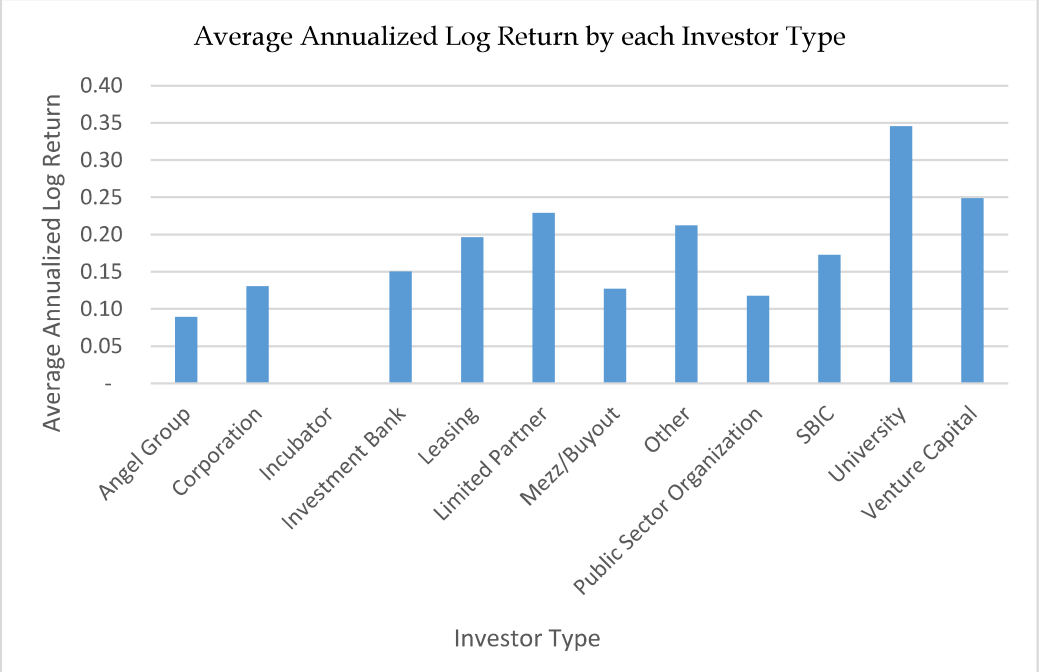


FIGURE 4
RETURN BY INVESTOR TYPE

The top plot shows cumulative log return by investor type, and the bottom plot the annualized cumulative return by investor type. The total number of investment observations in 82,818 taken from *VentureSource* database spanning the period 1965-2019. Variables are defined in the Appendix.





DETERMINANTS OF RETURNS

Univariate Analysis

Table 4, Panel A, compares the mean returns for investments made by the various types of investors, as well as the *t-statistics* of the difference of means with that of Venture Capital firm investor, while Panel B compares the mean returns for investments made by the various types of exits, as well as the *t-statistics* of the difference of means with that of IPO exit, the most sought-after exit.

There is a wide variation in average returns when the data is segregated by the type of investors with average log return ranging from 0.00048 to 0.96 and average annualized log return range from 0.00 to 0.35. The standard deviation for log returns (cumulative returns) is 1.31 and for annualized log returns is 0.59. Since there is a wide variation in return, the effect of investment proximity on returns would depend on the investor. Thus, in later tests, we re-examine our main results by investor type. Similarly, there is an even wider variation in the average cumulative returns, ranging from 0.08 to 1.72 for log return and from 0.008 to 0.75 for annualized log return, by exit types. In later tests, we also examine our main results by exit types also, for robustness.

Panel A shows that the mean log returns as well as the mean annualized log returns are significantly more for the venture capital firm investor as compared to most of the other types of investors, except for University fund investors, whose returns are significantly more than those of venture capital firm investors. Panel B shows both the mean log returns as well as the mean annualized log returns are significantly more for IPO exits as compared to all other exits.

As mentioned in the literature review section, Chen et al. (2010) examines the performance of venture capital investments from the VC centers of San Francisco, New York and Boston (the "VC Hubs"). We find that the top 3 clusters for VCs includes Palo Alto and Menlo Park and thus will include these two additional cities in our "Favored VC Location". We also control for industries that venture capitalists prefer to invest in, which have shown high clustering of such investments in the past (see, e.g., Hochberg, Ljungqvist, and Lu (2007)). These industries are "Communications and Networking", "Biopharmaceuticals", "Media and Content", "Software", "Healthcare Services". In later tests, we will refer to these industries as "Favored VC industries".

Our key explanatory variables in this paper are the *Cross-region* dummy variable that equals 1 if the state of the investor's main office (HQ) location is in a different state/country than that of the portfolio company, and 0 otherwise, and distance (in km) between the investor HQ and the portfolio company HQ. We test the difference of mean log return using *t-test* and test the difference of median using *Mood's test* with a cutoff level of 5%. For univariate tests using distance, we define "short-distance" for distances below or equal to the overall-sample median distance between the investor HQ and the portfolio company HQ and "long distance" as those above the median. In robustness checks, we examine results not only for our full sample but also for various subsamples of data as discussed above: by financing entry rounds, by types of investors, by types of exits, whether from VC hubs or not, and whether to favored VC industries or not.

TABLE 4
DESCRIPTIVE STATISTICS OF INVESTOR TYPES AND EXITS

Panel A of the table shows average of two types of returns – cumulative and annualized - for each investor type, while Panel B shows these for each exit type. Also reported are t-statistics of the difference of mean returns from that of VC firm investor, the most common investor type (in Panel A), and t-statistics of the difference of mean returns from that of IPO exit, the most sought-after exit (in Panel B), and their significances. All variables are defined in the Appendix.

Panel A

Investor Type	N	Mean Log Return	*t*-stat of difference of mean with that of Venture Capital Firm investor	Mean Annualized Log Return	*t*-stat of difference of mean with that of Venture Capital Firm investor
Venture Capital Firm	65,196	0.62	-	0.25	-
Angel Investor	1,445	0.41	-7.67***	0.09	-16.46***
Corporation	7,115	0.23	-29.45***	0.13	-18.24***
Investment Bank	1,468	0.28	-9.74***	0.15	-6.16***
Leasing Company	254	0.38	-3.01***	0.17	-1.6
Limited Partner	429	0.54	-1.36	0.24	-0.39
Mezzanine/Buyout Fund	2,861	0.34	-13.93***	0.12	-14.18***
Public Sector Organization	105	0.52	-1.12	0.11	-5.50***
SBIC	1034	0.48	-3.50***	0.17	-4.83***
University	317	0.93	3.98***	0.33	2.53**
Other	1,925	0.48	-4.76***	0.21	-2.42**
Unidentified	669	0.44	-4.44***	0.25	0.26

Panel B

Exit Type	N	Mean Log Return	*t*-stat of difference of mean with IPO exit	Mean Annualized Log Return	*t*-stat of difference of mean with IPO exit
Failure	26,135	0.08	-203.45***	0.008	-163.25***
Growth	7,353	0.29	-126.41***	0.13	-105.67***
IPO	21,480	1.72	-	0.75	-
M&A	23,862	0.16	-119.76***	0.05	-111.99***
Other	3,988	0.48	-42.24***	0.21	-36.99***

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

We first perform the test using cross-region flag with all data. The results are shown in Table 5A. We find that the difference in returns, both log returns and annualized log returns, of out-of-state (cross region)

investments and in-state investments are negative and statistically significant. In addition, the median of returns for out-of-state investments are lower than that of in-state investments. *Mood's tests* yield *p-value*s close to 0, signaling that the distribution of expected returns for in-state investment are more positively skewed than that of out-of-state (cross region) investments. Thus, we conclude that, using all available data, in-state investments yield superior returns than out-of-state investments. Using the short and long "distance" variable, the results are similar to the results using the cross-region and in-state indicator variables.

In Table 5B we examine different exit types - IPO, M&A, and Growth. We ignore "Failure" exits in this section since the returns for failed investments are generally very low and very close to each other. Returns of in-state investments are superior to that of out-of-state investments. There are instances where median returns for in-state investments equal that of out-of-state investments but the *Mood's test* yields very significant p-value of close to 0, and this test is meant to indicate whether one distribution is more positively or negatively skewed than the other (Mood, 1950). Since the mean in-state investments return is higher than mean out-of-state return for all exit types, we argue that the distribution of in-state investments are more positively skewed than that of out-of-state investments.

Examining different investor types (Table 5C), the effect is only truly consistent for venture capital firms and fairly consistent for angel investors. For other investors, the difference in mean returns and in distribution of returns are insignificant at the 5% level. It is possible that we find strong results for venture capital firms because they make up the vast majority of our data, and our samples of investments by the other investor classes are relatively small. For investment entry round sub-samples (Table 5D), the effect is consistent for early-stage investors. For late-stage investors, the proximity effect is not as strong, again perhaps because of the small sample. Finally, the proximity effect is also consistent when we segregate the data by investments from favored VC locations (the well-documented hubs) or not, (Table 5E); as well as to favored industries or not (Table 5F).

The results are also similar using the short and long "distance" variable, to the results using the cross-region and in-state indicator variables. The consistency of results across all panels, showing cumulative returns and annualized returns to be significantly higher for shorter distance investments than for longer distance ones, attests to the robustness of our findings.

TABLE 5
IN-STATE VERSUS CROSS-REGION INVESTMENTS, SHORT DISTANCE VERSUS LONG DISTANCE INVESTMENTS: UNIVARIATE RESULTS

This table shows difference in mean returns (and the associated t-statistics), and difference in median returns by whether venture investments are in-state or cross-region. Cross-region is defined as equal to 1 if the state of the investor's main office location is different than that of the portfolio company. Distance below and above whole sample median distance qualifies as short and long distance. Panel A shows tests with all data, Panel B shows by exit types, Panel C shows by investor types, Panel D shows by investment entry type, Panel E by from favored VC firm location or not, and Panel F by to VC favored industries or not.

Panel A

	Cross-region	In-State	Short Distance	Long Distance
Mean (Log Return)	0.50	0.64***	0.62	0.51 ***
Median (Log Return)	0.07	0.11 †††	0.10	0.07 †††
Mean (Annualized Log Return)	0.21	0.25***	0.24	0.22 ***
Median (Annualized Log Return)	0.01	0.12 †††	0.01	0.007 †††

Panel B

		Cross-region	In-State	Short Distance	Long Distance
IPO	Mean (Log Return)	1.59	1.90***	1.84	1.61 ***
	Median (Log Return)	1.42	1.69 †††	1.63	1.43 †††
	Mean (Annualized Log Return)	0.73	0.77***	0.75	0.74
	Median (Annualized Log Return)	0.56	0.58 †††	0.57	0.57
M&A	Mean (Log Return)	0.11	0.25***	0.25	0.09 ***
	Median (Log Return)	0	0 †	0	0 †
	Mean (Annualized Log Return)	0.02	0.10***	0.09	0.02 ***
	Median (Annualized Log Return)	0	0 †	0	0 †
Growth	Mean (Log Return)	0.27	0.31**	0.31	0.25 ***
	Median (Log Return)	0	0 †	0	0
	Mean (Annualized Log Return)	0.12	0.14***	0.14	0.11 ***
	Median (Annualized Log Return)	0	0 †	0	0

Panel C

		Cross-region	In-State	Short Distance	Long Distance
VC firm	Mean (Log Return)	0.56	0.69***	0.67	0.57 ***
	Median (Log Return)	0.10	0.14 †††	0.13	0.10 †††
	Mean (Annualized Log Return)	0.23	0.27***	0.26	0.24 ***
	Median (Annualized Log Return)	0.011	0.017 †††	0.015	0.01 †††
Angel	Mean (Log Return)	0.34	0.44*	0.42	0.38
	Median (Log Return)	0.03	0.12 †††	0.13	0.01 †††
	Mean (Annualized Log Return)	0.07	0.10	0.09	0.07
	Median (Annualized Log Return)	0.003	0.014 †††	0.01	0.001 †††
Corporation	Mean (Log Return)	0.23	0.22	0.23	0.22
	Median (Log Return)	0	0	0	0
	Mean (Annualized Log Return)	0.13	0.12	0.12	0.13
	Median (Annualized Log Return)	0	0	0	0
Investment Bank	Mean (Log_Return)	0.26	0.40*	0.28	0.28
	Median (Log Return)	0.04	0.05	0.05	0.04
	Mean (Annualized Log Return)	0.14	0.20	0.15	0.14

Mezzanine Fund	Median (Annualized Log Return)	0.005	0.006	0.005	0.004
	Mean (Log_Return)	0.326	0.42	0.38	0.32
	Median (Log Return)	0	0.01	0	0
	Mean (Annualized Log Return)	0.12	0.14	0.12	0.12
	Median (Annualized Log Return)	0	0.001	0	0
SBIC	Mean (Log_Return)	0.47	0.49	0.48	0.48
	Median (Log Return)	0.10	0.09	0.098	0.11
	Mean (Annualized Log Return)	0.18	0.16	0.16	0.18
	Median (Annualized Log Return)	0.011	0.01	0.01	0.01

Panel D

		Cross-region	In-State	Short Distance	Long Distance
Early Entry	Mean (Log_Return)	0.51	0.65***	0.62	0.51***
	Median (Log Return)	0.07	0.11 †††	0.10	0.07 †††
	Mean (Annualized Log Return)	0.21	0.25***	0.24	0.22 ***
	Median (Annualized Log Return)	0.007	0.012 †††	0.01	0.01 †
Late Entry	Mean (Log_Return)	0.33	0.41	0.47	0.28 ***
	Median (Log Return)	0.008	0.06 †	0.12	0 †††
	Mean (Annualized Log Return)	0.115	0.17	0.22	0.07 ***
	Median (Annualized Log Return)	0.0008	0.02 †	0.05	0.02 †††

Panel E

		Cross-region	In-State	Short Distance	Long Distance
From Favored VC Location	Mean (Log_Return)	0.53	0.72***	0.70	0.54 ***
	Median (Log Return)	0.086	0.15 †††	0.14	0.08 †††
	Mean (Annualized Log Return)	0.24	0.28***	0.28	0.24 ***
	Median (Annualized Log Return)	0.009	0.02 †††	0.017	0.01 †††
From Outside Favored VC Location	Mean (Log_Return)	0.49	0.58***	0.55	0.49 ***
	Median (Log Return)	0.07	0.10 †††	0.09	0.06 †††
	Mean (Annualized Log Return)	0.19	0.22***	0.20	0.20
	Median (Annualized Log Return)	0.007	0.01 †††	0.01	0.006 †††

Panel F

		Cross-region	In-State	Short Distance	Long Distance
To VC favored industries	Mean (Log Return)	0.54	0.69***	0.65	0.55 ***
	Median (Log Return)	0.10	0.16 †††	0.14	0.10 †††
	Mean (Annualized Log Return)	0.24	0.28***	0.27	0.25 **
	Median (Annualized Log Return)	0.01	0.02 †††	0.02	0.01 †††
To Outside VC favored industries	Mean (Log Return)	0.46	0.60***	0.58	0.46 ***
	Median (Log Return)	0.05	0.09 †††	0.08	0.05 †††
	Mean (Annualized Log Return)	0.17	0.21***	0.20	0.18 **
	Median (Annualized Log Return)	0.005	0.01 †††	0.009	0.005 †††

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level in difference of means *t*-test.
† denotes significant at 10% level, †† significant at 5% level, and ††† significant at 1% level in difference of median Mood's test.

Multivariate Analysis

We first test the Manigart et al. (2002) claim that early investors may get higher returns from startups than late round investors. We regress log returns against financing round (*Round Entry*) of entry of investors in the portfolio companies, and other control variables, using the following specification:

$$\begin{aligned}\text{Returns}_t = \boldsymbol{\beta T} + \boldsymbol{\beta I} + \beta 1 \times \text{Round Entry} + \beta 2 \times \text{US Annual GDP Growth}_{t-1} + \\ \beta 3 \times \text{US Annual Industrial Production Growth}_{t-1} + \\ \beta 4 \times \text{US Annual Consumer Sentiment Growth}_{t-1} + \varepsilon,\end{aligned} \quad (1)$$

βT is the vector of time fixed effects, and
βI is the vector of industry fixed effects, as defined in the Appendix.

We use three control variables - US Annual GDP Growth, and US Annual Industrial Production Growth, which are proxies for booming economic conditions that can influence returns at the time of investment exit, and US Annual Consumer Sentiment Index Growth that can also influence investor sentiment and, thereby, returns at the time of exit. We use cumulative log returns as the dependent variable, but in a robustness check, also test with annualized log returns.

We run the above specification over all data as well as separately over 2 exit groups - definite successful exit (IPO & M&A) and definite unsuccessful exit (Failure). Regression coefficients and their significance are reported in Table 6.

TABLE 6
INVESTMENT ENTRY TIMING AND RETURNS

This table shows regression coefficients and the associated t-statistics (in parenthesis) of log return on round entry, as the main explanatory variable, and control variables as well as industry and year fixed effects. Definite Successful exits for investments are defined as those portfolio firms that successfully has an IPO or were acquired. Definite unsuccessful exits are firm failures. Firm failures are defined as startups whose last financing round is a growth round that dates back more than 10 years from current date. All variables are defined in Appendix.

	Log Returns as Dependent Variable		
	All Exits	Definite Successful Exits (IPO, M&A)	Definite Unsuccessful Exit (Failure)
N	82,818	45,342	26,135
Round Entry	-0.10*** (-45.36)	-0.17*** (-49.08)	-0.02*** (-44.34)
US Annual GDP Growth	-0.28 (-0.71)	-0.29 (-0.66)	-0.06*** (-3.185)
US Annual Industrial Production Growth	-0.07*** (-2.84)	-0.23*** (-5.27)	0.01*** (2.81)
US Annual Consumer Sentiment Growth	-0.08 (-1.27)	-0.16** (-2.03)	0.003*** (2.71)
Industry Segment Fixed Effect	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes
Adjusted R-square	0.114	0.185	0.151

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level

In all regressions, *Round Entry* variable is negative and statistically significant at the 1% level, signaling a strong association. These results are also consistent with annualized log returns (untabulated). Thus, early stage investors realize a higher returns as compared to later round investor. We examine 2 types of return calculation methods - cumulative log return and annualized log return - and two proxies for geographical proximity - cross-region dummy and distance (in km), and other control variables that could influence venture capital returns, as discussed before, in regression specifications below:

$$\begin{aligned} Y_t = {} & \boldsymbol{\beta T} + \boldsymbol{\beta I} + \beta 1 \times \text{cross region} + \beta 2 \times \text{Round Entry} + \\ & \beta 3 \times \text{US Annual GDP Growth}_{t-1} + \beta 4 \times \text{US Annual Industrial Production Growth}_{t-1} \\ & + \beta 5 \times \text{US Annual Consumer Sentiment Growth}_{t-1} + \varepsilon, \end{aligned} \quad (2)$$

$$Y_t = \boldsymbol{\beta T} + \boldsymbol{\beta I} + \beta 1 \times \text{distance} + \beta 2 \times \text{Round Entry} + \beta 3 \times \text{US Annual GDP Growth}_{t-1} + \beta 4 \times \text{US Annual Industrial Production Growth}_{t-1} + \beta 5 \times \text{US Annual Consumer Sentiment Growth}_{t-1} + \varepsilon,$$

where Y_t is either cumulative log returns or annualized log returns.

We first use all data as well as for only definite successful exits (IPO and M&A). The results are shown in Tables 7A and 7B. For tests using cumulative log returns, both cross-region dummy and distance variable are negative and statistically significant at the 1% level. For tests using definite successful exits only, the absolute value of coefficients and *t-stats* for both the cross-region variable and distance increase, signaling even stronger effect than presented in the test with all data. For tests using annualized log returns, the cross-region dummy variable is significant when using all data and using data on definite exits only; the distance variable is only significant when for definite exits. Thus, if successful exits are the desired VC outcomes, returns reduce as VC investments become more distant.

There are 5256 observations in our sample, where the investors are from outside the US. In Table 7B, we examine the associations of investment proximity with both type of returns for domestic-only data (both investors and portfolio companies are located inside US), and find the results consistent: both cross-region indicator variable, and distance variable are significantly and negatively associated with both types of returns.

In table 7C, we examine the associations between investment proximity and returns by investment stage. For early round investor, containing nearly 99% of all data, the effect is, as expected, consistent with what is found in regression using all data seen in Table 7A and 7B. Next, we examine only Venture Capital firm investors, and the next most common investor class in our data, Corporations. The results are reported in Table 7D. For Venture Capital firm investors, we find that the associations of both cross-region variable and distance variable to be consistent with what we found in the regression using all data – proximity is significantly and positively associated with returns. However, for corporate investments, we find no significant effects. This might suggest the effect of geographical proximity on returns may be present more strongly only for venture capital firms because of the better monitoring such VC firms can provide.

Next, in Table 7E, we divide our data into those involving investors from favored VC locations – the 3 VC hubs (see Chen et al. (2010)) plus Palo Alto and Menlo Park - and from other VC locations. Both cross-region variable and distance variable are negative and statistically significant, confirming our initial hypothesis. This doesn't necessarily contradict Chen et al. (2010) finding since his group study only states that the VC's outsized performance, not all investment performance, come from outside the VC's location. For tests using annualized returns (untabulated), cross-region variable is negative and significant regardless of location while distance is insignificant in both tests.

Finally, in Table 7F, we examine investments to only favored VC industries and only to non-favored VC industries (see Hochberg, Ljungqvist, and Lu (2007)). For tests using cumulative log return, both cross-region variable and distance variable are negative and statistically significant. For tests using annualized log returns (untabulated), cross-region variable is significant while distance variable is not.

Overall, using two types of venture investment returns, cumulative and annualized, and all data as well as various data sub-samples, we find that the significant positive association of geographical proximity with investment returns is robust using the cross-region (out-of-state dummy) variable. The *distance* variable is also statistically significant, though not in all sub-samples like the *cross-region* variable, which is consistent with what we found in univariate tests.

TABLE 7
INVESTMENT PROXIMITY AND RETURNS: MULTIVARIATE ANALYSIS

This table shows regression coefficients and t statistics in parenthesis, of 2 types of returns – cumulative and annualized - regressed on several explanatory variables, of which *Cross_region* (whether an investment is out of state) and *Distance_km* (the distance in kilometers between the investor HQ and portfolio company HQ) are the key ones, as well as time and industry fixed effects. Panels A and B show regressions run over the full sample, and in the other panels, regressions run over various subsamples of data, as robustness checks. All variables are defined in Appendix.

Panel A: All data with both type of returns

Log Returns as Dependent Variable				
	All Exits		Definite Successful Exits (IPO and M&A)	
N	82,818	82,818	45,342	45,342
Cross_region	-0.131*** (-15.02)		-0.252*** (-17.79)	
Distance_km		-0.000017*** (-6.57)		-0.000042*** (-9.88)
Round Entry	-0.098*** (-44.63)	-0.099*** (-44.72)	-0.167*** (-48.72)	-0.166*** (-48.45)
US Annual GDP Growth	-0.255 (-0.67)	-0.277 (-0.73)	-0.274 (-0.61)	-0.305 (-0.67)
US Annual Industrial Production Growth	-0.069*** (-2.84)	-0.068*** (-2.8)	-0.222*** (-5.06)	-0.229*** (-5.20)
US Annual Consumer Sentiment Growth	-0.080 (-1.217)	-0.085 (-1.28)	-0.152* (-1.91)	-0.163** (-2.04)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.116	0.114	0.19	0.187

Annualized Log Return as Dependent Variable				
	All Exits		Definite Successful Exits (IPO and M&A)	
N	82,818	82,818	45,342	45,342
Cross_region	-0.035*** (-8.84)		-0.058*** (-9.21)	
Distance_km		-6.15e-07 (-0.51)		-5.09e-06*** (-2.69)
Round Entry	-0.011*** (-11.08)	-0.0115*** (-11.45)	-0.0263*** (-17.19)	-0.0264*** (-17.26)

US Annual GDP Growth	-0.059 (-0.34)	-0.065 (-0.37)	-0.055 (-0.27)	-0.062 (-0.31)
US Annual Industrial Production Growth	-0.017 (-1.57)	-0.017 (-1.57)	-0.047** (-2.39)	-0.049** (-2.48)
US Annual Consumer Sentiment Growth	-0.028 (-0.94)	-0.029 (-0.99)	-0.035 (-0.97)	-0.037 (-1.04)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.127	0.126	0.243	0.242

Panel B: Both type of returns on domestic-only data (both investors and portfolio companies are located inside US)

	Log Returns		**Annualized Log Returns**	
N	77,562	77,562	77,562	77,562
Cross_region	-0.126*** (-14.10)		-0.032*** (-7.96)	
Distance_km		-0.000017*** (-6.44)		-7.22e-07*** (-0.58)
Round Entry	-0.10*** (-43.73)	-0.101*** (-43. 83)	-0.011*** (-10.72)	-0.015*** (-11.03)
US Annual GDP Growth	-0.26*** (-0.69)	-0.287*** (-0.75)	-0.067 (-0.38)	-0.072 (-0.41)
US Annual Industrial Production Growth	-0.065** (-2.58)	-0.065** (-2.53)	-0.011 (-1.02)	-0.012 (-1.02)
US Annual Consumer Sentiment Growth	-0.079*** (-1.19)	-0.083*** (-1.26)	-0.027 (-0.88)	-0.028 (-0.92)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.116	0.114	0.127	0.126

Panel C: Log returns by Entry Round

Log Return as Dependent Variable				
	Early round (Before round 9)		Late Round (After round 9)	
N	81,809	81,809	1009	1009
Cross_region	-0.131*** (-14.90)		-0.044 (-1.02)	
Distance_km		-0.0000171*** (-6.35)		0.00000287 (0.24)

Round Entry	-0.113*** (-45.92)	-0.114*** (-45.98)	-0.028** (-2.38)	-0.0283** (-2.37)
US Annual GDP Growth	-0.251 (-0.66)	-0.27 (-0.71)	0.445 (1.32)	0.44 (1.30)
US Annual Industrial Production Growth	-0.077*** (-3.10)	-0.077*** (-3.06)	0.001 (0.03)	0.00005 (0.001)
US Annual Consumer Sentiment Growth	-0.088 (-1.33)	-0.093 (-1.39)	0.156*** (2.85)	0.154* (2.81)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.117	0.116	0.43	0.43

Panel D: Log returns by Major Investor Type

Log Return as Dependent Variable				
	Venture Capital		Corporation	
N	65,196	65,196	7115	7115
Cross_region	-0.123*** (-12.29)		0.011 (0.43)	
Distance_km		-0.000017*** (-5.47)		5.7e-06 (0.77)
Round Entry	-0.108*** (-41.98)	-0.108*** (-42.00)	-0.053*** (-9.20)	-0.053*** (-9.23)
US Annual GDP Growth	-0.152 (-1.12)	-0.148 (-1.09)	-0.202 (-0.64)	-0.20 (-0.63)
US Annual Industrial Production Growth	-0.022 (-0.36)	-0.024 (-0.39)	0.006 (0.09)	0.006 (0.09)
US Annual Consumer Sentiment Growth	-2.67 (-1.13)	-2.61 (-1.11)	-0.034 (-0.595)	-0.034 (-0.59)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.12	0.119	0.108	0.108

Panel E: Log returns by Favored VC Location (the VC centers of SF, NY, Boston, Palo Alto, and Menlo Park)

Log Return as Dependent Variable				
	Favored VC Location		Outside Favored VC Location	
N	31,325	31,325	51,493	51,493
Cross_region	-0.153*** (-10.31)		-0.088*** (-7.98)	

Distance_km		-0.000026*** (-6.02)		-7.34e-07** (-2.14)
Round Entry	-0.112*** (-29.55)	-0.113*** (-29.40)	-0.092*** (-34.27)	-0.093*** (-34.56)
US Annual GDP Growth	-0.251 (-0.63)	-0.283 (-0.71)	-12.84 (-0.99)	-0.124 (-0.96)
US Annual Industrial Production Growth	-0.116*** (-3.07)	-0.116*** (-3.05)	0.02 (0.47)	0.0268 (0.44)
US Annual Consumer Sentiment Growth	-0.098 (-1.43)	-0.105 (-1.50)	-2.24 (-1.01)	-2.15 (-0.97)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.136	0.129	0.108	0.107

Panel F: Log returns by Favored VC Industries ("Communications and Networking", "Biopharmaceuticals", "Media and Content", "Software", "Healthcare Services")

Log Return as Dependent Variable				
	Favored VC Industries		**Not Favored VC Industries**	
N	42,318	42,318	40,500	40,500
Cross_region	-0.138*** (-11.17)		-0.119*** (-9.71)	
Distance_km		-0.000019*** (-5.10)		-0.000016*** (-4.42)
Round Entry	-0.11*** (-35.52)	-0.112*** (-35.68)	-0.087*** (-27.95)	-0.087*** (-27.93)
US Annual GDP Growth	1.09*** (2.72)	1.107*** (2.75)	-0.325 (-0.87)	-0.345 (-0.92)
US Annual Industrial Production Growth	-0.095** (-2.55)	-0.093** (-2.47)	-0.004 (-0.13)	-0.00345 (-0.11)
US Annual Consumer Sentiment Growth	0.192*** (2.66)	0.196*** (2.72)	-0.091 (-1.39)	-0.095 (-1.46)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.152	0.15	0.089	0.088

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

Economic Significance

From Table 7A, our base model, transforming the coefficients to percentages, VC investors' out-of-state investments' cumulative returns is around 12.3% lower, and annualized returns 3.4% lower for out-of-state investments than for in-state investments. Likewise, for every 100 km increase in the distance between the VC and the portfolio companies, cumulative returns of the investments drop by 0.169%. For only successful exits – defined as M&A and IPO, the effects are more pronounced. Cumulative returns of these eventually-successful investments drop by 22%, while annualized returns drop 5.6% for out-of-state investments, as compared to in-state investments. Using the distance_km variable, a 100 km increase lowers cumulative returns by 0.42%, and annualized returns by 0.05%.

For IPO-exits only, cumulative returns drop by 20.47% for out-of-state investments. For M&A-exits only, cumulative returns for out-of-state investments drop by 14.96%. From Table 7D, for the models using VC firm-only data, cumulative returns drop by 11.57%, if investors change from in-state investments to out-of-state. For investors inside VC Hub, (Table 7E), moving investments to out-of-state results in a 14.19% drop in cumulative returns. Finally, for data segregated by favored VC industries (see Appendix), the result (Table 7F) is similar to our base model. For investments within favored VC industries, moving out-of-state results in an expected drop of 12.89% in cumulative investment returns.

Robustness Checks

Selection Bias Control

We employed a multivariate analysis that included control variable and fixed effects to explain returns. Nevertheless, it is still possible that returns appears to be associated with investment proximity simply because proximate investments occur more frequently in types of investments where the observed outcomes are more likely. To control for this form of selection bias, we employ an instrumental variable (IV) simultaneous equations regression model using limited information maximum likelihood (LIML) estimation (see Juergens and Lindsey (2009), Krishnan, Ivanov, Masulis, and Singh (2011), and Krishnan and Masulis (2011)), where *Cross_region* (whether an investment is out of state) or *Distance_km* (the distance in kilometers between the investor HQ and portfolio company HQ) is the alternative endogenous covariate. An IV should have the properties that while it strongly predicts the investment, it is unrelated to main dependent variable (returns) being examined.

The IVs we use are (i) *Hot Investment year*, an indicator variable that takes the value of 1 for those years when the frequency of venture investments in portfolio companies have been more than the whole sample median, and 0 otherwise, (ii) *Investor Capital under Management,* the amount of capital managed by the venture capital investor at the time of the investment taken from the *VentureXpert* database, (iii) *Investor Industry Preference Match,* an indicator variable that takes the value of 1 if the investor's preferred industry SIC code, taken from the *VentureXpert* database, matches the portfolio firm industry code, at the time of investment for the venture capital investor, and 0 otherwise, and (iv) *Investor Stage Preference Match*, that takes the value of 1, when whether the investor's preference for early or late stage investments, taken from the *VentureXpert* database, matches the stage of the current investment in a portfolio firm, and 0 otherwise.

Economically, the choice of these IVs is justified because investment conditions, the size and investment preferences of the venture capital investor, as at the time of investment, will strongly predict the investment. However, there is no compelling reason to expect these features at the time of investment to be significantly associated with returns in the long run, in the presence of control variables and fixed effects, other than through the investment itself. Statistically, we examine the validity of the instruments by performing overidentification tests. The *F*-statistics for the joint significance of the four IVs is 12.7 for *Cross_region* and 11.5 for *Distance_km*, which are above the critical value of 10 recommended by Staiger and Stock (1997). The *Anderson-Rubin* statistic for overidentification results in insignificant *p*-values for *cumulative log return*, after controlling for economic variables and fixed effects. Thus, we fail to reject the joint null that the IVs are uncorrelated with the error term, and hence statistically satisfy the exclusion requirement of valid instruments.

Table 8 reports the regression coefficients, and the associated *t* statistics in parenthesis for an IV model using the limited information maximum likelihood (LIML) estimation approach, where *Cross_region* (whether an investment is out of state) or *Distance_km* (the distance in kilometers between the investor HQ and portfolio company HQ) is the alternative endogenous covariate. We find that our main results are robust. Cross region indicator continues to be significantly negatively associated with returns, and distance of investment is also significantly negatively associated with cumulative returns.

TABLE 8
INVESTMENT PROXIMITY AND RETURNS: SELECTION BIAS CONTROL

This table reports the regression coefficients, and the associated *t* statistics in parenthesis for an IV model using the limited information maximum likelihood (LIML) estimation approach, where *Cross_region* (whether an investment is out of state) or *Distance_km* (the distance in kilometers between the investor HQ and portfolio company HQ) is the alternative endogenous covariate instrumented with *Hot Investment Year*, *Investor Capital under Management, Investor Industry Preference,* and *Investor Stage Preference* and regressed on several explanatory variables, including time and industry fixed effects. All variables are defined in Appendix.

	Log Return as Dependent Variable	
Cross_region	-0.10*** (-9.65)	
Distance_km		-0.000015*** (-5.81)
Round Entry	-0.09*** (-40.33)	-0.09*** (-39.99)
US Annual GDP Growth	-0.26 (-0.65)	-0.27 (-0.69)
US Annual Industrial Production Growth	-0.06*** (-2.75)	-0.06*** (-2.76)
US Annual Consumer Sentiment Growth	-0.04 (-0.78)	-0.06 (-0.82)
Industry Segment Fixed Effect	Yes	Yes
Exit Year Fixed Effect	Yes	Yes
Adjusted R-squared	0.109	0.111

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

Matched Sample Analysis

Next, we examine whether proximity of investments continue to be significantly and positively associated with returns using matched-sample analysis. Investment pairs – one cross-region and the other in state, or alternatively, one long distance and the other short distance - are matched by (i) investor name, (ii) industry segment of the portfolio company, and (iii) by minimum difference in investment year (i.e., we determine the matched pair with lowest investment year difference). Table 9A shows results for difference in mean cumulative or annualized returns by investment proximity determined by cross region or in-state investment. Table 9B shows results for difference in mean cumulative or annualized returns by investment proximity determined by long distance or short distance investment. Both panels shows, that over a large matched sample, cumulative log returns as well as annualized log returns are significantly higher for the more proximate investments. Thus, the relationship between proximity of venture capital investments and returns looks robust, significant, and positive.

TABLE 9
INVESTMENT PROXIMITY AND RETURNS:
ROBUSTNESS CHECK WITH MATCHED SAMPLE ANALYSIS

Investment pairs – one cross-region and the other in state, or alternatively, one long distance and the other short distance - are matched by (i) venture capital investor name, (ii) industry segment of the portfolio company, and (iii) by minimum difference in investment year (i.e., determine the matched pair with lowest investment time difference). Panel A shows results for difference in mean cumulative or annualized returns by investment proximity determined by cross region or in-state investment. Panel B shows results for difference in mean cumulative or annualized returns by investment proximity determined by long distance or short distance investment.

Panel A

	In State	Cross Region	Difference of Means t-stat
Mean (Log Return)	0.949	0.768	9.86***
Mean (Annualized Log Return)	0.364	0.308	6.80***
N	12,414	12,414	

Panel B

	Short Distance	Long Distance	Difference of Means t-stat
Mean (Log Return)	0.849	0.697	9.25***
Mean (Annualized Log Return)	0.325	0.285	5.39***
N	15,245	15,245	

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

REASONS FOR RETURNS

To understand the positive association between geographical proximity and returns, we examine some portfolio company features that might influence stock performance. Lerner (1995) and Bernstein et al (2016), for example, have argued that geographical proximity may lead to better monitoring of portfolio companies especially by Venture Capital firms. The two ways to do this, are by continued investment by VCs, and through board independence (and better monitoring) via VC appointed directors and other outside independent directors. For this analysis, for portfolio companies with multiple venture capital investors, companies with more proximate (in-state) VC investors than distant (out-of-state) ones, are classified as relatively "In-state". We use a similar methodology for classifying the relatively "short-distance" companies, based on the long/short distance dummy variable. Table 10 shows that the average length of VC investments is significantly shorter for relatively "in-state" or, alternatively the relatively "short-distance" companies, presumably because the venture capital investors are able to exit, on average, sooner. However, VC Director and other outside directors as a proportion of all Directors on the boards on portfolio companies is significantly higher, on average, for in-state and short distance investments indicating better involvement (and likely better monitoring) by VCs in the more proximate investments.

To check the association of the length of VC investment in a portfolio company and the percentage of VC appointed and other outside directors in a portfolio company on returns, we start with our regression model and add these two portfolio company variables as explanatory variables. Table 11 shows that percentage of VC appointed and other outside directors in a portfolio company is significantly and

positively associated with both cumulative and annualized returns, while the relationship of the length of VC investment on returns is more complicated – while it is positively and significantly associated with cumulative returns, it is significantly and negatively associated with annualized returns, i.e. cumulative log returns divided by the number of years of investment, perhaps implying diminishing returns over time. Further, as we noted before, in the more proximate (higher return) investment, venture capital firms exit quicker: so investing for longer does not lead to higher annualized returns.

TABLE 10
RETURNS AND PORTFOLIO COMPANY FEATURES: UNIVARIATE ANALYSIS

This table reports mean statistics of the length of VC investments and other portfolio company board features by in-state and out-of-state investors, and by short and long distance investments. If a portfolio company has more in-state investors than out-of-state ones, then it would be classified as "in- state", otherwise Cross Region. Similarly, if a portfolio company has more short distance investors than long distance ones, then it would be classified as "Short Distance", otherwise "Long Distance".

	Portfolio Company Features	
	Length of VC investment	% VC & Other Outsider Directors on Board
In State	1.95	0.31
Cross Region	2.11	0.30
t-stat of difference of means	-4.50***	3.38***
Short Distance	1.98	0.31
Long Distance	2.12	0.30
t-stat of difference of means	-4.20***	3.50***
N	15,983	15,983

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

TABLE 11
RETURNS AND PORTFOLIO COMPANY FEATURES: MULTIVARIATE ANALYSIS

This table shows regression coefficients and t statistics in parenthesis, of 2 types of returns – cumulative and annualized - regressed on several explanatory variables, of which *Cross_region* and *Distance_km* are the key ones, as well as time and industry fixed effects, run over the full sample. The new explanatory variables added are *% VC & Other Outsider Directors on Board*, and *Length of VC investment*. All variables are defined in Appendix.

	Log Returns as Dependent Variable			
	Log Return		Annualized Log Return	
Cross_region	-0.137*** (-16.09)		-0.034*** (-8.73)	
Distance_km		-0.00002*** (-7.72)		-5.23e-07 (-0.433)
Round Entry	-0.077*** (-34.79)	-0.077*** (-34.88)	-0.014*** (-13.83)	-0.0146*** (-14.19)
US Annual GDP Growth	-0.125 (-0.34)	-0.148 (-0.40)	-0.067 (-0.38)	-0.072 (-0.42)

US Annual Industrial Production Growth	-0.073*** (-3.09)	-0.073*** (-3.05)	-0.0164 (-1.50)	-0.0165 (-1.50)
US Annual Consumer Sentiment Growth	-0.066 (-1.02)	-0.07 (-1.08)	-0.028 (-0.95)	-0.029 (-0.98)
% VC & Other Outsider Directors on Board	0.191*** (6.16)	0.186*** (6.52)	0.039*** (2.87)	0.038*** (2.85)
Length of VC investment	0.099*** (67.20)	0.099*** (67.09)	-0.01*** (-14.86)	-0.01*** (-14.92)
Industry Segment Fixed Effect	Yes	Yes	Yes	Yes
Exit Year Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-squared	0.163	0.161	0.129	0.128

* denotes significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

CONCLUSION

Using a very large dataset of 82,818 investments in startup companies, spanning a 54-year period from 1965 through 2019, we document that venture capital firms are biggest group of investors by far, (about 79% of all observations), among all early state investors like corporations, angel investors, university endowment funds etc. Late-stage investments are classified as those happening after 9th round shows a slightly increasing trend in recent years. The number of venture investments peaked in the 1990's. Cross-region (across state/country) investment proportion shows a declining trend, but remain more than half of all investments. The majority of in-state investments are found clustered in California (San Francisco Hub), Massachusetts (Boston Hub) and New York. Out-of-state investments are more widely spread-out over the map, suggesting that regardless of the clusters, venture investors do make investments over a wider region in the United States. The trend of returns generated from early round investing to late round decreases generally: early investors get a higher return than late stage investors, and IPO exits entail the highest return for the investor both in terms of log returns and annualized of log returns, justifying the conclusion that IPO would be the most sought after exit. University fund investments yield the highest return both in terms of log returns and annualized log returns, while venture capital firm investments (which, as discussed above, comprise of the dominant portion of all investments in our dataset), the second highest return.

We examine the effect of geographical proximity on VC investment returns using two different proxy for geographical proximity - cross-region indicator variable and the actual distance between the portfolio company HQ and investor HQ. We examine the effect using two different types of returns, cumulative log return and annualized log return, and perform univariate and multivariate tests. We analyze various subsections of our large dataset by exit types – all, successful and unsuccessful, by investor-firm types, by entry timing – early-stage or late-stage investments, whether investment was made from favored VC locations (VC hubs), whether investment is to favored VC industries. We find that the cross-region variable is a significant predictor of returns in our full dataset as well as across almost all data subsets. The actual distance variable, while not as strong as the cross-region variable, is also robust in that it is significant for cumulative returns.

The relation between investment proximity and returns continues to be robust even when we control for selection bias (using Instrumental Variables and Limited Information Maximum Likelihood estimation procedure), and in a matched-sample analysis where we pair each proximate investment with a matched distant investment. Given our robust results, we conclude that proximity of investments beneficially affect the returns for VCs, perhaps because they can better monitor such investments (Bernstein et al, 2016).

Indeed, we find that VC Director and other outside directors as a proportion of all Directors on the boards on portfolio companies is significantly higher, on average, for in-state and short distance investments as compared to out-of-state or longer distance investments, indicating better involvement by VCs in their more proximate investments. The percentage of VC-appointed and other outside directors in a portfolio company is also significantly and positively associated with both cumulative and annualized returns, indicating better governance, and perhaps monitoring by VCs in their more proximate investments.

REFERENCES

Barry, C., Muscarella, C., Peavy, J., & Vetsuypens, M. (1990). The role of venture capital in the creation of public companies: Evidence from the going-public process. *Journal of Financial Economics*, *27*, 447–472.

Bernstein, S., Xiroud, S., & Townsend, R. (2016). The Impact of Venture Capital Monitoring. *Journal of Finance*, *71*(4), 1591-1622.

Chemmanur, T., Hull, T., & Krishnan, K. (2016). Do Local and International Venture Capitalists Play Well Together? The Complementarity of Local and International Venture Capitalists. *Journal of Business Venturing*, *31*(5). 573-594.

Chen, H., Gompers, P., Kovner, A., & Lerner, J. (2010). Buy local? The geography of venture capital. *Journal of Urban Economics*, *67*, 90-102.

Coval, J., & Moskowitz, T.J. (2002). Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54(6), 2045-2073.

Cumming, D., & Dai, N. (2010). Local Bias in venture capital investments. *Journal of Empirical Finance*, *17*(3), 362-380.

Ellis, J.A., Madureira, L., & Underwood, S. (2020). The Causal Effects of Proximity on Investment: Evidence from Flight Introductions. *Journal of Financial and Quantitative Analysis*, *55*(6), 1978-2004.

Gompers, P.A. (1995). Optimal investment, monitoring, and the staging of venture capital. *Journal of Finance*, *50*(5). 1461–1489.

Gorman, M., & Sahlman, W. (1989). What do venture capitalists do? *Journal of Business Venturing*, pp. 231-248.

Hellmann, T., & Puri, M. (2002). Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence. *Journal of Finance*, *57*(1), 169-197.

Hochberg, Y., Ljungqvist, A., & Lu, Y. (2007). Whom you know Matters: Venture Capital networks and Investment performance. *Journal of Finance*, *62*(1), 251-301.

Juergens, J., & Lindsey, L. (2009). Getting Out Early: An Analysis of Market Making Activity at the Recommending Analyst's Firm. *Journal of Finance*, *64*, 2327-2359.

Krishnan, C.N.V., & Masulis, R. (2013). Law firm Expertise and Merger and Acquisition Outcomes, *Journal of Law and Economics*, *56*(1), 189-226.

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

Lerner, J. (1995). Venture capitalists and the oversight of private firms. *Journal of Finance 50*, 301-388.

Manigart, S., Waele, K., Wright, M., Robbie, K., Desbrières, P., Sapienza, H., & Beekman, A. (2002). Determinants of required return in venture capital investments: a five-country study. *Journal of Business Venturing*, *17*, 291-312.

Mood, A.M., & Brown, G.W. (1951). On median tests for linear hypotheses. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 159-166.

APPENDIX
DEFINITIONS OF VARIABLES

Return Variables	Description
Log Return or Cumulative Log Return	= ln (Valuation at Exit Round / Valuation at Entry Round)
Annualized Log Return	= Log Return / Investment Time in Years

Proximity Variable	Description
Cross_Region	An indicator Variable that takes the value of 1 when the portfolio firm main location/HQ is not in the same U.S. state/country as compared to VC investment firm main location/HQ, and 0 otherwise
Distance	The actual distance in kilometers between the portfolio firm main location/HQ and VC investment firm main location/HQ
Short Distance	An indicator value that takes the value of 1 when Distance as defined above is equal to or less than the full sample median value, and 0 otherwise.
Long Distance	An indicator value that takes the value of 1 when Distance as defined above is more than the full sample median value, and 0 otherwise.

Investor Type	Description
Venture Capital firm	Includes "Venture Capital", "Corporate Venture Capital" and "Diversified Private Equity" in the original dataset
Mezzanine/Buyout Fund	"Mezz/Buyout" in *VentureSource* database
Non-identified	Observations that doesn't have a specific investor classification in *VentureSource* database

Investing round	Description
Early round	Investment Round before round 9
Late round	Investment Round at or after round 9

Investor exits	Description
IPO	As defined in *VentureSource* database
M&A	M&A consists of "ACQ" (Acquisition), "AACQ" (Asset Acquisition), "MER" (Merger), "Add-on", "PE" (Private Equity Acquisition), "Carveout", "revMER" (Reverse merger), "Spin" (Spin-off), "MBO" ("Management Buyout", "LBO" (Leveraged Buyout), "MBI" (Management Buy-in) from the *VentureSource* database.
Growth	As defined in *VentureSource* database
Failure	If the last financing round of a startup is a growth round that dates back more than 10 years (2009 in our case) then it is defined as "Failure"

VC Hub	Description
Favored VC location	VC Hub are San Francisco, New York City, Boston, Menlo Park and Palo Alto
Outside Favored VC location	Observations whose investors' locations are outside of the cities listed above

Industry	Description
VC Favored Industry	Startups whose industries are in "Communications and Networking", "Biopharmaceuticals", "Media and Content", "Software" and "Healthcare Services"
Outside VC Favored Industry	Observations whose industries are outside the industries defined above

Control Variables	Description
Round Entry	The variable shows entry round of the investor
US Annual GDP Growth	US GDP growth from January to December of a certain year. Data is collected from World Bank site. The variable is lagged by 1 year to exit year of investments.
US Annual Industrial Production Growth	US Industrial Production growth from January to December of a certain year. Data is collected from World Bank site. The variable is lagged by 1 year to the exit year of investments.

US Annual Consumer Sentiment Index Growth	US Consumer Sentiment Index growth from January to December of a certain year. Data is collected from University of Michigan website. The variable is lagged by 1 year to the exit year of investments.
Industry Segment Fixed Effects	A vector of indicator variables, taken from the original dataset, that classify the industry of the startup portfolio firm. There are 27 industries: "Aerospace and Defense", "Agriculture and Forestry", "Biopharmaceuticals", "Business Support Services", "Communications and Networking", "Construction and Civil Engineering", "Construction and Civil Engineering", "Consumer Information Services", "Electronics and Computer Hardware", "Financial Institutions and Services", "Food and Beverage", "Healthcare Services", "Household and Office Goods", "Machinery and Industrial Goods", "Materials and Chemicals", "Media and Content", "Medical and Information Services", "Non-renewable Energy", "Personal Goods", "Renewable Energy", "Retailers", "Semiconductors", "Software", "Travel and Leisure", "Utilities", "Vehicles and Parts", "Wholesale Trade and Shipping"
Exit Year Fixed Effects	A vector of indicator variables corresponding to the exit year of investments.
Hot Investment Year	An indicator variable that takes the value of 1 if the number of venture capital investments that year is more than the median of number of yearly investments for our entire sample, and 0 otherwise. This is used as an Instrumental Variable.
Investor Capital under Management	The amount of capital managed by the venture capital investor at the time of the investment taken from the *VentureXpert* database. This is used as an Instrumental Variable.
Investor Industry Preference Match	An indicator variable that takes the value of 1 if the investor's preferred industry SIC code or investment, taken from the *VentureXpert* database, matches the portfolio firm industry code, at the time of investment for the venture capital investor, and 0 otherwise. This is used as an Instrumental Variable.

Investor Stage Preference Match	An indicator variable that takes the value of 1, when whether the investor's preference for early or late stage investments, taken from the *VentureXpert* database, matches the stage of the current investment in a portfolio firm, and 0 otherwise. This is used as an Instrumental Variable.
Portfolio Company Variables	Description
Length of VC investment	The length in years from the time of venture capital investment entry to exit.
% VC & Other Outsider Directors on Board	The percentage of VC and outside directors on the board of a portfolio company as a percentage of the total number of directors.