How Does Venture Capital Involvement Affect Deal Flow? Evidence from Billions of Smartphone Signals

Jack Xiaoyong Fu

October 26, 2026

Abstract

I measure venture capital (VC) involvement intensity and deal flow using 163 billion smartphone pings around VC and startup offices. Exploiting plausibly exogenous variation in travel ease, I show that a one-standard-deviation increase in involvement raises new pitches by 72%. The new pitches tilt toward the visited portfolio company's industry and toward startups raising their first round, consistent with word of mouth among founders. Although the quality of the new pitches is largely unchanged, involvement still improves the VC's future investment outcomes. My model-based estimates imply the larger deal flow alone could explain about a third of this improvement.

JEL classifications: G24, L26, D83
Keywords: Venture capital, deal flow, active involvement, geolocation data

---

Jack Xiaoyong Fu is at the University of Hong Kong (jackfxy@hku.hk). I am extremely grateful to Itay Goldstein, Huan Tang, Luke Taylor, and Yao Zeng for their invaluable support and continuous guidance. I would like to thank Rus Abuzov (discussant), Jan Bena, Shai Bernstein (discussant), Philip Bond, William Diamond, Ran Duchin (discussant), Michael Ewens, Daniel Garrett, Vincent Glode, Paul Goldsmith-Pinkham, Juanita Gonzalez-Uribe, John Griffin, Marius Guenzel, Gerard Hoberg, Johan Hombert, Shenje Hshieh (discussant), David Hsu, Alex Imas, Sasha Indarte, Hyeik Kim (discussant), Yingxiang Li, Tong Liu, Song Ma, Gregor Matvos, Jean-Marie Meier, Marcus Painter (discussant), Gordon Phillips, Manju Puri, Neng Wang, Michael Weisbach, Emmanuel Yimfor (discussant), and Jonathan Zandberg, as well as seminar and conference participants at AFA (2025), FRA (2024), WFA (2024), ASU Sonoran Winter Finance Conference (2024), MFA (2024), Wharton-INSEAD Doctoral Consortium (2023), and WEFI PhD Workshop (2023), for their helpful comments and suggestions. I am grateful for funding from the Mack Institute and the Jacobs Levy Equity Management Center. The Institutional Review Board at the University of Pennsylvania determined that the research meets the eligibility criteria for IRB review exemption authorized by 45 CFR 46.104, category 4. This paper previously circulated under the title "How Does Active Involvement Benefit Investors? Evidence from 85 Billion Cell Phone Signals." First version: May 7, 2024.

1. Introduction

Deal flow, the pool of startups that pitch to a venture capitalist (VC) for funding, matters for the VC's investment performance. VCs rank it among the three most important factors for their success (Gompers et al., 2020), and differences in deal flow explain much of the performance gap between VCs (Sørensen, 2007; Nanda et al., 2020). But what determines deal flow? This question has been hard to answer, largely due to a lack of deal-flow data. Anecdotal evidence suggests one important determinant could be a VC's involvement with its portfolio companies: before choosing a VC, founders are often advised to ask the VC's existing portfolio companies how much help they received (Suster, 2010). By helping its portfolio companies, a VC may therefore attract other startups seeking hands-on support. In this paper, I study the causal relation between a VC's involvement and its future deal flow. The paper makes two contributions. First, it provides granular measures of VC involvement and deal flow at scale. Second, it shows that a VC can generate more deal flow by being more involved with its portfolio companies. The larger pool, in turn, improves the VC's investment performance. Involvement therefore pays off not only in the portfolio company that receives it (Bernstein et al., 2016) but also in the VC's deal flow and future investments.

To study the causal relation between involvement and deal flow, I face two empirical challenges. First, neither involvement nor deal flow has a direct granular measure. Surveys report aggregate patterns, but no existing dataset records how much time each VC spends with its portfolio companies each month, or which startups later pitch to it. Second, even with both measured, involvement is endogenous. Omitted variables, such as a VC's reputation, could drive both how much it visits its portfolio companies and how many startups pitch to it. Identifying the effect of involvement therefore requires exogenous variation.

To address the first challenge, I construct direct granular measures of both involvement and deal flow from smartphone geolocation data. The data cover January 2018 through January 2023. They contain roughly 163 billion location pings around the office buildings of 17,065 investors and 41,875 US startups. I measure involvement in two steps. First, for each VC and each startup, I identify its employees as the smartphones regularly present at its office on weekdays. Second, I measure a pair's involvement as the minutes a VC's employees and a portfolio company's employees spend at each other's offices in a month. A pitch, by contrast, is the first visit between a VC and a startup that have no prior deal. Together, the two measures record how much time each VC spends with each of its portfolio companies every month, and which startups later pitch to it.

To address the second challenge, I use an instrumental variable (IV) for involvement:

passersby. A passerby is a smartphone that appears near both the VC's office and the portfolio company's office in the same month and is not identified as a likely employee of *any* VC or startup in the sample. The number of passersby proxies for how easy it is to travel between these two particular offices that month. When the trip gets easier, the VC visits more. The first-stage F statistic is well above the conventional threshold of 10. Intuitively, the IV extends the direct-flight shocks used by Giroud (2013) and Bernstein et al. (2016). The direct-flight shocks capture travel conditions between distant cities, for pairs far enough apart to fly. Passersby capture those as well, and add the travel conditions of pairs close enough to drive. For example, a wildfire in California keeps people off the roads for weeks: passersby between a nearby VC and its portfolio company fall, and so do the VC's visits.

Passersby proxy for travel ease, but imperfectly, because passersby rise for one of two reasons: travel gets easier, or the demand for travel gets higher. The exclusion restriction in this setting therefore has two parts, both conditional on the fixed effects: travel demand, which is plausibly unrelated to involvement, must be unrelated to future pitches, and travel ease must affect future pitches only through involvement.

The first concern with the exclusion restriction is that travel demand might be related to pitches. Such a relation can arise in four ways. The first is a static difference across pairs: some pairs sit in dense downtown areas, so passersby are high in every month, and pitches may be high too, for reasons unrelated to involvement. To address this, I include VC–portfolio-company pair fixed effects, which absorb all static differences across pairs, so identification uses only changes within a pair over time. The second is a shock to one city: a local economic boom puts more people on the street, so passersby rise, and more founders in the market, so pitches rise. To address this, I include VC-city-by-quarter and portfolio-company-city-by-quarter fixed effects, which absorb local economic booms and any other shock common to either city in a quarter. The third is a shock to both cities at once. The most salient example is an industry boom: a biotech boom moves biotech people between Boston and San Francisco, so passersby rise, and it sends more biotech founders to pitch, so pitches rise. To address this, I add industry-by-quarter fixed effects in the robustness checks to absorb shocks of this kind. The fourth is a shock at the VC's own building: a VC that has just had an extremely successful exit is in the news, so more people (e.g., journalists and service providers) come to its building, some of them also pass near its portfolio companies, and measured passersby rise, while the same news draws more pitches. To address this, in the robustness checks, I add as a control the average passersby between the VC and its other portfolio companies in the same quarter. A shock at the VC's building raises passersby for all of its portfolio companies at once, so the control likely absorbs it.

The second concern with the exclusion restriction is that travel ease might affect pitches directly. Specifically, when the trip between a VC and one of its portfolio companies gets easier, other startups might also reach the VC more easily and pitch more. Two features of the design limit this concern. First, passersby measure one specific trip, between the VC's office and one portfolio company's office, while the pitches being counted come from other startups, whose routes to the VC are different. A shock that does make the VC easier to reach for everyone, like a new airport, is largely absorbed by the VC-city-by-quarter fixed effect, so the travel ease that remains is improvement on one specific route, toward the focal portfolio company. Second, if travel ease brought pitches directly, the new pitchers should be located where travel improved, near the focal portfolio company. I show in Section 5 that they are not: pitchers that share only a state with the focal portfolio company account for essentially none of the increase.

The 2SLS results show that involvement generates deal flow. The outcome variable is the number of new startups that pitch the VC in a later quarter. A one-standard-deviation increase in involvement brings the VC about 70 more new pitches in the next quarter, a 72% increase over the sample mean of 96 pitches per quarter. The effect persists and fades: the additional pitches drop from 70 to 25 two quarters out, and 20 three quarters out.

Where do the new pitches come from? A natural channel is word of mouth. For word of mouth to bring pitches, two things should hold: the portfolio company that receives the involvement should like it, and other founders should be able to hear about it. On the first, more frequent on-site involvement by a VC improves portfolio companies' performance (Bernstein et al., 2016), and anecdotal evidence suggests that founders want regular interaction with their investors.[1] On the second, VCs often give the startups that pitch to them a list of portfolio companies to call, including ones that struggled (Suster, 2010). Founders also review their investors online for other founders to read.[2] Word of mouth could also explain why pitches rise within a quarter and fade afterward: news of the involvement may spread quickly and then go stale. If the news spreads out from the portfolio company, more of the new pitches should come from startups connected to it. I therefore ask how the 70 additional pitches are linked to the portfolio company that received the involvement. From PitchBook, I construct three links between the pitching startup and the portfolio company: whether

[1] David Politis, founder of BetterCloud, describes initially resisting additional investor calls, then finding that "once we started doing those bi-weekly calls, it immediately improved my relationship with my investors" (Politis, 2024). Carly Zakin, cofounder of theSkimm, identifies regular access as one of the best parts of working with Homebrew: "a specific weekly forum to discuss anything that comes up on either side on a regular basis" (quoted in Patel, 2017).

[2] TheFunded, a site on which entrepreneurs rate venture capitalists anonymously, claimed 16,101 registered CEO or founder members and over 25,000 ratings and reviews by 2011 (Mangalindan, 2011).

they are in the same industry, whether their CEOs attended the same school, and whether they are based in the same state. About 27% of the additional pitches come from startups that share only an industry with the portfolio company. The alumni link contributes a small but statistically significant share. Startups that share only a state contribute essentially none. These patterns suggest the news travels along professional lines, not geographic ones.

Who initiates the new pitches? They could be VC initiated: spending time at the portfolio company, the VC learns about startups in its founders' circles and invites them to pitch. They could also be startup initiated: the portfolio company spreads word of the VC's involvement, and more founders decide to pitch. The composition of the new pitchers helps separate the two. If founders seek involved investors because they value hands-on help, the new pitchers should be the ones who need that help most: inexperienced startups raising money for the first time. That is what I find: a one-standard-deviation increase in involvement raises the share of new pitchers raising money for the first time from 22% to 26% two quarters out. A VC inviting startups would have little reason to favor unproven ones.

Involvement increases deal flow, but are the new pitches as good as the old ones? Measuring a startup's quality at the time of the pitch is difficult, because no clear indicator of quality exists at that stage. I therefore use an ex post measure: whether the startup eventually exits through an IPO or M&A. I find no decline in quality: the exit rate of new pitchers is largely unchanged at every horizon. Deal flow grows without getting worse.

How does the increased deal flow affect the VC's investments? The number of investments does not move: although pitches rise, the VC does not invest more often. The investments it makes, however, succeed more often. The 2SLS results show that a one-standard-deviation increase in involvement raises the share of new investments with an IPO or M&A exit recorded in PitchBook by 3.6 percentage points, against a sample mean of 15.5%, a 23% improvement. Taken together, deal flow grows, its quality stays about the same, and the VC picks about the same number of startups, yet the picks turn out better.

Why does more deal flow of the same quality lead to better investments? Because the VC does not invest in the average pitch: it invests in the best pitches it can find, and the best of many is better than the best of few. However, the 2SLS results cannot say by how much. They show that involvement increases deal flow. They show that involvement improves future investment outcomes. But they do not show how much of the improvement in future investment outcomes can be attributed to the increased deal flow rather than to other channels: involvement may, for example, enhance VC learning and thereby improve

the VC's selection ability itself. I use a model to assess how large the gain from larger deal flow alone could be.

I build a simple model in which a VC receives noisy signals of startup quality and chooses a fixed number of investments to maximize their expected success rate. Greater selection ability means more precise signals. The model shows how a larger pool can improve outcomes even when average quality and selection ability remain fixed. To calibrate the model, I rank startups within each VC's pool using subsequent valuation growth as a proxy for quality. I then choose signal precision so that the model's predicted average quality rank of selected startups is as close as possible to the average rank of the VC's actual investments. Finally, I use the coefficients from the 2SLS estimates: a one-standard-deviation increase in involvement raised deal flow by 72% and raised the exit rate of new investments by 3.6 percentage points. When I feed the 72% increase in deal flow into the baseline calibrated model, it predicts an exit-rate gain of 1.1 percentage points, about 32% as large as the estimated 3.6-percentage-point gain. This comparison suggests that larger deal flow alone can make an economically meaningful contribution to investment outcomes.

I validate the involvement measure in three ways. The first is a smell test. I sort every possible VC-startup pair into three groups: the VC eventually led a deal with the startup, the VC invested but never led, or the VC never invested in the startup. I find that pairs where the VC led show roughly 60 times more smartphone signal than never-invested pairs, and pairs where it invested without leading show 25 times more. The second check is about timing. I find that visits rise in the two to three months just before a pair's deal, when due diligence takes place, and stay elevated afterward, when the post-investment work begins. The third check compares statistics with surveys and prior studies, and I find comparable results (Gompers et al., 2020; Bernstein et al., 2016). I also use the measure to provide new stylized facts about how VCs allocate their attention. The most striking is the mismatch between where time goes and where profit comes from: a VC spends about 80% of its meeting time on portfolio companies whose return multiples end up below one, and those companies, in aggregate, contribute little to its profits. While the direct benefit seems too small to justify the cost, the deal-flow benefit I document may explain why VCs keep helping these underperforming companies: founders learn how a VC behaves when its portfolio companies struggle, and a VC known for helping them attracts more pitches. I also document other stylized facts: involvement concentrates on early-stage portfolio companies and declines when a deal has more co-investors or when the investor is larger.

Nevertheless, the measure has clear limitations. It misses virtual meetings and phone calls. It misses in-person meetings when a phone is off or its location apps are not running

in the background. Involvement and deal flow, as measured here, are therefore their in-person components rather than totals. The in-person component, however, is a central one. VCs add value in large part by coaching founders and by helping with hiring and firing decisions, and both benefit from in-person interaction. On the pitching side, evaluating a new startup requires judging the management team, which also benefits from being in person, so pitching typically begins with an in-person meeting. Beyond coverage, several measurement concerns remain. First, in shared buildings, visits could be assigned to the wrong tenant. To address this, I restrict the sample to addresses shared with no other sample company, and the main result holds. Second, some meetings moved online during COVID, which the smartphone signals cannot capture. To address this, I re-estimate on the pre-COVID period, when meetings were mostly in person, and the result holds. Third, VCs and founders also meet outside their offices. Anecdotal evidence suggests that coffee shops are a common venue. To address this, I locate the coffee shops nearest to each VC and startup office using SafeGraph points-of-interest data, purchase additional geolocation data around those venues, and rebuild the measure to include meetings there. The main result still holds.

The paper speaks first to the literature on VCs' active involvement in portfolio companies. Classic work documents intensive post-investment engagement: time spent advising and monitoring (Gorman and Sahlman, 1989; Sahlman, 1990; Lerner, 1995), and the professionalization of young startups, including founder replacement (Hellmann and Puri, 2002; Ewens and Marx, 2018). To study the effect of involvement, Sørensen (2007) uses structural estimation to separate the treatment effect from selection. Bernstein et al. (2016) provide the first causal evidence that involvement improves portfolio-company outcomes. Ewens and Sosyura (2023) further show that the individual VC director's irreplaceable value to the portfolio company lies largely in the personal financing network. Founders appear to value what a VC brings beyond capital: they accept lower valuations from more reputable VCs (Hsu, 2004), and they weigh whether an investor will keep supporting them after the money arrives (Khanna and Mathews, 2022). I contribute to this literature in two ways. First, I provide a direct measure of VC involvement at scale: the meeting time between a VC and a portfolio company, measured monthly and pair by pair, across the whole industry. Existing evidence on involvement comes from surveys (Bottazzi et al., 2008; Gompers et al., 2020), from the detailed records of one program (Sariri, 2025), or from exogenous shocks to involvement, such as new direct flights (Bernstein et al., 2016). Second, prior work focuses on the benefits of involvement to the portfolio company. I study its benefit to the investor: involvement attracts the deal flow from which future investments are drawn.

The paper speaks second to the literature on VC deal flow. Existing work shows that deal

flow matters (Gompers et al., 2020), but the drivers documented so far are largely outside the VC's control. Success in early investments, partly luck, improves later access (Nanda et al., 2020; Cong and Xiao, 2022), public recognition raises it (Jackson and Strebulaev, 2026), and litigation reduces it (Atanasov et al., 2012). In theory, VCs can build deal flow deliberately, by restraining fund size and fees to attract founders (Marquez et al., 2015) or by committing to add value (Sannino, 2024), but this channel has lacked direct empirical evidence. Part of the reason is that deal flow itself is rarely observed. Related work studies the networks, information frictions, and local capital supply that determine who meets whom (Hochberg et al., 2007; Howell, 2020; Howell and Nanda, 2024; Cook et al., 2026; Chen and Ewens, 2025). Previous work collects observations from a single VC (Kirsch et al., 2009; Petty and Gruber, 2011; Petty et al., 2023) or a single seed fund (González-Uribe et al., 2026). My contribution is again twofold. First, I provide a measure of deal flow at large scale: observed pitches from startups to each individual VC, month by month, across the whole industry. Second, I show that involvement generates deal flow. Unlike luck or recognition, involvement is a choice: deal flow becomes something a VC can build rather than inherit.

Finally, the paper joins a young literature that measures economic activity with smartphone geolocation data. This paper is among the first to use such data in finance.[3] The most closely related is Fu and Taylor (2026), who use smartphone signals to measure VCs' due diligence before the deal and its effect on capital misallocation. In this paper, the data supply both sides of the question: the involvement a VC gives and the deal flow it receives.

2. Data and measure

2.1. VC involvement and deal flow

Venture capitalists spend most of their working hours on two activities: sourcing deals and assisting their portfolio companies. The average VC in the Gompers et al. (2020) survey reports a 55-hour work week, of which 22 hours go to networking and sourcing deals and 18 hours to working with portfolio companies. Venture investors add value by sitting on boards, recruiting executives, shaping strategy, and connecting portfolio companies to customers and later-round investors (Gorman and Sahlman, 1989; Sahlman, 1990; Lerner, 1995; Hellmann and Puri, 2002; Bottazzi et al., 2008; Ewens and Malenko, 2026). The interaction is frequent: about 60% of VCs in the same survey report interacting with their portfolio companies at

[3]Applications in the broader economics literature include Atkin et al. (2022) and Miyauchi et al. (2025). Besides Fu and Taylor (2026), the only other finance application I am aware of is Gerken et al. (2026), who measure the SEC's interactions with the firms it oversees.

least once a week. It is also often onsite: almost 90% of the VCs surveyed by Bernstein et al. (2016) agree that they would visit a portfolio company more frequently following the introduction of a direct flight. In-person contact is not incidental: Lee (2026) shows it causally generates informational advantages for investors, with interpersonal trust-building as one key benefit. Throughout the paper I call this in-person value-adding activity *involvement*.

Sourcing, in turn, depends on deal flow. Unlike stock market investors, who can buy any listed stock, a venture investor faces a much narrower menu: in practice, its choices are largely limited to the startups that are willing to pitch it and to accept its money. Deal flow therefore shapes the investor's choice set. Few pitches survive the funnel: for each deal it closes, the average VC in the Gompers et al. (2020) survey considers roughly one hundred potential opportunities and meets the management of about 28.[4] An investor that attracts more pitches chooses from a deeper pool.

This paper connects the two activities by asking whether involvement generates deal flow. The connection is hard to study because neither side of it appears in traditional data. Commercial databases record completed deals, not the pitches that precede them, and no dataset records how much time investors spend with their portfolio companies. Surveys provide totals but not variation: they cannot say which VC, in which quarter, was more involved, and what happened to its deal flow next. Smartphone geolocation data therefore provide a new way to study this question: they record, for every investor-startup pair and every month, both how much time the two sides spend together and whether a startup pitches to an investor.

2.2. Measuring involvement and pitches

I collected the headquarters and branch office addresses of the investors and startups from PitchBook. Using the Google Maps API, I converted each address into coordinates, and I purchased smartphone location signals within a 200-meter radius of these coordinates from January 2018 to January 2023 from the data vendor. The data vendor collects these data from hundreds of popular apps spanning messaging, social media, navigation, music, weather, travel, health, and other categories, and reports coverage of 220 to 240 million monthly active users in the US, roughly 80% of all smartphones.[5] Those apps record the

[4]Individual VCs report similar funnels. Kindred Capital reports roughly 500–600 first meetings for about ten investments per year (Kindred Capital, 2019), and Homebrew reports 399 first meetings for nine investments (Patel, 2014).

[5]According to a 2023 Pew Research Center report (see `https://www.pewresearch.org/internet/fact-sheet/mobile/`), 90% of US adults report owning a smartphone, and the US population is approx-

longitude and latitude of a phone with timestamps, subject to user permissions, when they are open or backgrounded. The location estimates can be accurate to within 20 meters. Each signal carries a device identifier that is hashed to comply with privacy standards.

The construction then proceeds in two phases. The first phase identifies the devices likely belonging to each VC's and each startup's employees. To separate employees from passersby, I use the frequency and timing of a device's appearance near the office: an employee's device should be detected there regularly on working days. Since a typical month has about 20 working days, I flag a device as a likely employee of a building if it appears near the building on at least five weekdays in a month, in at least two months. To further separate employees from frequent visitors such as delivery workers, I exclude devices that would qualify as employees at more than ten employers in a month.

The second phase detects meetings. When a device belonging to one side's likely employees is detected at the other side's building, I record a visit event: its date, how long the device stayed, and in how many distinct minutes it pinged during the stay. I keep visit events in which the device pinged in at least three distinct minutes and stayed between 30 and 300 minutes. Both bounds reduce false positives. The lower bound drops brief pass-bys such as couriers and pedestrians, and the upper bound drops all-day presences that look like a shared workspace rather than a meeting. Visit events come in two directions, with a VC employee visiting the startup's building or a startup employee visiting the VC's building. In the baseline I pool both directions. Section 7 varies the duration thresholds and reports a version that keeps only one direction on each side. The effect survives both changes.

From the qualifying visit events I build the involvement measure. For each pair of investor and portfolio company and each month, I take, for each device and day, that device's longest visit, and I add these durations across devices and days within the month. For example, if one partner visits on two days for two hours each and a colleague visits once for 30 minutes, the pair's total for the month is 270 minutes. I call this total the pair's total visit minutes for the month. The regression variable averages this monthly total over the months of the quarter and enters as the log of one plus the average for each pair, which I label Log (Involvement).

From the same visit events I build the pitch measure. An ideal measure of deal flow would be a comprehensive list of the entrepreneurs who want to pitch each VC, but no such list exists. I define a pitch as the first qualifying visit, in either direction, between an investor and a startup that have no deal with each other. The definition includes pairs that never transact at any point, so the pitch panel is not limited to pitches that succeeded. Visits in or

imately 335 million in 2023, so the number of smartphone users is approximately 300 million.

after the month of a pair's first deal are excluded, because they reflect post-deal work rather than pitching. Each pair enters the pitch panel once, at the quarter of its first visit. The measure covers every pair among the 17,065 investors and 41,875 startups in the smartphone data, so it is limited to startups that appear in PitchBook. It captures a startup's pitches made before its first funding round, as long as the startup eventually raises a round that PitchBook records, but it misses startups that never raise from anyone.

The two measures could contain measurement error in three ways. The first source of measurement error is coverage. A phone can be turned off, lose reception, or have the relevant apps neither open nor running in the background, so some meetings leave no trace. However, whether a given meeting is captured largely depends on individual phone habits, such as how often someone clears background apps, so this error is plausibly classical measurement error once fixed effects are in place.

The second source of measurement error is venue. Meetings held away from the two sides' offices are not counted. In particular, after COVID, many meetings are held online. However, on-site visits and video calls are likely substitutes for each other. When travel becomes easier, a VC makes more on-site visits and fewer video calls, so the effect I estimate is the effect of on-site involvement relative to the video calls it replaces. Because a video call likely provides some value as well, the true effect of on-site involvement is even larger, so my estimate is more likely an underestimate than an overestimate. Section 7 reports a version of the pitch measure that adds first meetings at coffee houses near either side's office, with nearly identical results.

Another concern is that different locations can have different false-positive rates. A startup office in Times Square attracts more false positives than one in a quiet suburb, and a pair whose buildings sit within a mile of each other attracts more than a pair in different cities. However, location and distance do not change over time, so the pair fixed effects absorb these differences.

The third source of measurement error is attribution. In a building shared by several tenants, a visit can be assigned to the wrong one. Section 7 mitigates this concern by restricting the sample to addresses shared with no other sample company and to headquarters buildings, and the estimates are similar.

2.3. Sample formation

The sample starts from PitchBook. I collect all deals satisfying the following criteria: both the investor and the portfolio company are headquartered in the US with available office addresses, the deal date falls between January 2017 and January 2023, so that every focal window overlaps the smartphone coverage, which begins in January 2018, the founding years of both parties are available, PitchBook lists the investor's status as active, and the investor's primary type is not a corporation, a PE-backed company, a VC-backed company, or "other." These criteria yield 157,675 deal records, each pairing one investor with one portfolio company in one financing round. The smartphone purchase described in Section 2.2 covers the headquarters and branch offices of every investor and portfolio company in these records, and the raw delivery contains roughly 163 billion location pings over the five-year window. Among these records, I keep those whose deal is an explicit venture round (Seed through late-stage series, which drops accelerator-only, angel-only, and exit transactions) and whose deal date and round sequence are recorded. The 2023 PitchBook vintage defines this deal universe and all deal characteristics. A second PitchBook vintage, downloaded in September 2025, supplies the forward-looking outcomes: later financing rounds, exits, and each investor's subsequent investments. Outcomes are therefore measured with two extra years of history rather than with the data available at the time of the deal.

Each retained record defines a focal deal with a focal window: the 24 months following the investment date or the period from the investment date to the portfolio company's next qualifying round, whichever is shorter. One observation is one investor in one deal in one quarter, so a deal with three investors contributes three observations per quarter. I follow each focal deal over its window, at most nine quarters, and the panel has 661,259 observations.

Deal flow is measured over a far larger universe than the focal deals: the pitch measure covers every investor-startup pair in the smartphone data, not only the pairs in the focal deals. For an investor i and quarter $t + h$, the outcome *number of new pitches* counts the startups whose first visit with i falls in quarter $t + h$. Outcomes at horizon h are set to missing when $t + h$ falls after the smartphone coverage ends in 2023Q1, because those cells are unobserved rather than zero.

The focal deals are venture financing rounds, but the investors that participate in them are broader than venture capital firms: PitchBook classifies them into venture capital, private equity, and other types. With slight abuse of terminology, I refer to these investors as VCs.

On the deal-flow side, an average VC receives about 96 new pitches per quarter and makes 5.8 new investments (Table 1). On the involvement side, 10.1% of pairs of VC and portfolio company record at least one qualifying visit during their focal windows, and 27.2% of portfolio companies record at least one qualifying visit from one of their investors. In quarters with a visit, the median pair records 85 visit minutes per month.

Why do so few pairs record a visit? Some investors, especially non-lead investors, may never visit a portfolio company in person. But the main reason is likely measurement error. The data capture only a fraction of the meetings that take place. A visit enters the data only when the visitor's phone pings during the stay, which requires a location-reporting app to be open or running in the background. The filters against false positives in Section 2.2 remove some true meetings as well.

Even so, a zero is informative. If the chance of capturing a meeting rises with the number of meetings held, then a pair-quarter with no recorded visit likely had few meetings. I therefore keep the pair-quarters with no recorded visit in all regressions. Dropping them would select the sample on the regressor itself.

3. Stylized facts and validation

This section first provides two validation tests of my involvement measure and a comparison against survey benchmarks, and then uses the measure to document several stylized facts about VC involvement.

3.1. Validation

The first test asks whether the intensity of smartphone signals is higher when a pair has a true investment relation. I take every combination of a VC and a startup in the PitchBook deal sample and split the pairs into three mutually exclusive groups: pairs in which the VC ever invested in the startup as the lead VC on at least one deal, pairs in which the VC ever invested but never as the lead, and pairs in which the VC never invested. To keep the data-cleaning filters from affecting the result, I apply no visit filters and use the simplest statistic: total unique ping minutes. For example, if a device assigned to an employee of a VC is observed near a startup's building at three timestamps on January 1, 2020, at 10:00:01, 10:00:05, and 10:10:49, the number of unique ping minutes for that day is two, because the three pings fall in two distinct minutes, 10:00 and 10:10. I sum unique ping minutes for each

pair over the entire coverage window, January 2018 to January 2023, regardless of whether or when the pair has a deal. If the smartphone signal data were purely noise, the three groups would show similar signal intensity. Instead, the mean signal intensity per pair is roughly 60 times higher for lead pairs and 25 times higher for non-lead pairs than for never-invested pairs (Figure 1), and the ordering follows the depth of the relation. Even this simplest statistic, computed before any visit definition or filter is applied, ranks the three groups the way true investment relations would predict.

The second test asks whether the measure moves with the investment relation in time. If the measured visits reflect portfolio involvement, they should intensify after the investment date and stay elevated while the VC works with the portfolio company. Figure 2 plots monthly meeting minutes between the pair by month relative to the pair's first investment date, for pairs whose first deal falls between January 2019 and December 2021, so that each pair has a full window of 12 months before and 12 months after the investment date within the smartphone coverage.

Panel (a) plots raw means, separately for lead and non-lead VCs. Three patterns stand out. First, the two series move together, and the lead VC's interaction is almost always higher than the non-lead VC's. Second, meeting minutes are positive before the investment date, which is expected: the months before the deal contain startup pitching and VC due diligence. Third, the two roles differ after the investment date. Lead-VC involvement rises after the deal and peaks eight to nine months after the investment date. Non-lead involvement instead shows a distinct peak at one year after the investment date, consistent with a syndicate member that is not expected to be involved frequently and meets the portfolio company around a full year after investing.

Panel (b) plots, for lead VCs, the coefficients on months relative to the investment date after absorbing calendar-month fixed effects. Relative to the investment month, meeting minutes before the investment are mostly lower, with the exception of the two to three months just prior, which are likely the peak of the due diligence period. Meeting minutes after the investment are mostly higher. Involvement as measured here begins when the investment relation begins.

Table IA.1 compares the levels of the smartphone measures with two external benchmarks. Among investors that PitchBook classifies as venture capital, the median VC-year sees 33.7 distinct pitching startups for every startup it invests in,[6] and Gompers et al.

[6]The Table 1 means, 96 new pitches and 5.8 new investments per quarter, imply a smaller ratio because they weight each VC by its number of focal deals and so describe the most active VCs. The median here describes the typical VC-year.

(2020) report that 28 potential investments reach the "met management" stage per closed deal. Within focal deals that have at least one captured visit, the mean number of distinct visit days is 1.21 per month, and the survey respondents of Bernstein et al. (2016) report 9.93 visits per year to a portfolio company on average, or 0.83 per month. While the absolute levels potentially suffer from the measurement errors described in Section 2.2, the magnitudes are roughly in line with the survey benchmarks.

3.2. Stylized facts

3.2.1. Stylized fact 1: VCs spend most of their time on underperforming portfolio companies

Figure 3 sets the allocation of involvement against the distribution of returns. I sort each lead VC's deals by their eventual return multiple and compare the share of meeting time spent on each return group with the share of total venture returns that the group produces, the latter from Horsley Bridge data reported by Benedict Evans (2016). The allocation of time is nearly the reverse of the distribution of returns. Deals returning less than one times capital absorb about 80% of lead-VC meeting time but produce less than 3% of total returns. Deals returning more than ten times capital produce 60% of total returns but absorb less than 5% of meeting time. The returns side of this pattern is well documented: venture returns are dominated by a few large winners. The time side is what the smartphone data add: they show how a VC divides its time across its portfolio companies.

Figure 4 shows the same pattern with exits. Portfolio companies that never reach an IPO or M&A receive roughly 1,250 minutes of meeting time over the focal window, against roughly 730 for those that eventually exit successfully.

Why is the allocation of involvement negatively correlated with portfolio-company returns? I consider three interpretations. The first interpretation is that VCs cannot tell which portfolio companies will succeed and spread their time evenly across them. Because failures outnumber successes, most total time would then end up on those that fail. Figure 4 contradicts this interpretation: if time were spread evenly, time per portfolio company would be flat across eventual outcomes, but those that end badly receive nearly twice the attention of those that end well. The second interpretation is a negative treatment effect, in which VC attention disrupts the founders it aims to help. The existing causal evidence contradicts this interpretation as well: Bernstein et al. (2016) show that easier access to portfolio companies raises innovation and the probability of a successful exit. The third interpretation, and the

one most consistent with the evidence, is that VCs direct help where it is needed. This needs-based allocation echoes Lerner (1995), who finds that venture capitalists are more likely to join the boards of their portfolio companies around CEO turnover, when the need for oversight is greatest.

If the portfolio companies that need the most help still mostly underperform, why do VCs keep helping them? A likely motive is future deal flow: by helping struggling portfolio companies, a VC builds a founder-friendly image that spreads by word of mouth and attracts future pitches. Practitioners describe the motive in exactly these terms. Mark Suster of Upfront Ventures points founders to his portfolio companies as references because "in tough times I tend to roll up my sleeves rather than head for the door" (Suster, 2010).

3.2.2. Stylized fact 2: Early-stage investments receive more involvement

Involvement declines steeply with the portfolio company's maturity. Total meeting minutes fall from about 1,200 for Seed deals to under 100 for Series E and later (Figure 5). Figure 5 measures the lead VC's involvement. A Series E company has many more investors on its cap table than a Seed company, so the time it receives from all of its investors combined may fall less steeply, or not at all, even if each investor spends little. The direction is consistent with the evidence on where involvement pays: Bernstein et al. (2016) find that the effect of involvement is several times larger for early-stage portfolio companies, whose teams are smaller and whose strategy is less settled.

3.2.3. Stylized fact 3: Deals with more co-investors receive less involvement per VC

When more VCs share a deal, each VC spends less time on the portfolio company. Meeting minutes per VC fall from about 2,000 for deals with a single VC to about 340 when five or more VCs share the deal (Figure 6). This pattern is consistent with the moral hazard in value addition modeled in the VC syndication literature (Bayar et al., 2020), in which each syndicate member counts on the others to do the work.

3.2.4. Stylized fact 4: Larger VCs visit each deal less

Involvement per deal also declines with VC size. Lead-VC meeting minutes decline with assets under management, with the number of investments, and with the number of invest-

ment professionals (Figure 7). This pattern speaks to a finding in Kaplan and Schoar (2005): returns are concave in fund size, and funds that grow quickly after strong performance tend to perform worse. One proposed explanation is partner attention: a bigger fund backs more portfolio companies per partner, leaving less time for each, a tradeoff that Fulghieri and Sevilir (2009) model formally. The meeting-time data show this attention channel directly, consistent with the shift toward larger and less hands-on portfolios that Ewens et al. (2018) document as experimentation became cheaper.

Together these facts describe involvement as a scarce resource that VCs ration deliberately: toward early-stage portfolio companies, toward deals with few co-investors to share the work, and toward portfolio companies that are struggling rather than succeeding. Hands-on time flows to the portfolio companies that need it most. That is why founders who expect to need help would seek out a VC known for providing it.[7] The rest of the paper formally tests the causal relation between involvement and deal flow: does involvement bring the VC more pitches?

4. Empirical strategy

The main research question of the paper is whether involvement causes deal flow. A simple OLS regression of deal flow on involvement does not answer it, because the OLS coefficient mixes two main effects. One is selection. As VCs grow and establish a reputation, they spend less time with each portfolio company (Figure 7) yet receive more pitches, so selection alone produces a negative relation between involvement and deal flow. The other is the treatment effect, the positive relation this paper is after. An OLS coefficient cannot separate the two, so it need not resemble the treatment effect in size or even in sign.

To deal with the endogeneity issue, I use an instrumental variable: the pair's quarterly number of *passersby*. A device counts as a passerby if, within the same month, it is observed at least once within 200 meters of the VC's building and at least once within 200 meters of the portfolio company's building, and it is not identified as a likely employee of *any* VC or startup in the sample. The instrument is the monthly number of distinct passerby devices, averaged over the months of the quarter. Since a device flagged as a likely employee does not enter the passersby count, and both the involvement measure and the pitch measure are built from visits by devices flagged as likely employees, the instrument mechanically does

[7] Practitioner advice points founders to exactly this information. Suster (2010) advises founders to check a VC by speaking with its portfolio companies, including those that stumbled, because a VC's behavior in hard times is the most informative reference.

not have any overlap with either measure.

Passersby are a proxy for the ease of travel between the two specific buildings. Weather, road work, transit changes, and route openings shift the number of passersby who happen to appear near both locations, and when it is easier to travel, the VC is likely to make more on-site visits. Because passersby measure the ease of travel between the two specific buildings, the instrument picks up both travel shocks for long-distance pairs, such as the new airline routes that Giroud (2013) and Bernstein et al. (2016) exploit as monitoring-cost shocks, and local shocks, such as a wildfire in California that reduces travel locally. In this sense, the instrument extends the direct-flight instrument by also capturing travel conditions for pairs at close distance. With the pair fixed effect, the VC-city-by-quarter fixed effect, and the startup-city-by-quarter fixed effect absorbed, the main source of variation is at the VC-location by startup-location by quarter level. Distance, for instance, is well known to be endogenous to investment, but it does not change over time, so the pair fixed effect absorbs it, along with everything else constant about the pair. The intuition is simple: in some quarters, travel between the VC and one of its existing portfolio companies is easier than usual, the VC visits that portfolio company more, gains more total exposure to the startup community, and receives more pitches in the following quarters.

The exclusion restriction of the instrument is as follows: conditional on the fixed effects, travel ease between the VC and a portfolio company, as measured by the passersby between the two buildings, affects the VC's deal flow only through the pair's involvement. Formally, the exclusion restriction requires the instrument to be uncorrelated with the error term of the second-stage regression. In my setting, the exclusion restriction could fail in two ways: a confounder could move passersby and deal flow at the same time, or travel ease could affect deal flow directly rather than through involvement.

The first concern is a local economic boom. A recovery from COVID in San Francisco, for example, could raise both the foot traffic near a San Francisco VC and that VC's deal activity. However, the VC-city-by-quarter and startup-city-by-quarter fixed effects already absorb anything happening in either city in a given quarter, so shocks of this kind are not a concern.

The second concern is an industry boom, a third factor that could move passersby and deal flow at the same time. Suppose most biotech startups sit in Boston and most biotech VCs in San Francisco. A biotech boom sends more biotech people back and forth between the two cities, so more devices appear near both buildings and the passersby count rises. The same boom sends more biotech founders to pitch the VC, so deal flow rises. The city-

by-quarter fixed effects absorb each city's own conditions, but not a shock that loads on the two cities' pairing through one industry. In Section 7, I add industry-by-quarter fixed effects to absorb shocks of this kind.

The third concern is a shock at the VC's own building. Suppose a VC has just had an extremely successful exit and is in the news. More people come to its building, such as journalists and service providers, and some of them also pass near its portfolio companies, so measured passersby rise for all of the VC's pairs at once. The same news draws more pitches. The VC-city-by-quarter fixed effect absorbs shocks common to the city, not a shock at one building. In Section 7, I add as a control the mean of Log (Passersby) between the VC and its other portfolio companies in the same quarter, which carries such a shock, and show that it does not predict the pair's involvement and leaves the main result largely unchanged.

The fourth concern is that easier travel improves deal flow directly, with founders finding it easier to travel to the VC, rather than through involvement. Two features of the design limit this concern. First, the instrument measures travel ease between the VC and one focal portfolio company, not the VC's overall accessibility. If the VC's city opens a new airport, for example, travel to that VC becomes easier for every startup at once. Shocks of this kind are common to the VC's city and quarter, so the VC-city-by-quarter fixed effect already absorbs them. Second, if easier travel drove the inflow directly, new pitchers should largely be located in the same place as the focal portfolio company. Section 5.2 shows that they are not: new pitchers that share only a state with the focal portfolio company account for essentially none of the increase, which comes instead from pitchers that share only the portfolio company's industry or have no observable industry, state, or alumni link at all.

The unit of observation is VC by deal by quarter. For each focal deal, I keep the quarters within 24 months of the investment date, or up to the portfolio company's next qualifying round, whichever comes first. The two stages are

$$\log\big(1 + X_{ic,t}\big) \;=\; \gamma\, \log\big(1 + Z_{ic,t}\big) \;+\; \alpha_{ic} \;+\; \alpha_{\text{city}(c),\,t} \;+\; \alpha_{\text{city}(i),\,t} \;+\; u_{ic,t}, \tag{1}$$

$$Y_{i,\,t+h} \;=\; \beta \log\big(\widehat{1 + X_{ic,t}}\big) \;+\; \alpha_{ic} \;+\; \alpha_{\text{city}(c),\,t} \;+\; \alpha_{\text{city}(i),\,t} \;+\; \varepsilon_{ic,t}, \tag{2}$$

where $X_{ic,t}$ is the total visit minutes per month between VC i and portfolio company c, averaged within quarter t, as constructed in Section 2.2, $Z_{ic,t}$ is the pair's passersby, and $Y_{i,t+h}$ is the number of new startups that pitch VC i for the first time in quarter $t+h$. The same fixed effects appear in both stages: a pair effect α_{ic}, and city-by-quarter effects $\alpha_{\text{city}(c),t}$ and $\alpha_{\text{city}(i),t}$ for the portfolio company's city and the VC's city.

5. Main results

5.1. Involvement and the number of new pitches

The first question is whether involvement brings the VC more pitches. Table 2 reports the IV estimates of equation (2) for horizons of one to three quarters. The unit of observation is VC by deal by quarter. The independent variable is Log (Involvement), instrumented by Log (Passersby). The outcome is the number of new startups that pitch the VC for the first time in quarter $t + h$, with the three columns reporting horizons $h = 1$, 2, and 3. Observations whose outcome quarter falls after the smartphone coverage ends are dropped, which is why the sample shrinks with the horizon. All columns absorb the pair, VC-city-by-quarter, and startup-city-by-quarter fixed effects, and standard errors are clustered by VC.

Involvement brings pitches. The coefficient at the one-quarter horizon is 70.2, significant at the 1% level (column 1). The standard deviation of log involvement is 0.99, so a coefficient of 70.2 means that a one-standard-deviation increase in involvement brings the VC about 70 more pitching startups in the next quarter. Against a sample mean of 96 new pitches per quarter (Table 1), this is a 72% increase.

How large is an inflow of 70 pitches? The median VC sees 33.7 pitching startups for every startup it invests in (Table IA.1). The 70 additional pitches amount to roughly two investments' worth of new deal flow from a one-standard-deviation increase in involvement.

The effect is front-loaded and fades. The coefficient falls to 24.8 two quarters out and to 19.6 three quarters out, both significant at the 1% level (columns 2 and 3). Most of the response arrives within a quarter of the involvement itself, which is the timing one would expect if founders act on recent news about the VC. The Kleibergen–Paap F statistic exceeds 400 at every horizon.

Table IA.2 reports the first stage for each column: the coefficient on Log (Passersby) is 0.092 in the baseline sample and about the same in the other two, with t-statistics above 20. The coefficient is positive and modest, as it should be. VCs visit more when travel is easier, but travel conditions are a small part of why a VC visits a portfolio company. Most visits reflect features of the VC and the company, such as a VC whose expertise matches what the company needs, and the same features are likely to shape the VC's deal flow as well. To identify a causal relation, the 2SLS estimates use only the part of the variation in involvement that is plausibly exogenous.

In what follows, the one-quarter estimate is the headline number, and the longer horizons trace its decay. The next two subsections explore who the new pitchers are and how the news reached them.

5.2. How new pitchers are connected to portfolio companies

An on-site visit is observed by the portfolio company that receives it. How does it reach the wider founder community? In the venture capital industry, a natural channel is word-of-mouth recommendation: founders evaluate a VC by asking the founders who have worked with it, and VCs themselves invite this. Fred Wilson of Union Square Ventures gives founders "a list of every entrepreneur I've ever worked with" and tells them to "throw a dart at that list and talk to four or five of them randomly" (quoted in Suster, 2010). Personal contact of this kind matters for who raises venture funding: Howell and Nanda (2024) show that participants randomly exposed to more VC judges at a business-plan competition are more likely to raise venture capital afterward, in part by reaching out to the VCs they met.

If word of mouth spreads outward from the portfolio company, the founders connected to it should hear the news first. Table 3 splits the total effect into eight mutually exclusive groups according to the pitcher's observable links to the focal portfolio company. I measure three links in PitchBook. An industry link means the pitching startup operates in the same primary industry sector as the portfolio company. An alumni link means their CEOs attended at least one common educational institution, according to PitchBook's executive biography records. A state link means they are headquartered in the same US state. The eight groups are the seven combinations of the three links plus a group with none of them. Because the groups partition the new pitchers, the eight coefficients sum exactly to the total effect of Table 2, column 1.

The new pitches line up with the observable network, and industry ties matter most. Pitchers who share only an industry with the portfolio company contribute about 27% of the total effect, significant at the 1% level (column 1). Such startups account for only 19% of a VC's new pitches on average, so the added pitches tilt toward the focal portfolio company's industry.[8] Pitchers with only an alumni link contribute a small but significant flow (column 2), and the industry-plus-alumni and industry-plus-state combinations add further small significant flows (columns 4 and 5). The remaining link combinations are near

[8]For each observation, I sort the VC's new pitchers in the next quarter into the eight groups of Table 3, defined relative to the focal portfolio company. The 19% is the sample mean of the industry-only count divided by the sample mean of all new pitches, over the observations of Table 1. Computed the same way, startups that share only a state are 17% of new pitches, and startups with no recorded link are 55%.

zero and insignificant: state alone, alumni plus state, and all three links together (columns 3, 6, and 7). The multi-link groups are small by construction, since each requires several links to hold at once, so there is little flow for those columns to detect. Industry appears in almost every significant group, and industry is where word of mouth should travel farthest: founders talk most to founders in their own field, at industry events and in the reference calls that precede a fundraise. The industry pattern is consistent with Eldar and Grennan (2024), who find that startups backed by the same VC benefit mainly from the same-industry startups in its portfolio. Alumni ties matter in venture capital as well: Garfinkel et al. (2025) find that about one third of VC deals pair a founder with an investor from the same university. Their tie runs from founder to investor and shapes who gets funded. The link here runs from founder to founder and carries word of mouth. Geography on its own carries nothing, which already suggests that the news travels through professional circles rather than physical proximity.

The largest group, however, is pitchers with no measured link to the portfolio company, at about 68% of the total effect, significant at the 1% level (column 8). The likely reason this group is so large is that the three links cover only a small part of the ways founders actually know each other. A pitch from a startup with no measured tie suggests the news traveled through connections that PitchBook does not record. Founders who passed through the same accelerator compare VCs with their batchmates. Startups backed by a common VC meet at that VC's portfolio events. Founders who once worked at the same employer stay in touch after they leave. Founder communities, from demo days to online groups, circulate news about VCs continuously. None of these connections appear in PitchBook's industry, education, or location fields, so the pitches they generate land in the no-link group. The one VC whose complete proposal log has been studied looks similar: Petty et al. (2023) report that 37.9% of its proposals came directly from entrepreneurs with no prior tie to the VC and 48.5% through arm's-length intermediaries, paths that these fields would also miss. A natural interpretation of Table 3 is therefore one word-of-mouth process observed at two resolutions: the linked groups show it where PitchBook records a tie, and the no-link group collects it where the tie is real but unrecorded. Involvement raises the VC's *visibility* among founders, and the traceable channels are the part of that process the data can show directly. Founders respond to demonstrated investor characteristics rather than to self-described ones (Ebrahimian and Zhang, 2025), so credible information about a VC's involvement is most likely to come from those who observe it first-hand, which is what word of mouth from portfolio companies provides. If visibility is what draws founders in, the founders it draws should be the ones who value involvement most. The next subsection checks who they are.

5.3. The experience of new pitchers

If founders seek involved VCs because they value hands-on help, the marginal pitcher should be the kind of founder who needs hands-on help: young, unproven, and raising for the first time. If the new pitches instead reflect founders responding to the VC's name recognition alone, there is no reason for the composition to tilt toward first-timers. Table 4 reports the share of new pitchers raising funding for the first time, at each horizon.

The composition tilts toward first-timers, with a lag. The share of first-time funded pitchers is largely unchanged in the first quarter (column 1), then rises by 4.3 percentage points at the two-quarter horizon and by 2.8 at three quarters, both significant at the 1% level (columns 2 and 3). Against a sample mean of 22% (Table 1), the two-quarter estimate raises the share to about 26%. One possible reason for the lag is that founders differ in how quickly they can act on news about a VC. A startup already raising can add the VC to its pitch list right away. A previously unfunded startup may need time to prepare a pitch, so the response takes longer to appear, in which case the volume response arrives first and the composition shift follows.

Figure 5 shows that VCs spend the most involvement time on their earliest-stage deals. Table 4 is consistent with that pattern: the pitchers newly attracted by involvement are also the least established ones. Involvement is a service, and the pitchers it attracts are the service's natural customers. The remaining question is what the larger deal flow is worth to the VC.

5.4. Quality of new pitches and performance of new investments

A larger deal flow is worth little if the marginal pitch is bad. Pitches from previously unfunded startups could dilute the pool. If the pitches that involvement adds are worse than those that would have arrived anyway, the eventual success rate of each cohort of new pitchers should fall with involvement. Table 5 measures each cohort by the share of its startups that eventually exit through an IPO or M&A, at each of the three horizons (columns 1 to 3). Later cohorts have had less time to exit, but the city-by-quarter fixed effects compare cohorts within the same quarter, so that difference is absorbed. Exits are read from the September 2025 PitchBook vintage, so a startup that pitched in 2018 has had about seven years to exit and one that pitched in 2022 about three.

The marginal pitches are about as good as the average pitch. The point estimates are

−0.002, 0.002, and −0.007 across the three horizons, each with a standard error of about half a percentage point against a sample mean of 7.7% (Table 1). Involvement scales the pool without measurably changing pitcher quality.

What does the VC gain from the larger deal flow? Table 6 reports the VC's future investment outcomes: the number of new investments (column 1), the number of those with an IPO or M&A exit recorded in PitchBook (column 2), and the exit rate among them (column 3). These outcomes concern the VC's subsequent investments, rather than the focal portfolio company. The investment count is largely unchanged (column 1): a 72% larger intake leaves the number of new investments about the same, so the VC becomes more selective rather than more acquisitive. But the investments it does make succeed more often. The number of new investments with an IPO or M&A exit rises by 0.24, significant at the 5% level (column 2), an increase of about a third against the sample mean of 0.75 (Table 1). That column keeps every quarter, including those with no new investment. The rate of such exits among new investments, defined only in quarters with at least one, rises by 3.6 percentage points, significant at the 10% level (column 3), a 23% improvement on the sample mean of 15.5%.

More deal flow is not automatically better. Abuzov (2025) shows that deals closed by busy VC partners perform substantially worse because the partners screen them less carefully. A larger inflow could also increase screening demands. Here, investment counts remain roughly unchanged while exit rates rise. These results are consistent with a larger pool allowing the VC to choose better investments. The sample means already show that VCs on average have some selection ability: new investments exit at 15.5%, twice the pitch pool's 7.7% (Table 1). When involvement enlarges the choice set, that bar rises further.

The three results together make a simple point and raise a question. The pool of pitches grows, its average quality stays the same, and the VC makes the same number of investments, yet the investments turn out better. If the average pitch is no better, why does a larger choice set help?

A simple example gives the intuition. Suppose 10% of startups are good and the rest are average, and the VC invests in ten startups each period. First consider a VC that can rank startups perfectly. Facing 50 pitches, it invests in the top ten of its pool, which contain the pool's five good startups and five average ones, so five of its ten investments succeed. Facing 100 pitches, the top ten are all good, and all ten succeed. The share of good startups in the pool never moved and the number of investments never moved, yet the success rate doubled. Now consider the opposite extreme, a VC that cannot rank startups at all and picks at

random. Its success rate is 10% at any pool size, so a larger pool does not help. In reality, a VC sits between these extremes, with some but not perfect ability to rank, and for such a VC a larger pool can still help. Section 6 writes down a model of this more realistic case, estimates the VC's selection ability from where its actual picks rank within its constructed startup pool, and finds a model-implied gain from the larger pool about 32% as large as the estimated performance gain.

6. Model and calibration

Section 5 shows that greater involvement brings a VC more pitches, with little change in average pitch quality or the number of investments. Yet its subsequent investments perform better. The intuition is that the VC does not invest in the average pitch. It invests in the most promising startups it can identify. A larger pool gives it more chances to find a strong startup, even if average pitch quality remains the same. I develop a model showing that a VC can benefit from a larger pool even if it cannot rank startups perfectly. I then calibrate the model to assess the magnitude of this mechanism relative to the improvement reported in Table 6.

6.1. Model

A VC receives pitches from N startups and makes k investments ($1 \leq k < N$). Both the number of investments and the amount invested in each startup are fixed. The VC chooses only which startups to fund, maximizing the expected number of successes. I assume that every selected startup accepts the VC's investment.

Each startup j has quality Q_j, drawn independently from a Pareto distribution:

$$\Pr(Q_j > q) = q^{-\alpha}, \qquad q \geq 1, \quad \alpha > 0.$$

The highly skewed returns on startup investments motivate using a Pareto distribution for startup quality (Levine, 2014; Othman, 2019). Most startups have relatively low quality, while a few have much higher quality. A smaller α gives more weight to these exceptional startups. If financed, startup j succeeds with probability

$$p(Q_j) = \min\{cQ_j, 1\}, \qquad 0 < c < 1.$$

The parameter c determines the overall likelihood of success, and the cap at one keeps the expression a probability.

For each startup, the VC observes a noisy signal S_j:

$$\log S_j = \log Q_j + \frac{Z_j}{\sqrt{\theta}}, \qquad Z_j \sim N(0,1), \quad 0 < \theta < \infty.$$

Signal errors are independent across startups and independent of quality. The parameter θ measures the VC's selection ability: a higher θ means more precise signals and better informed choices. After observing the signals, the VC updates its beliefs about startup quality. A higher signal implies a higher expected probability of success, so the VC invests in the k startups with the highest signals.

Let A_N be the set of startups receiving investment. The expected success rate is

$$R(N, k; \alpha, \theta, c) = E\left[\frac{1}{k}\sum_{j \in A_N} p(Q_j)\right].$$

Proposition 1. Under the model's assumptions, a larger pitch pool raises the expected success rate when the number of investments, the quality distribution, and selection ability remain fixed:

$$R(N+1, k; \alpha, \theta, c) > R(N, k; \alpha, \theta, c).$$

Appendix A.1 gives the proof. The intuition is that an additional startup receives investment only if its signal exceeds the lowest signal among the VC's current choices. Replacing that choice raises the portfolio's expected success rate, although the replacement may still perform poorly. Selection ability is essential: as θ approaches zero, the VC's choices approach random selection. In that limit, a larger pool provides no gain.

6.2. Calibration

Estimating the magnitude of this mechanism is challenging because the simple model omits many features of venture investing, including competition for startups and screening constraints. Nevertheless, the calibration can indicate whether a larger pitch pool can generate an economically meaningful improvement in outcomes.

The model assumes that the VC can finance any startup it selects. In practice, startups also choose their investors, as in Sørensen's (2007) model of two-sided matching. Hsu (2004) finds that startups are more likely to accept offers from VCs with stronger reputations. To make the model's assumption more plausible, I focus on the top 10% of VCs by assets under management (AUM), using size as a proxy for reputation and bargaining power.

The unit of observation is a VC in a quarter. I start from quarters with at least five new

investments. The startup pool combines observed pitches and new investments, including funded startups whose pitches were not observed in that quarter.

I proxy quality with subsequent valuation growth from completed transactions. I regress log valuation growth on stage, industry, and year-month fixed effects and include the number of months between valuations as a control variable. These controls account for systematic differences in subsequent valuation growth associated with stage, industry, and timing. I use the residual as a proxy for startup quality and rank pitched and funded startups within each VC-quarter's pool, with rank one best.

I set the Pareto parameter α to 1.150 by fitting a Pareto tail to the startup investment returns reported by Levine (2014) using Correlation Ventures data. This return-based estimate provides a benchmark for startup quality.

I estimate the VC's selection ability, θ, by matching the ranks of its investments. For VC-quarter i, let N_i be the number of startups in its pool and k_i the number of those startups that the VC funds. In the data, $\bar{p}_i$ is the realized average rank of the VC's investments in that quarter, divided by N_i. For given N_i, k_i, and θ, the model predicts the average quality rank of the k_i startups the VC selects. Dividing this prediction by N_i gives $m(N_i, k_i; \alpha, \theta)$. I estimate θ by nonlinear least squares:

$$\hat{\theta} = \arg\min_{\theta \geq 0} \sum_i \left[\bar{p}_i - m(N_i, k_i; \alpha, \theta)\right]^2 .$$

The estimate makes the model's predicted investment ranks as close as possible to the observed ranks, giving each observation equal weight.

Finally, I choose c to match the observed exit rate. For VC-quarter i, let e_i be the number of its k_i investments with an observed exit. The mean of e_i/k_i across VC-quarters is $R_0 = 14.077\%$. I choose $\hat{c}$ to satisfy $R(\bar{N}, \bar{k}; \alpha, \hat{\theta}, \hat{c}) = R_0$, where $\bar{N}$ and $\bar{k}$ are the average pool size and investment count.

To assess the magnitude, I put the fitted parameters into the expected success rate:

$$R(N, k; \alpha, \hat{\theta}, \hat{c}) = E\left[\frac{1}{k} \sum_{j \in A_N(\hat{\theta})} \min\{\hat{c}Q_j, 1\}\right] .$$

Here, $A_N(\hat{\theta})$ contains the k startups with the highest signals under the estimated selection ability. Each term $\min\{\hat{c}Q_j, 1\}$ is a selected startup's success probability, so the expression inside the expectation is their average. The expectation averages over startup qualities drawn from the Pareto distribution with parameter α and over signal errors. Appendix A.2.3 gives the numerical calculation.

6.3. Magnitude and interpretation

I expand the pool using the response estimated in Table 2, column 1. Its coefficient of 70.180 implies that a one-standard-deviation increase in log involvement brings about 69 additional pitches. Relative to the sample mean of 96.365 pitches reported in Table 1, this is an increase of $g = 71.777\%$. I apply this proportional increase to the startup pool of the top 10% AUM sample, holding the investment count, selection ability, quality distribution, and c fixed. A strong assumption here is that the pitch response in the main sample also applies to large VCs and to the startup pool constructed for the calibration.

Evaluating the expected exit rate at the larger pool size gives

$$R_1 = R((1+g)\bar{N}, \bar{k}; \alpha, \hat{\theta}, \hat{c}) = 15.208\%.$$

For the same increase in involvement, Table 6 implies an exit-rate gain of 3.572 percentage points. The model's gain is therefore about 32% as large.[9] Table 7 reports the inputs and the comparison.

The precise fraction depends on the assumptions and data used in the calibration, with alternatives reported in Appendix A.3. Nevertheless, the exercise suggests that deal flow can make an economically meaningful contribution. Choosing the same number of investments from a larger pool generates a gain roughly one third as large as the exit-rate increase associated with involvement in Table 6, holding average quality and selection ability fixed. Involvement may also improve investment outcomes through VC learning, helping VCs evaluate startups more accurately and increasing selection ability, θ.

7. Robustness

7.1. Specification and sample robustness

Each paragraph below names a threat to the design, the test that confronts it, and the result. Tables IA.3 through IA.9 report the details, and all but Table IA.4 repeat the baseline of 70 new pitches in their first column. The main results survive in all of the robustness checks.

The first concern, about the exclusion restriction, comes from Section 4: an industry shock, such as a biotech boom, can move people and founders between the pair's two cities

[9] The model's exit rate rises from 14.077% to 15.208%, a gain of 1.131 percentage points. Dividing this gain by Table 6's gain of 3.572 percentage points gives $1.131/3.572 \times 100 \approx 32\%$.

at once, and neither city's own fixed effects absorb it. To address this concern, I add industry-by-quarter fixed effects, defined by the focal portfolio company's primary industry in PitchBook. Table IA.3 reports the estimates with the seven industry sectors and with the forty finer industry groups. The coefficient moves from 70.2 to 68.9 and 68.1, within 3% of the baseline, remains significant at the 1% level, and the first stage is as strong as before.

A second concern, also from Section 4, is a shock at the VC's own building, such as a burst of attention after an extremely successful exit: it draws people to the VC's building, some of whom also pass near its portfolio companies, and it draws pitches. Such a shock has a signature: because the crowd is at the VC's building, it raises passersby for all of the VC's pairs in that quarter, not only for the pair in question. I therefore compute, for each VC, portfolio company, and quarter, the average of Log (Passersby) between the VC and each of its other portfolio companies in the same quarter, and I add it as a control in both stages. If a shock at the VC's building were driving the results, the average would predict the pair's involvement, and the estimate would move substantially once the average is controlled for. Table IA.4 shows neither: the average has no explanatory power for the pair's involvement, and with the control the second-stage coefficient is 62, still significant at the 1% level. The results therefore mainly come from passersby on the pair's own route, as the instrument intends, and not from a crowd at the VC's building.

A third concern is physical proximity. When the two buildings are very close, a device near one building is also near the other, so recorded visits and passersby are more likely to be false positives. Table IA.5 drops all pairs closer than 5 km and then 10 km. The estimate does not shrink. It grows, to 92 and 105 respectively. The effect does not come from nearby pairs.

A fourth concern is the pandemic. The sample spans 2018 to 2023, and COVID moved foot traffic and startup activity together. Restricting to focal quarters before 2020Q2 (Table IA.6, column 2) yields an estimate of 97, larger than the baseline, on less than a third of the sample. The effect is not an artifact of pandemic-era co-movement.

A fifth concern is building attribution. When a building houses several tenants, visits and passersby can be assigned to the wrong one. Column 3 of Table IA.6 keeps only portfolio companies whose address is shared with no other portfolio company in the sample. The estimate is 59, of similar magnitude to the baseline, so attribution noise in shared buildings does not appear to generate the result. Table IA.7, column 2 makes the related restriction to headquarters buildings only, dropping branch locations on both sides. The estimate is 78 with a tighter standard error.

A sixth concern is the construction of the involvement measure itself. Table IA.8 swaps the baseline for two alternatives. Per-visitor minutes divides the month's total minutes by the number of distinct devices that visited. Daily-peak minutes keeps only the single longest visit on each day and adds these daily peaks over the month. The estimates are 78 and 71, within about 12% of the baseline. Table IA.9 varies the two construction choices: the focal-window length (6, 12, and 18, versus 24 months) and the visit-duration filter (10–300, 60–300, and 30–600, versus 30–300 minutes). Estimates range from 53 to 78, always significant at the 1% level.

A final concern is mechanical overlap between treatment and outcome, since both are built from visits. By construction the two sides use different pairs: the involvement measure uses the focal pair's own visits, while the outcome counts first visits from other startups. Table IA.7, column 3 imposes full separation anyway. Involvement is measured only from visits by the VC's employees to the portfolio company's building, and pitches only from visits by startup employees to the VC's building, so no visit direction enters both sides of the regression. The estimate roughly doubles, to 139, consistent with each side being measured with less noise once the directions are unmixed. Under every variation in Tables IA.3 through IA.9, the conclusion of Section 5.1 stands.

7.2. Meetings at coffee shops

A further concern is venue. The pitch measure counts first visits at the pair's own buildings, but early conversations famously happen over coffee. Suster (2011) tells founders raising money to take fifty coffee meetings, and gives VCs the same advice: "Get out of your offices and go have coffee meetings. Preferably at startup HQs. Why do they always need to come to you? Increase your deal flow." At Palo Alto's Coupa Café, founders buy lattes before their pitch sessions with VCs (Mitchell, 2017). A first meeting held at a coffee shop near either office therefore goes uncounted by the office-based measure. To address this concern, I rebuild the pitch measure to include meetings at coffee shops.

I identify coffee shops from SafeGraph points-of-interest data, keeping places with NAICS code 722515 whose category tags contain Cafe, Coffee, Hong Kong Style Cafe, Internet Cafe, or Tea House. Each VC and each startup is assigned the coffee shop nearest its headquarters, if one exists within one mile. The assignment follows how people use coffee shops: for a meeting near the office, the natural choice is the closest one. On this rule, 97.5% of VCs and 96.0% of startups in the estimation sample have an assigned shop, and 96.5% of focal deal-quarter observations have both sides assigned.

A coffee meeting is a same-day co-location of a VC-side device and a startup-side device at an assigned shop, with overlapping stays. I count a meeting only at a shop assigned to exactly one side of the pair, so one side is at its own shop and the other has traveled to it. If the shop is assigned to both sides, the two offices are neighbors, and a co-location there likely reflects the shared neighborhood rather than a meeting. If the shop is assigned to neither, a co-location there cannot be tied to the pair. Coffee meetings then enter the pitch panel under the same rules as office visits. Meetings on or after the pair's first deal month are excluded. Each pair enters once, at the quarter of its first meeting, taking the earlier of the office-based and coffee-based dates. The preferred specification requires the daily overlap to last 30 to 300 minutes, the same duration window the office measure uses. Appendix B details the construction.

The augmentation is large. It enlarges the pitch panel by 43% and moves the first pitch to an earlier quarter for 14% of the pairs the office measure records. The estimates barely move. At the one-quarter horizon the coefficient is 69.8, against the office-based baseline of 70.2 (Table IA.10). Looser meeting definitions, any overlapping stay or a shared clock hour, give 69.8 and 72.5. At the two- and three-quarter horizons the coefficients attenuate somewhat and remain significant. The first-stage F statistic and the sample are identical across columns because only the outcome changes.

In this exercise, only the deal flow measure adds the coffee-shop data. Pitching is pre-deal first contact between strangers, and it plausibly happens at a third venue. Involvement is post-investment support, more likely to happen where the work is: at the portfolio company's office or the VC's. I therefore rebuild the pitch measure and leave involvement and the instrument as constructed in Sections 2 and 4. The office-based measure remains the baseline because each side's own building identifies its people cleanly, while a public shop is a shared venue where co-location is a noisier signal of a meeting and may produce more false positives. Still, the exercise serves as a meaningful robustness check.

8. Conclusion

Does a VC's involvement with its portfolio companies bring it more deal flow? The question has been hard to answer because neither involvement nor deal flow is observed directly. I measure both from smartphone location data. Involvement is the time a VC's employees and a portfolio company's employees spend at each other's offices in a month. A pitch is the first visit between a VC and a startup with no prior deal. Using passersby between the two offices as an instrument for involvement, I find that involvement generates deal flow. A

one-standard-deviation increase in involvement brings the VC about 70 more new pitches in the next quarter, a 72% increase over the mean.

The new pitches are consistent with word of mouth spreading from the portfolio company. The new pitchers concentrate in the portfolio company's industry, not in its state. More of them are raising money for the first time, the startups that need hands-on help most. The pattern is consistent with these startups seeking out the VC rather than the reverse.

Involvement increases deal flow, but it does not raise the quality of the startups that pitch: the eventual exit rate of the new pitchers is largely unchanged. If involvement had raised the quality of pitches, better investment outcomes would be no surprise. Instead, the larger choice set is a potential source of the improvement in the VC's investment outcomes. It does not invest more often, but its new investments exit through an IPO or M&A 3.6 percentage points more often.

The model explains why. The best of many pitches is better than the best of few, even when the average quality is the same. The model predicts a gain from larger deal flow alone that is about 32% of the estimated improvement in investment outcomes.

The results have two implications. First, the benefit of involvement does not stop at the portfolio company. Part of it returns to the VC as deal flow, and this return may explain why VCs keep spending most of their meeting time on portfolio companies that contribute little to their profits. Second, deal flow is something a VC can build. Involvement is a choice, and the choice brings more pitches of the same quality and better investments.

References

Rustam Abuzov. Busy venture capitalists and investment performance. *Journal of Financial and Quantitative Analysis*, 60(6):2753–2787, 2025.

Vladimir Atanasov, Vladimir Ivanov, and Kate Litvak. Does reputation limit opportunistic behavior in the VC industry? Evidence from litigation against VCs. *Journal of Finance*, 67(6):2215–2246, 2012.

David Atkin, M. Keith Chen, and Anton Popov. The returns to face-to-face interactions: Knowledge spillovers in Silicon Valley. Working Paper 30147, National Bureau of Economic Research, 2022.

Onur Bayar, Thomas J. Chemmanur, and Xuan Tian. Peer monitoring, syndication, and the dynamics of venture capital interactions: Theory and evidence. *Journal of Financial and Quantitative Analysis*, 55(6):1875–1914, 2020.

Benedict Evans. In praise of failure. Blog post, ben-evans.com, August 10, 2016. URL `https://www.ben-evans.com/benedictevans/2016/4/28/winning-and-losing`.

Shai Bernstein, Xavier Giroud, and Richard R. Townsend. The impact of venture capital monitoring. *Journal of Finance*, 71(4):1591–1622, 2016.

Laura Bottazzi, Marco Da Rin, and Thomas Hellmann. Who are the active investors? Evidence from venture capital. *Journal of Financial Economics*, 89(3):488–512, 2008.

Jun Chen and Michael Ewens. Venture capital and startup agglomeration. *Journal of Finance*, 80(4):2153–2198, 2025.

Lin William Cong and Yizhou Xiao. Persistent blessings of luck: Theory and an application to venture capital. *Review of Financial Studies*, 35(3):1183–1221, 2022.

Lisa D. Cook, Matt Marx, and Emmanuel Yimfor. Funding black high-growth startups. *Journal of Finance*, 81(3):1619–1660, 2026.

Mehran Ebrahimian and Ye Zhang. Matching and bargaining in entrepreneurial finance: Evidence from an experimental system. SSRN Working Paper No. 3724424, 2025.

Ofer Eldar and Jillian Grennan. Common venture capital investors and startup growth. *Review of Financial Studies*, 37(2):549–590, 2024.

Michael Ewens and Nadya Malenko. Board dynamics over the startup life cycle. *Journal of Finance*, 81(5):2543–2594, 2026.

Michael Ewens and Matt Marx. Founder replacement and startup performance. *Review of Financial Studies*, 31(4):1532–1565, 2018.

Michael Ewens and Denis Sosyura. Irreplaceable venture capitalists, 2023. Working paper.

Michael Ewens, Ramana Nanda, and Matthew Rhodes-Kropf. Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128(3):422–442, 2018.

Xiaoyong Fu and Lucian A. Taylor. Due diligence and the allocation of venture capital. Working Paper 33987, National Bureau of Economic Research, 2026.

Paolo Fulghieri and Merih Sevilir. Size and focus of a venture capitalist's portfolio. *Review of Financial Studies*, 22(11):4643–4680, 2009.

Jon A. Garfinkel, Erik J. Mayer, Ilya A. Strebulaev, and Emmanuel Yimfor. Alumni networks in venture capital financing. *Journal of Financial and Quantitative Analysis*, pages 1–31, 2025.

William C. Gerken, Steven Irlbeck, Marcus Painter, and Guangli Zhang. Watching the watchdogs: The information asymmetry of SEC interactions. SSRN Working Paper No. 4941708, 2026.

Xavier Giroud. Proximity and investment: Evidence from plant-level data. *Quarterly Journal of Economics*, 128(2):861–915, 2013.

Paul A. Gompers, Will Gornall, Steven N. Kaplan, and Ilya A. Strebulaev. How do venture capitalists make decisions? *Journal of Financial Economics*, 135(1):169–190, 2020.

Juanita González-Uribe, Robyn Klingler-Vidra, Su Wang, and Xiang Yin. The broader role of venture capital due diligence. *Review of Financial Studies*, 39(7):2018–2063, 2026.

Michael Gorman and William A. Sahlman. What do venture capitalists do? *Journal of Business Venturing*, 4(4):231–248, 1989.

Thomas Hellmann and Manju Puri. Venture capital and the professionalization of start-up firms: Empirical evidence. *Journal of Finance*, 57(1):169–197, 2002.

Yael V. Hochberg, Alexander Ljungqvist, and Yang Lu. Whom you know matters: Venture capital networks and investment performance. *Journal of Finance*, 62(1):251–301, 2007.

Sabrina T. Howell. Reducing information frictions in venture capital: The role of new venture competitions. *Journal of Financial Economics*, 136(3):676–694, 2020.

Sabrina T. Howell and Ramana Nanda. Networking frictions in venture capital, and the gender gap in entrepreneurship. *Journal of Financial and Quantitative Analysis*, 59(6): 2733–2761, 2024.

David H. Hsu. What do entrepreneurs pay for venture capital affiliation? *Journal of Finance*, 59(4):1805–1844, 2004.

Blake Jackson and Ilya A. Strebulaev. Human capital in venture capital: Evidence from 100,000 venture capitalists. Working Paper 35501, National Bureau of Economic Research, 2026.

Steven N. Kaplan and Antoinette Schoar. Private equity performance: Returns, persistence, and capital flows. *Journal of Finance*, 60(4):1791–1823, 2005.

Naveen Khanna and Richmond D. Mathews. Skill versus reliability in venture capital. *Journal of Financial Economics*, 145:41–63, 2022.

Kindred Capital. The venture capital selection process: The numbers. Medium post, 2019. URL https://kindredcapital.medium.com/the-venture-capital-selection-process-the-numbers-dc4a4e13b2eb.

David Kirsch, Brent Goldfarb, and Azi Gera. Form or substance: The role of business plans in venture capital decision making. *Strategic Management Journal*, 30:487–515, 2009.

Robin Y. Lee. Face-to-face social interactions and local informational advantage. *Journal of Financial Economics*, 180:104280, 2026.

Josh Lerner. Venture capitalists and the oversight of private firms. *Journal of Finance*, 50 (1):301–318, 1995.

Seth Levine. Venture outcomes are even more skewed than you think. VC Adventure (blog), August 12, 2014. URL https://sethlevine.com/archives/2014/08/venture-outcomes-are-even-more-skewed-than-you-think.html.

JP Mangalindan. TheFunded: Still the Valley's anonymous gossip shop. Fortune, March 24, 2011. URL https://fortune.com/2011/03/24/thefunded-still-the-valleys-anonymous-gossip-shop/.

Robert Marquez, Vikram Nanda, and M. Deniz Yavuz. Private equity fund returns and performance persistence. *Review of Finance*, 19(5):1783–1823, 2015.

Deana Mitchell. Silicon Valley's hot café: Where digirati pitch ideas over Venezuelan coffee. VOA News, July 29, 2017. URL https://www.voanews.com/a/silicon-valley-coupa-cafe/3964021.html.

Yuhei Miyauchi, Kentaro Nakajima, and Stephen J. Redding. The economics of spatial mobility: Theory and evidence using smartphone data. *Quarterly Journal of Economics*, 140(4):2507–2570, 2025.

Ramana Nanda, Sampsa Samila, and Olav Sorenson. The persistent effect of initial success: Evidence from venture capital. *Journal of Financial Economics*, 137(1):231–248, 2020.

Abraham Othman. Startup growth and venture returns. AngelList research paper, 2019. URL https://angel.co/pdf/growth.pdf.

Satya Patel. Homebrew's 1%: The VC metrics behind investing in one of every 100 companies we meet. Venture Generated Content (blog), 2014. URL https://venturegeneratedcontent.com/2014/01/09/homebrews-1-the-vc-metrics-behind-investing-in-one-of-every-100-companies-we-meet/.

Satya Patel. How we work at Homebrew. Venture Generated Content (blog), July 19, 2017. URL https://venturegeneratedcontent.com/2017/07/19/how-we-work-at-homebrew/.

Jeffrey S. Petty and Marc Gruber. "In pursuit of the real deal": A longitudinal study of VC decision making. *Journal of Business Venturing*, 26:172–188, 2011.

Jeffrey S. Petty, Marc Gruber, and Dietmar Harhoff. Maneuvering the odds: The dynamics of venture capital decision-making. *Strategic Entrepreneurship Journal*, 17:239–265, 2023.

David Politis. Startup founder survival guide: 50+ rules to live by, learned the hard way. Not Another CEO (Substack), January 24, 2024. URL https://notanotherceo.substack.com/p/startup-founder-survival-guide.

William A. Sahlman. The structure and governance of venture-capital organizations. *Journal of Financial Economics*, 27(2):473–521, 1990.

Francesco Sannino. The equilibrium size and value-added of venture capital. *Journal of Finance*, 79(2):1297–1352, 2024.

Amir Sariri. The economics of advice: Evidence from start-up mentoring. *Management Science*, 71(12):10022–10046, 2025.

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

Mark Suster. Why it's critical that you reference check your VC. Both Sides of the Table (blog), December 15, 2010. URL https://bothsidesofthetable.com/why-its-critical-that-you-reference-check-your-vc-81470cd7f14c.

Mark Suster. Why you need to take 50 coffee meetings. Both Sides of the Table (blog), August 16, 2011. URL https://bothsidesofthetable.com/why-you-need-to-take-50-coffee-meetings-835c2edf85b9.

Main figures

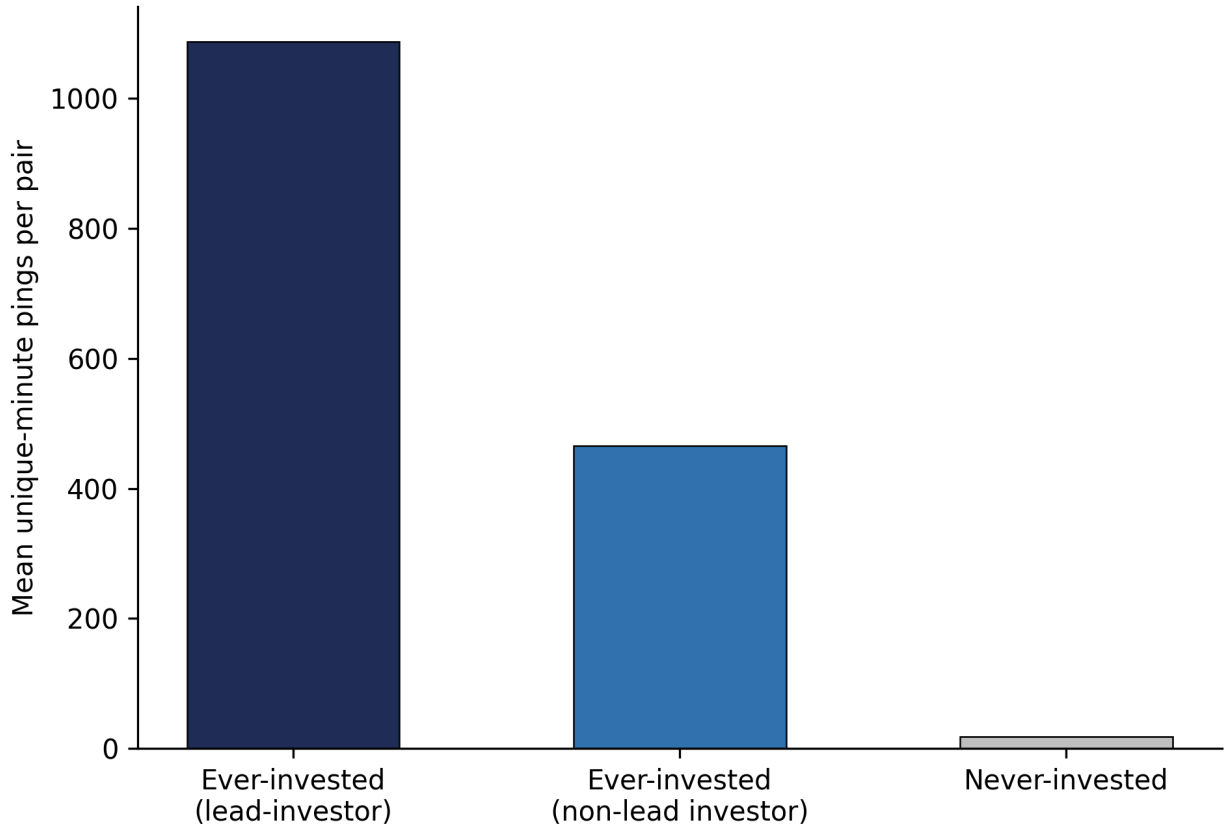

Figure 1. Intensity of smartphone signals across all VC-startup pairs. This figure plots the mean signal intensity per VC-startup pair for three mutually exclusive groups: pairs in which the VC ever invested in the startup as the lead VC on at least one deal, pairs in which the VC ever invested but never as the lead, and pairs in which the VC never invested. Investment relations are taken from PitchBook records as of February 2023 and include deals of any date. The universe is every combination of a VC and a startup in the PitchBook deal sample, and the same smartphone data purchase covers all of their buildings, so differences across groups reflect behavior rather than data availability. Pairs with no recorded pings count as zero. Signal intensity is the total number of unique ping minutes recorded by either side's devices at the other side's building over January 2018 to January 2023, with no visit filters, and regardless of whether or when the pair has a deal.

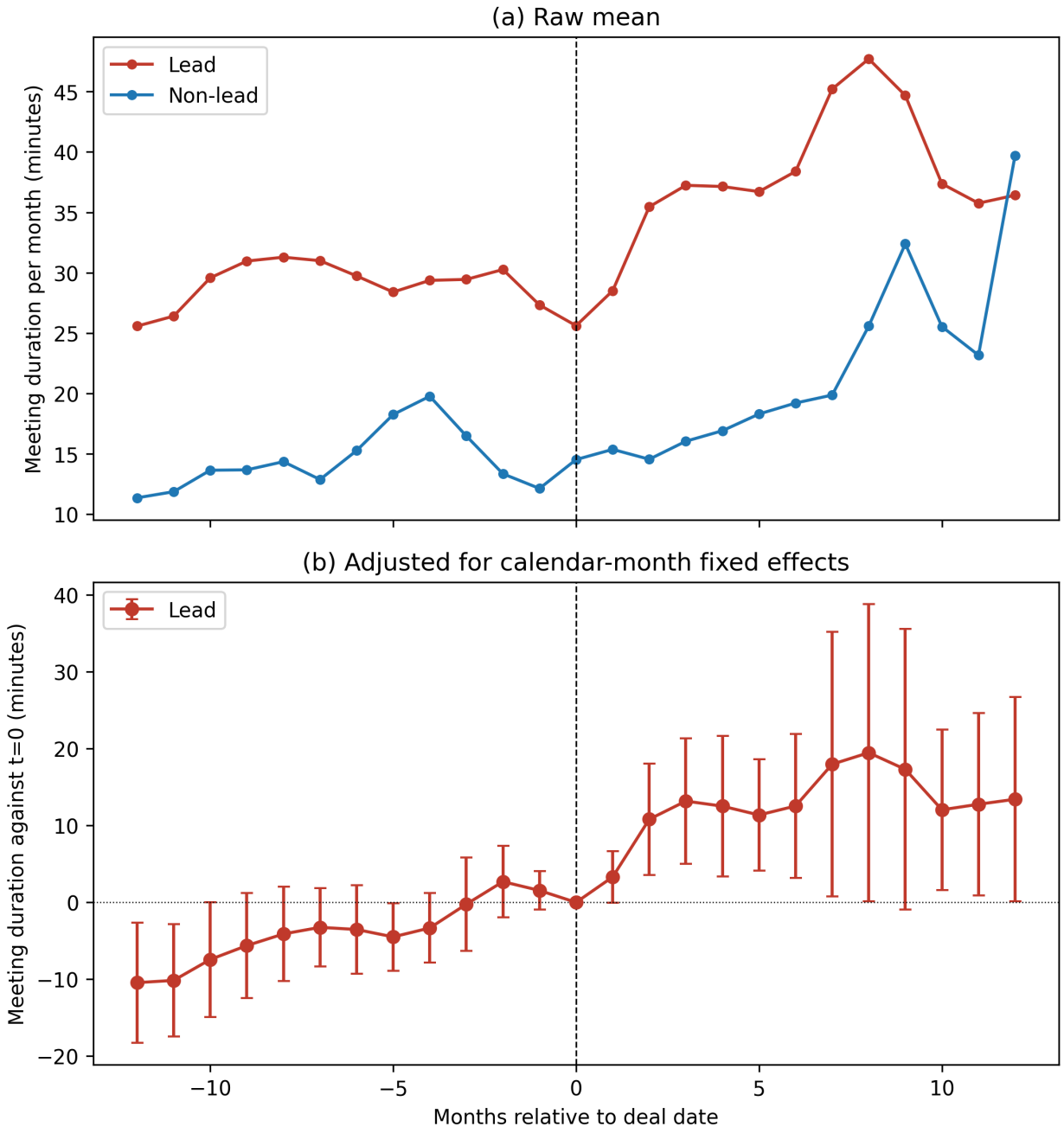

Figure 2. VC involvement around the investment date. This figure plots monthly meeting minutes between the VC and the startup, by month relative to the pair's first investment date, for pairs whose first deal falls between January 2019 and December 2021, so that each pair has 12 months before and 12 months after the investment date within the smartphone coverage. Panel (a) plots raw means, separately for lead and non-lead VCs. Panel (b) plots coefficients from a regression of lead-VC meeting minutes on indicators for months relative to the investment date and calendar-month fixed effects. Error bars show 95% confidence intervals based on standard errors clustered by VC.

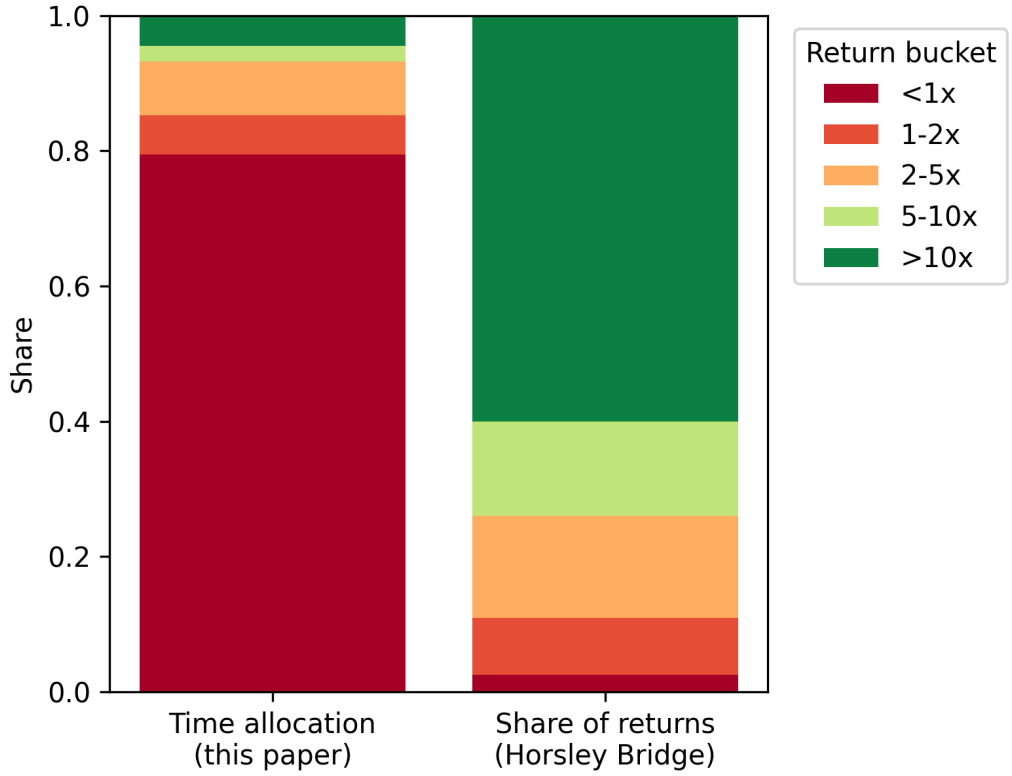

Figure 3. VC time allocation versus return concentration. This figure sorts lead-VC deals into five groups by their eventual return multiple, below 1x, 1–2x, 2–5x, 5–10x, and above 10x, and shows two shares side by side: the share of lead-VC meeting time that goes to each group, and the share of total venture returns that each group produces. For each focal deal, I take the portfolio company's first financing round dated at least 24 months after the focal deal and compute the return multiple as that round's valuation divided by the focal-deal valuation. A fixed minimum horizon makes multiples comparable across deals. Deals with no round at least 24 months later get a multiple of zero. Meeting time is summed over each deal's focal window. PitchBook valuations allow sorting the sample deals into return groups but not measuring the returns each group produces, so the returns distribution comes from an outside benchmark: Horsley Bridge aggregate data on more than 7,000 investments made by funds in which it invested between 1985 and 2014, as reported by Benedict Evans (2016). The two bars therefore describe different sets of investments. Segment values are digitized from the published chart.

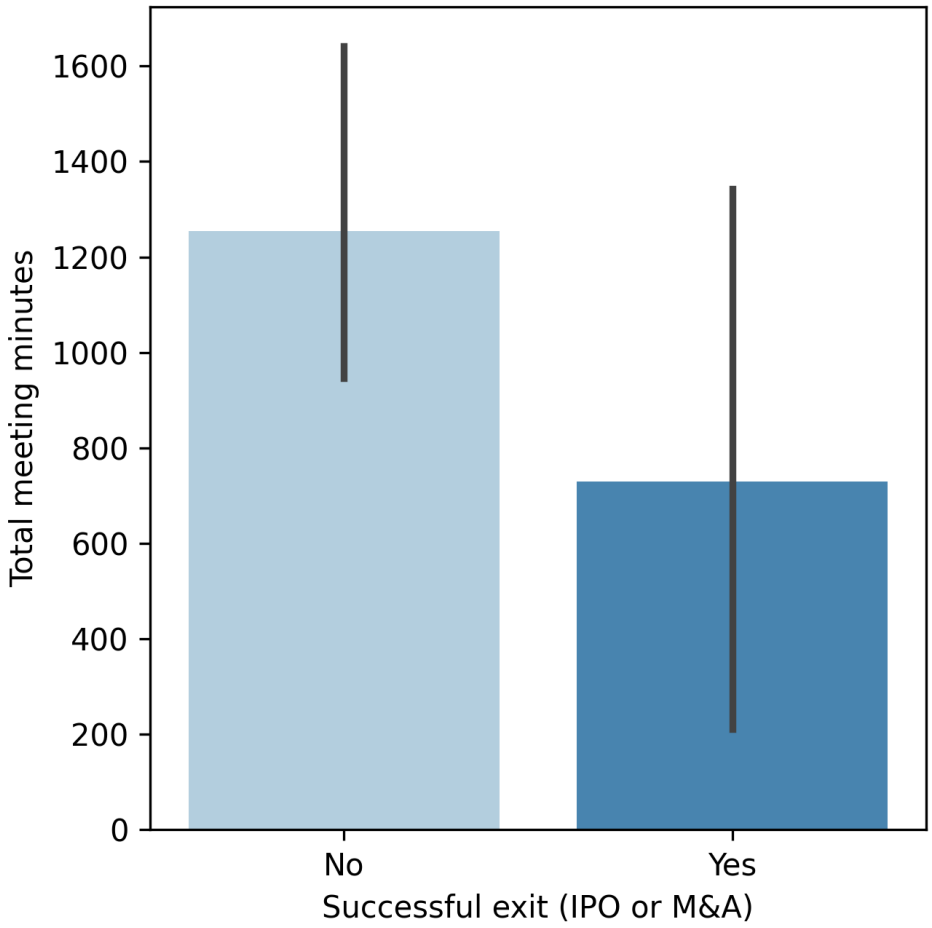

Figure 4. VC involvement by portfolio company exit outcome. This figure plots mean total meeting minutes over the focal window for two groups of focal deals: portfolio companies that never exit through an IPO or M&A, and those that eventually do. The sample keeps only focal deals in which PitchBook lists the VC as a lead investor. Error bars show bootstrap 95% confidence intervals.

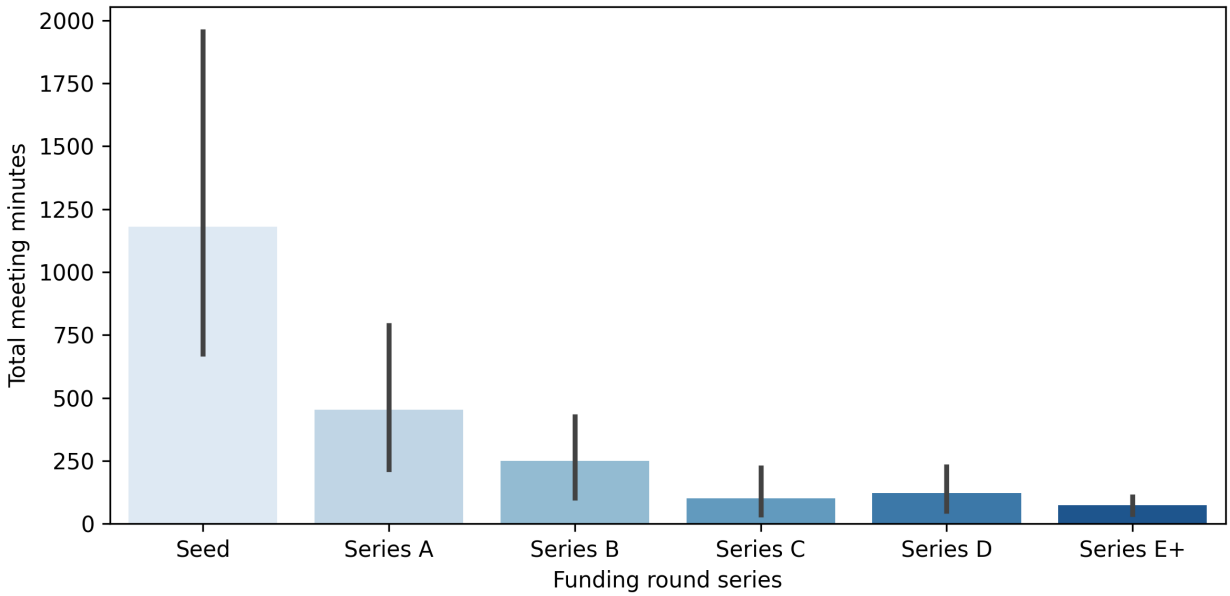

Figure 5. VC involvement by deal stage. This figure groups the sample deals by funding-round type, from Seed to Series E and later, and plots, for each group, the mean total meeting minutes over the up to 24 months after the investment date. The sample keeps only deals in which PitchBook lists the VC as a lead investor. Series E and later rounds are grouped as Series E+. Error bars show bootstrap 95% confidence intervals.

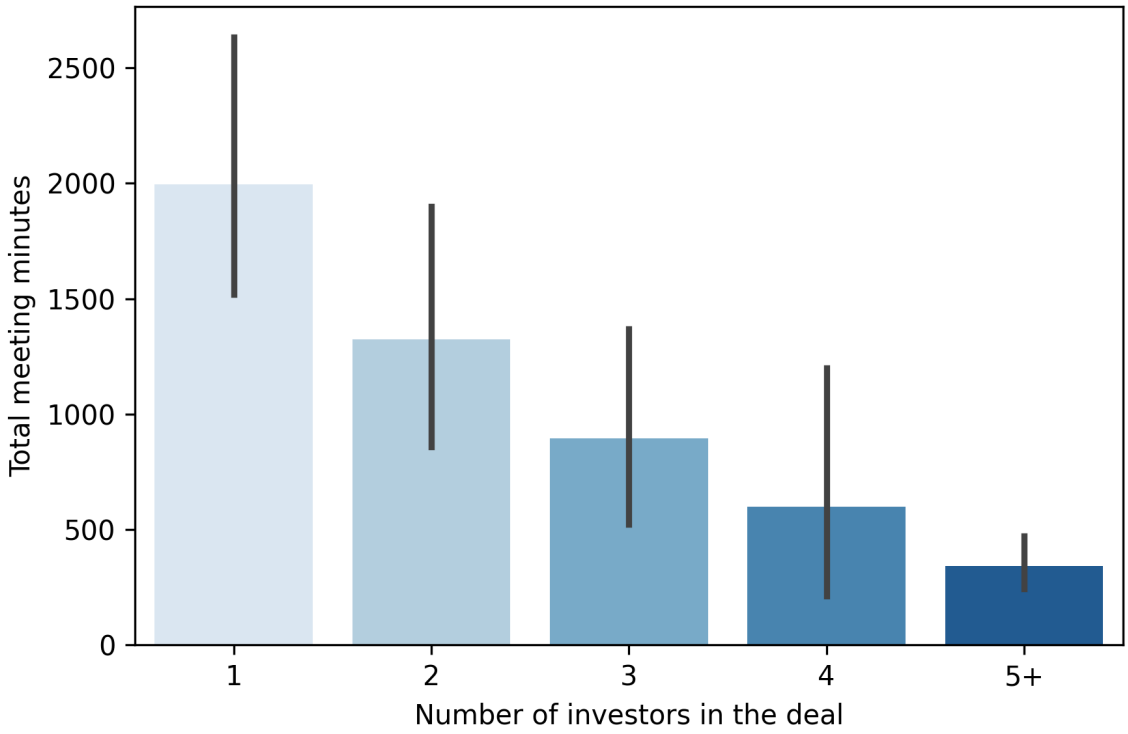

Figure 6. Per-VC involvement by syndicate size. This figure groups the sample deals by the number of VCs in the deal, from one to five or more, and plots, for each group, the mean total meeting minutes per VC over the up to 24 months after the investment date. The unit of observation is VC by deal. Error bars show bootstrap 95% confidence intervals.

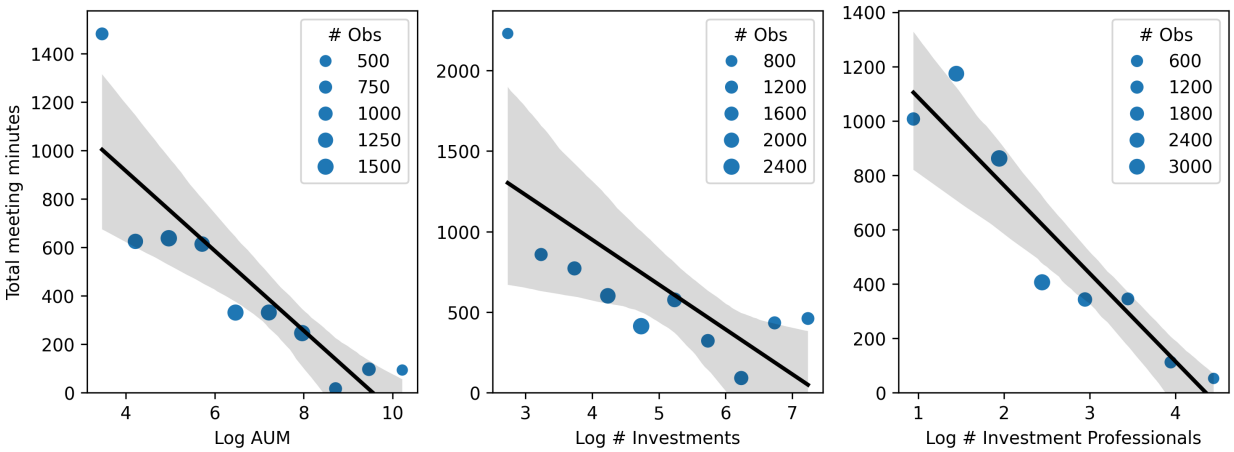

Figure 7. VC involvement by investor size. This figure plots binned means of total meeting minutes over the up to 24 months after the investment date against three measures of the lead VC's size: log assets under management, log number of investments, and log number of investment professionals. The sample keeps only deals in which PitchBook lists the VC as a lead investor and classifies it as a venture capital investor. Each x-variable is trimmed at its 5th and 95th percentiles. Marker size is proportional to the number of deals in the bin, and the line is a linear fit with its 95% confidence band.

Main tables

Table 1. Summary statistics. This table reports summary statistics for the estimation sample. The unit of observation is VC by deal by quarter: a focal deal is one deal record pairing one VC with one portfolio company, observed over each calendar quarter of its focal window. For each PitchBook deal in the sample, I define the focal window to run up to 24 months after the investment date, ending early if the portfolio company raises a qualifying next round. Each row reports the number of observations, the mean, and the standard deviation. Panel A reports the independent variable: Log (Involvement) is the log of one plus the pair's total visit minutes per month, averaged within the quarter. Visits count both directions, a VC employee visiting the portfolio company's building and a portfolio company employee visiting the VC's building. Panel B reports the instrumental variable: Log (Passersby) is the log of one plus the number of distinct devices that appear near both the VC's and the portfolio company's buildings within the same month, excluding devices identified as likely employees of any VC or startup in the sample, averaged within the quarter. Panel C reports outcome variables in raw levels, all measured one quarter ahead. Number of new pitches counts startups whose first observed visit with the VC occurs in that quarter, including both startups the VC never invests in and startups it invests in only later. Share first-time funded is the share of those new pitchers with no funding round recorded in PitchBook before the focal quarter. Pitcher IPO/M&A rate is the share of new pitchers that eventually exit through an IPO or M&A. Number of new investments counts companies whose first-ever investment from the VC occurs in that quarter, Number of new investments with IPO/M&A counts those with an IPO or M&A exit recorded in PitchBook, and Rate of new investments with IPO/M&A is their ratio, defined only when at least one new investment occurs. Observation counts are smaller for the pitcher outcomes and the exit rate because those variables are undefined in quarters with no new pitchers or no new investments.

	Observations	Mean	SD
Panel A. Independent variable			
Log (Involvement)	661,259	0.189	0.986
Panel B. Instrumental variable			
Log (Passersby)	661,259	1.382	2.104
Panel C. Outcome variables			
Number of new pitches	661,252	96.365	200.300
Share first-time funded	627,858	0.222	0.197
Pitcher IPO/M&A rate	627,858	0.077	0.079
Number of new investments	661,252	5.775	16.328
Number of new investments with IPO/M&A	661,252	0.745	1.925
Rate of new investments with IPO/M&A	464,371	0.155	0.249

Table 2. Involvement and the number of new pitches. This table reports second-stage results from two-stage least squares (2SLS) estimation. The unit of observation is VC by deal by quarter. The independent variable Log (Involvement) is the log of one plus the pair's total visit minutes per month, averaged within the quarter. The instrumental variable is Log (Passersby), the log of one plus the number of distinct devices that appear near both the VC's and the portfolio company's buildings within the same month, excluding devices identified as likely employees of any VC or startup in the sample, averaged within the quarter. The outcome in column (h) is the number of new pitches h quarters ahead, for $h = 1, 2, 3$: startups whose first observed visit with the VC occurs in that quarter, including both startups the VC never invests in and startups it invests in only later. All columns absorb pair, startup city by quarter, and VC city by quarter fixed effects. KP F-stat is the Kleibergen–Paap rk Wald F statistic. Standard errors in parentheses are clustered by VC. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)	New pitches (t+2)	New pitches (t+3)
	(1)	(2)	(3)
Log (Involvement)	70.180***	24.817***	19.562***
	(10.286)	(7.865)	(5.025)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	479.3	439.9
Observations	645,990	602,623	555,636

Table 3. How new pitchers are connected to portfolio companies. This table decomposes the outcome of Table 2, column (1), into eight mutually exclusive groups of new pitchers, defined by the pitcher's links to the focal portfolio company. An industry link means the pitching startup and the portfolio company share the same PitchBook primary industry sector; an alumni link means their CEOs attended at least one common educational institution in PitchBook's biography records; a state link means they are headquartered in the same US state. Each column counts the new pitchers one quarter ahead in one group. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	Number of new pitches							
Link type:	Industry only	Alumni only	State only	Industry + alumni	Industry + state	Alumni + state	All three links	No observed link
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log (Involvement)	18.967***	0.454***	0.151	0.147***	2.989**	-0.005	0.024	47.454***
	(2.291)	(0.067)	(2.665)	(0.027)	(1.284)	(0.084)	(0.043)	(5.974)
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KP F-stat	454.8	454.8	454.8	454.8	454.8	454.8	454.8	454.8
Observations	645,990	645,990	645,990	645,990	645,990	645,990	645,990	645,990

Table 4. Involvement and the experience of new pitchers. This table reports second-stage results from 2SLS estimation, where the outcome variable is the share of new pitchers that are raising funding for the first time. New pitchers are the startups whose first observed visit with the VC occurs h quarters ahead, and the three columns report $h = 1, 2, 3$. First-time funded means the pitcher has no funding round recorded in PitchBook before the focal quarter. If a VC by deal by quarter observation has no new pitchers, that observation is excluded from the regression. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	Share first-time funded (t+1)	Share first-time funded (t+2)	Share first-time funded (t+3)
	(1)	(2)	(3)
Log (Involvement)	0.003	0.043***	0.028***
	(0.006)	(0.008)	(0.009)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	446.4	447.5	419.6
Observations	612,165	569,274	523,540

Table 5. Involvement and the quality of new pitches. This table reports second-stage results from 2SLS estimation, where the outcome variable is the share of new pitchers that eventually exit through an IPO or M&A. New pitchers are the startups whose first observed visit with the VC occurs h quarters ahead, and the three columns report $h = 1, 2, 3$. If a VC by deal by quarter observation has no new pitchers, that observation is excluded from the regression. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	Pitcher IPO/M&A rate (t+1)	Pitcher IPO/M&A rate (t+2)	Pitcher IPO/M&A rate (t+3)
	(1)	(2)	(3)
Log (Involvement)	-0.002	0.002	-0.007
	(0.005)	(0.004)	(0.005)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	446.4	447.5	419.6
Observations	612,165	569,274	523,540

Table 6. Involvement and the VC's future investment outcomes. This table reports second-stage results from 2SLS estimation, where the outcome variables are the VC's future investment outcomes, all measured one quarter ahead. Number of new investments counts companies whose first-ever investment from the VC occurs in that quarter, Number of new investments with IPO/M&A counts those with an IPO or M&A exit recorded in PitchBook, and Rate of new investments with IPO/M&A is their ratio, defined only when at least one new investment occurs. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New investments	New investments with IPO/M&A	IPO/M&A rate
	(1)	(2)	(3)
Log (Involvement)	-0.072	0.237**	0.036*
	(0.751)	(0.097)	(0.019)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	454.8	212.2
Observations	645,990	645,990	446,841

Table 7. Calibration inputs and results. The unit of observation is a VC-quarter, and the calibration sample is restricted to the top 10% of VCs by AUM. Panel A reports the calibration inputs and the pool expansion from Tables 1 and 2. Panel B reports the observed baseline exit rate, the model's prediction after expansion, and the comparison with Table 6. pp denotes percentage points. Numerical entries are rounded to three decimal places; calculations use unrounded values.

Symbol	Meaning	Value	Source or method
Panel A. Calibration inputs			
$\bar{N}$	Pool size	92.938	Average number of startups
$\bar{k}$	Investment count	7.590	Average number of investments
α	Pareto parameter	1.150	Pareto tail fitted to returns on startup investments (Levine, 2014)
$\hat{\theta}$	Selection ability	0.038	Nonlinear least squares estimation of investment ranks
$\hat{c}$	Success scale	0.023	Solves $R(\bar{N}, \bar{k}; \alpha, \hat{\theta}, \hat{c}) = R_0$
g	Increase in pool size	71.777%	Table 2 pitch coefficient $\times$ involvement SD / mean pitches from Table 1
Panel B. Exit rates and comparison			
R_0	Baseline exit rate	14.077%	Number of exits divided by the investment count in each VC-quarter, averaged across VC-quarters
R_1	Exit rate after expansion	15.208%	$R((1+g)\bar{N}, \bar{k}; \alpha, \hat{\theta}, \hat{c})$
	Model gain	1.131 pp	$100(R_1 - R_0)$
	Table 6 gain	3.572 pp	Column 3 coefficient $\times$ the same involvement SD, converted to pp
	Gain relative to Table 6	31.673%	Model gain / Table 6 gain $\times 100$ (unrounded)

A. Model and calibration

A.1. Proof of Proposition 1

Write $W = \log Q$ and $L = \log S$. Under the model, W is exponential with rate α and $L = W + Z/\sqrt{\theta}$. Define $r(w) = \min\{ce^w, 1\}$ and $h(\ell) = E[r(W) \mid L = \ell]$. The posterior density is

$$f(w \mid \ell) = C(\ell) \exp\left\{-\alpha w - \frac{\theta}{2}(\ell - w)^2\right\} \mathbf{1}\{w \geq 0\},$$

where $C(\ell)$ normalizes the density. Differentiating the posterior expectation gives

$$h'(\ell) = \theta \operatorname{Cov}_\ell(r(W), W) > 0.$$

The differentiation is valid because r is bounded and the posterior has Gaussian tails, uniformly over compact intervals of ℓ. To establish strictness, let W' be an independent draw from the same posterior. Then

$$\operatorname{Cov}_\ell(r(W), W) = \frac{1}{2} E_\ell\left[(r(W) - r(W'))(W - W')\right].$$

The integrand is nonnegative. It is positive with positive probability because r is strictly increasing on $(0, \log(1/c))$ and the posterior density is positive there.

Independence across startups implies that $E[p(Q_j) \mid L_1, \ldots, L_N] = h(L_j)$. Thus, selecting the k highest signals maximizes the expected number of successes. This uses the model's assumption that the expected financed outcome, conditional on qualities and signal errors, is $p(Q_j)$.

Now retain the original N startups and add one independent startup. Let b_N be the kth highest original log signal and L_{new} the additional startup's log signal. Conditional on all signals, the gain in the portfolio's expected success rate is

$$\frac{h(L_{\text{new}}) - h(b_N)}{k} \mathbf{1}\{L_{\text{new}} > b_N\}.$$

The added startup does not change beliefs about the original startups. The gain is therefore nonnegative and is strictly positive when the new startup enters the selected set. Signals are continuous, so ties have probability zero. By exchangeability, the entry event has probability $k/(N+1) > 0$. Taking expectations gives

$$R(N+1, k; \alpha, \theta, c) > R(N, k; \alpha, \theta, c),$$

which proves Proposition 1. The proof does not require a finite mean of Q, because the success probability is bounded.

As θ approaches zero, rankings are equivalently generated by $\sqrt{\theta}W + Z$ and converge to random rankings. Bounded convergence then gives $R = E[p(Q)]$ in the limit, independently of pool size.

A.2. Calibration method

A.2.1. Constructing startup ranks

Sample and pool. I use VC-quarters in 2018–2022 with at least five new completed Seed, Early Stage VC, or Later Stage VC investments. The main sample is the top 10% by AUM among VCs with positive AUM in 2023.

The pool combines startups first observed pitching the VC in the quarter and its new investments, counting each startup once per VC-quarter.

Valuation growth. I use completed transactions in the September 2025 PitchBook snapshot. The implied valuation is PitchBook's *DealSize* divided by the fraction acquired (*PercentAcquired*/100). For VC rounds, I subtract the deal size to obtain the pre-money valuation. For exits, I make no subtraction.

The initial valuation is the startup's valuation as of the focal quarter, and the next valuation is the first observed afterward.[10] I use only positive valuations.

Regression adjustment. Log valuation growth is the natural log of the ratio of the next valuation to the initial valuation. I estimate the following regression, where FE denotes fixed effects:

$$\begin{aligned}\text{Log valuation growth} = \text{Stage FE} + \text{Industry FE} + \text{Year-month FE} \\ + \beta \times \text{Months between valuations} + u.\end{aligned} \tag{A.1}$$

Stage is measured at the initial valuation and classified as Seed, Early VC, Later VC, Other, or Unknown. Industry is the startup's primary industry group, with missing values assigned to Unknown. The year-month fixed effect uses the month and year of the initial valuation. The residual u measures valuation growth after accounting for these controls.

The regression uses 16,373 pairs of initial and next valuations, each with equal weight. Each startup's pair enters once, even if it appears in several VC pools. I estimate the regression before the exclusions below and use the same residuals across AUM groups.

Missing values and ranks. I exclude startups without an initial valuation and assign the lowest ranks to those without a next valuation. I also exclude funded startups with an observed IPO or acquisition after investment but no next valuation, to avoid assigning the lowest rank to a known exit. Appendix A.3 tests alternatives.

Before this exit exclusion, each VC-quarter must contain at least 20 startups with initial valuations, including at least one funded and one unfunded startup. After the exclusion,

[10] I use the latest usable valuation at or before quarter-end. For investments without a pitch observed in that quarter, I use the investment date instead.

I recompute pools, investment counts, and ranks, and require at least one funded and one unfunded startup to remain. Valuation restrictions can leave fewer than five investments, and the exit exclusion can leave fewer than 20 startups. The rank sample contains 513 VC-quarters from 72 VCs.

Within each pool, I rank residuals from highest to lowest, with rank one best, and give ties average ranks. The estimation target $\bar{p}_i$ is the average rank of the VC's retained investments divided by the number of startups in its pool. I use these observed performance ranks as a proxy for the quality ranks in the model.

A.2.2 Estimating selection ability

Tail parameter. I set $\alpha = 1.150$ by fitting a Pareto tail to the distribution of startup investment returns reported by Levine (2014) using Correlation Ventures data.[11]

Rank regression. Define $X = \alpha \log Q \sim \text{Exp}(1)$ and $T = \lambda X + Z$, where $\lambda = \sqrt{\theta}/\alpha$. Ranking T is equivalent to ranking the original signal. For fixed λ, the relation between signal ranks and quality ranks does not depend on α. I therefore estimate λ and recover $\theta = (\alpha\lambda)^2$.

Let $\mu(N, k; \lambda)$ be the expected average quality rank of the selected startups, divided by N. The main-text function is $m(N, k; \alpha, \theta) = \mu(N, k; \sqrt{\theta}/\alpha)$. I fit

$$\hat{\lambda} = \arg\min_{\lambda \geq 0} \sum_i \left[\bar{p}_i - \mu(N_i, k_i; \lambda)\right]^2. \tag{A.2}$$

Each VC-quarter receives equal weight and enters with its own N_i and k_i.

Computing the rank moment. Let ϕ and Φ denote the standard normal density and CDF. The signal CDF is

$$F_\lambda(t) = \int_0^\infty e^{-x} \Phi(t - \lambda x)\, dx, \qquad \bar{F}_\lambda(t) = 1 - F_\lambda(t).$$

Write $B_{n,b}(q) = \Pr\{\text{Binomial}(n, q) \leq b\}$, with $B_{n,b} = 0$ for $b < 0$ and $B_{n,b} = 1$ for $b \geq n$. Conditioning on a focal startup's $X = x$ and $Z = z$ gives

$$\begin{aligned}\mu(N, k; \lambda) = \frac{1}{N} + \frac{N-1}{k} \int_0^\infty \int_{-\infty}^\infty e^{-2x} \phi(z) \big[\bar{F}_\lambda(z) B_{N-2,k-2}(q) \\ + F_\lambda(z) B_{N-2,k-1}(q)\big]\, dz\, dx,\end{aligned} \tag{A.3}$$

where $q = \bar{F}_\lambda(\lambda x + z)$. The term $1/N$ counts the startup itself in its quality rank. A competitor has higher quality with probability e^{-x}. Conditional on this event, its excess

[11] Seth Levine, 2014, "Venture outcomes are even more skewed than you think," VC Adventure, August 12.

scaled log quality is again exponential, so its signal exceeds the focal signal with probability $\bar{F}_\lambda(z)$. If it does, at most $k-2$ other signals may exceed the focal signal; otherwise, at most $k-1$ may do so. These two cases give the terms in brackets.

I evaluate rank moments by numerical integration and interpolate them with a monotone cubic function. The grid for λ runs from 0 to 1, with increments of 0.005 up to 0.300 and wider spacing thereafter. I minimize (A.2) by nonlinear least squares and compare the objective across starting values and the random-selection boundary. The limits

$$\mu(N,k;0)=\frac{N+1}{2N}, \qquad \mu(N,k;\infty)=\frac{k+1}{2N}$$

provide checks. Since μ decreases with λ, the objective for $\lambda > 1$ is bounded below by $\sum_i[\max\{\bar{p}_i-\mu(N_i,k_i;1),0\}]^2$. This bound exceeds the attained minimum for every reported specification. The baseline gives $\hat{\lambda}=0.170$ and $\hat{\theta}=0.038$.

A.2.3 Computing success rates and the expansion

Expected success rate. Let x be a possible value of scaled log quality, $X=\alpha\log Q$, and let z be a possible value of the standard normal signal error, Z. A startup with these values has scaled signal $\lambda x+z$. It is selected if at most $k-1$ of the other $N-1$ signals are higher, with probability $B_{N-1,k-1}(\bar{F}_\lambda(\lambda x+z))$. The expected success rate is

$$R(N,k;\alpha,(\alpha\lambda)^2,c)=\frac{N}{k}\int_0^\infty\int_{-\infty}^\infty p(e^{x/\alpha})e^{-x}\phi(z)B_{N-1,k-1}\big(\bar{F}_\lambda(\lambda x+z)\big)\,dz\,dx. \tag{A.4}$$

The integral averages a startup's success probability times its selection probability over quality and signal noise. Multiplying by N sums across startups, and dividing by k gives the portfolio's expected success rate. I evaluate this integral numerically.

Matching the exit rate. Let e_i count the retained investments with an observed completed IPO or acquisition after investment. I use VC-quarters with exit information for all retained investments and average N_i, k_i, and e_i/k_i across these quarters. This gives $\bar{N}=92.938$, $\bar{k}=7.590$, and $R_0=14.077\%$. Each VC-quarter receives equal weight.

For noninteger N and k, I average the four adjacent integer-pool rates. Write $N=n+a$ and $k=m+b$, where n,m are integers and $a,b\in[0,1)$. Then, suppressing the parameters,

$$\begin{aligned}R(N,k)=(1-a)(1-b)R(n,m)+a(1-b)R(n+1,m)\\+(1-a)bR(n,m+1)+abR(n+1,m+1).\end{aligned} \tag{A.5}$$

This preserves both means. All component pools in the calculations satisfy $1\le k\le N$. I choose $\hat{c}$ by solving

$$R(\bar{N},\bar{k};\alpha,\hat{\theta},\hat{c})=R_0.$$

For the capped probability, the left side is continuous and strictly increasing in $c \in (0, 1)$, with limits zero and one. A one-dimensional root search gives $\hat{c} = 0.023$. I hold this value fixed when the pool expands.

Expansion and comparison. The Table 2 coefficient of 70.180 and the standard deviation of log involvement of 0.986 imply 69.168 additional pitches. Dividing by the mean of 96.365 pitches gives a pool expansion of $g = 71.777\%$. Reported values are rounded to three decimal places; calculations use unrounded inputs.

I replace $\bar{N}$ with $(1+g)\bar{N}$, leaving $\bar{k}$, α, $\hat{\theta}$, and $\hat{c}$ fixed. The expected exit rate rises from 14.077% to $R_1 = 15.208\%$, a gain of 1.131 percentage points. Table 6's coefficient is 3.624 percentage points per unit of log involvement. Multiplying it by the same standard deviation gives a gain of 3.572 percentage points. Writing this benchmark as ΔR_{IV} in probability units, the model's gain relative to it is

$$\frac{R_1 - R_0}{\Delta R_{\text{IV}}} \approx 0.31673 = 31.673\%. \tag{A.6}$$

A.3. Robustness checks

The table below varies the sample, valuation measure, quality distribution, and success probability, using the same 71.777% expansion and 3.572-percentage-point benchmark throughout. The columns broaden the AUM sample from the top 10% to the top 20% and top 25%.

Sample and valuation checks. I first reapply the 20-startup minimum after the baseline exclusion. I then change how missing follow-up valuations enter: either retain all such startups at the bottom, including funded startups with observed exits, or exclude all of them. Finally, I restrict both valuations to completed Seed, Early Stage VC, or Later Stage VC rounds, using implied pre-money values and requiring an observed follow-up. The last two specifications require at least 20 retained startups. Each change recomputes pools, investment counts, ranks, and the eligible VC-quarters, then refits λ and recalibrates c. The broad-transaction residual adjustment is held fixed when only the missing-value rule changes; venture-only values use their own adjustment.

Distribution and probability checks. The alternative $\alpha = 1.420$ is the survival-tail exponent corresponding to the density exponent 2.420 reported by Othman (2019).[12] I also replace the capped probability with $p(Q) = 1 - \exp(-cQ)$. When α changes, I hold $\hat{\lambda}$ fixed and set $\hat{\theta} = (\alpha\hat{\lambda})^2$, preserving fitted ranking behavior. I recalibrate c for each combination. Both success functions are bounded and increasing, so the argument in Appendix A.1 continues to apply.

[12] Abraham Othman, 2019, "Startup growth and venture returns," AngelList research paper.

Entries report the model's gain as a percentage of the Table 6 gain for the same one-standard-deviation increase in log involvement.

Specification	Top 10% AUM	Top 20% AUM	Top 25% AUM
Panel A. Sample and valuation			
Baseline	31.673	28.558	28.727
Reapply the 20-startup minimum	31.737	28.589	28.775
Rank all missing follow-ups last	27.220	24.769	24.176
Exclude all missing follow-ups	15.204	13.661	15.950
Use observed venture-financing pairs	15.359	14.836	16.698
Panel B. Distribution and success probability			
Alternative tail: $\alpha = 1.420$	28.135	25.259	25.415
Exponential success probability	28.571	25.758	25.935
Exponential probability and $\alpha = 1.420$	24.699	22.182	22.341

Across these specifications, the model's gain is about 14% to 32% as large as the exit-rate increase associated with involvement in Table 6. These results suggest that a larger pool can make an economically meaningful contribution across the alternative calibrations.

B. The coffee-shop meeting measure

This appendix details the construction of the augmented pitch measure used in Section 7.2.

B.1. Venue list and anchors

I identify coffee shops from SafeGraph points-of-interest data. I first keep places with NAICS code 722515, and then keep places whose category tags contain Cafe, Coffee, Hong Kong Style Cafe, Internet Cafe, or Tea House. This yields the candidate shop universe in the United States, each place with an identifier and coordinates. For every VC and every startup in the deal universe, I assign the shop nearest its headquarters address, provided the shop lies within one mile. The assignment gives one shop per VC or startup. It covers 97.5% of VCs and 93.5% of startups. The remainder have no shop within a mile and cannot generate coffee meetings, though they keep their office-based pitches. Within the estimation sample, 97.5% of VCs and 96.0% of startups have an assigned shop, and 96.5% of focal deal-quarter observations have both sides assigned. I call the assigned shop its anchor. Anchoring uses headquarters only, so branch offices do not carry anchors.

B.2. Device visits at coffee shops

The location signals are the same smartphone data as in Section 2.2, delivered for the anchored shops over January 2018 to January 2023. I keep signals from devices on the employee roster of Section 2.2 and drop devices that qualify as employees of more than ten employers in a month. For each device, shop, and local day, I form one visit from the first to the last signal of that day and keep it when the span reaches ten minutes and the device pings in at least three distinct minutes. No hour-of-day or weekday restriction applies, so evening and weekend meetings count.

B.3. Co-location and the exactly-one-anchored rule

For every shop and local day with at least two visiting devices, I classify each device's employers as anchored, when this shop is that employer's anchor, or non-anchored. Devices split into an anchored group and a visitor group, and I pair devices only across the two groups. Each pair must have overlapping stay intervals, checked pairwise. Every counted meeting therefore has exactly one side at its home venue and one side that traveled to it. This rule discards meetings where neither side is anchored to the shop, which are unattributable, and meetings where both are, which a shared amenity could explain. A device holding both anchored and non-anchored employers joins the anchored group only, so the rule undercounts rather than overcounts candidate meetings. For each employer pair I record the direction, the overlap in minutes, and whether the two devices pinged in the same clock hour. I aggregate to the shop-day-pair level, keeping the maximum overlap across device pairs.

B.4. From co-locations to first pitches

I convert co-locations to pitch events under three alternative meeting definitions. The first counts any same-day overlapping co-location. The second requires the two devices to ping in the same clock hour. The third, the preferred definition, requires the day's maximum overlap to lie between 30 and 300 minutes, the same duration window the office-visit measure uses, for the same reasons: the lower bound drops brief pass-bys, and the upper bound drops all-day co-presence that looks like a shared workspace rather than a meeting.

The pitch definition then follows Section 2.2. I drop co-locations on or after the pair's first deal month, keep quarters from 2018Q1, and collapse each pair to its earliest qualifying quarter, so each pair enters once, at the quarter of its first meeting.

B.5. The augmented panel and estimation

I merge the coffee first pitches with the office-based pitch panel at the pair level. A pair present in both sources keeps the earlier of its two first-pitch quarters. A pair present in one source keeps that source's quarter. Adding a venue can therefore only reveal a new pitching pair or move an existing pair's first pitch earlier. Everything downstream of the pitch panel is unchanged: the same outcome construction, the same estimation sample, and the same instrument, fixed effects, and clustering as Table 2.

Internet Appendix

Table IA.1. Comparison to external survey benchmarks. This table compares two statistics from the smartphone data with external benchmarks. Both statistics restrict to investors classified by PitchBook as venture capital investors. Pitches per investment is computed for each VC and year as the number of distinct startups pitching the VC divided by the number of distinct startups it invests in, and the table reports the median. Visits per month is the mean number of distinct visit days per month over the focal window, among focal deals with at least one captured visit. The Gompers et al. (2020) benchmark is the number of potential investments reaching the "met management" stage per closed deal in their survey. The Bernstein, Giroud, and Townsend (2016) benchmark is their surveyed mean number of visits per year to a portfolio company, divided by twelve.

	Pitches per investment	Visits per month
This paper (cellphone)	33.7	1.21
External benchmarks		
Gompers et al. (2020)	28	—
Bernstein, Giroud, Townsend (2016)	—	0.83

Table IA.2. First stage regression of Table 2. This table reports the first stage regression of the 2SLS estimates in Table 2. The outcome variable is Log (Involvement). The independent variable is Log (Passersby). Columns (1) to (3) use the exact estimation samples of Table 2, columns (1) to (3). The sample shrinks with the horizon because observations whose outcome quarter falls after the smartphone coverage ends are dropped. Fixed effects and clustered standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	Log (Involvement)		
	(1)	(2)	(3)
Log (Passersby)	0.092***	0.099***	0.096***
	(0.004)	(0.005)	(0.005)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	479.3	439.9
Observations	645,990	602,623	555,636

Table IA.3. Robustness test: adding industry-by-quarter fixed effects. This table re-estimates Table 2, column (1), adding industry-by-quarter fixed effects. The industry is the focal portfolio company's primary industry in PitchBook. Column (1) repeats the baseline. Column (2) adds sector-by-quarter fixed effects, using PitchBook's primary industry sector. Column (3) adds group-by-quarter fixed effects, using PitchBook's primary industry group. Specification, instrument, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)		
	(1)	(2)	(3)
Log (Involvement)	70.180***	68.914***	68.119***
	(10.286)	(10.132)	(10.184)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
Industry sector × Quarter FE	No	Yes	No
Industry group × Quarter FE	No	No	Yes
KP F-stat	454.8	456.6	452.6
Observations	645,990	645,990	645,970

Table IA.4. Robustness test: passersby at the VC's other portfolio companies. This table re-estimates Table 2, column (1), adding the passersby on the VC's other routes as a control. For each VC, deal, and quarter, Log (Passersby on other routes) is the average of Log (Passersby) between the VC and each of its other portfolio companies in the same quarter. It enters both stages as a control. The sample keeps VC-quarters with at least two portfolio companies. Panel A reports the first stage and Panel B the 2SLS estimate. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. First stage. Dependent variable: Log (Involvement)	
Log (Passersby)	0.088***
	(0.005)
Log (Passersby on other routes)	0.000
	(0.003)
Panel B. 2SLS. Dependent variable: New pitches (t+1)	
Log (Involvement)	61.965***
	(10.869)
Pair FE	Yes
Startup city × Quarter FE	Yes
VC city × Quarter FE	Yes
KP F-stat	373.7
Observations	580,308

Table IA.5. Robustness test: excluding pairs with nearby buildings. This table re-estimates Table 2, column (1), on subsamples restricted by the distance between the VC's and the portfolio company's buildings. Column (1) repeats the baseline. Columns (2) and (3) keep only pairs whose buildings are farther apart than 5 km and 10 km. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)		
Sample:	Baseline	Distance > 5 km	Distance > 10 km
	(1)	(2)	(3)
Log (Involvement)	70.180***	91.973***	104.909***
	(10.286)	(15.356)	(18.674)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	257.9	187.2
Observations	645,990	547,757	520,299

Table IA.6. Robustness test: pre-COVID and unshared-address subsamples. This table re-estimates Table 2, column (1), on two subsamples. Column (1) repeats the baseline. Column (2) keeps focal quarters strictly before 2020Q2. Column (3) keeps only portfolio companies whose address is shared by no other portfolio company in the sample. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)		
Sample:	Baseline	Pre-COVID	Unshared address
	(1)	(2)	(3)
Log (Involvement)	70.180***	97.238**	59.049***
	(10.286)	(39.305)	(8.690)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	269.4	328.7
Observations	645,990	198,131	296,070

Table IA.7. Robustness test: headquarters-only and one-direction restrictions. This table re-estimates Table 2, column (1), restricting how visits map to the pair. Column (1) repeats the baseline, which includes headquarters and branch buildings and both visit directions. Column (2) keeps only visits to headquarters buildings, for both the involvement measure and the pitch outcome. Column (3) measures involvement only from visits by the VC's employees to the portfolio company's building, and pitches only from visits by startup employees to the VC's building. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)		
Construction:	Baseline	Headquarters only	One direction
	(1)	(2)	(3)
Log (Involvement)	70.180***	77.988***	138.990***
	(10.286)	(7.527)	(19.500)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	526.7	226.5
Observations	645,990	645,990	645,990

Table IA.8. Robustness test: alternative involvement measures. This table re-estimates Table 2, column (1), replacing the involvement measure. Total minutes, the baseline, keeps each device's longest visit on each day, and adds these minutes across all devices and all days in the month. Per-visitor minutes divides the month's total minutes by the number of distinct devices that visited. Daily-peak minutes keeps only the single longest visit on each day, by any device, and adds these daily peaks over the month. Each measure is averaged over the quarter's months and enters as the log of one plus its value. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)		
Involvement measure:	Total minutes (baseline)	Per-visitor minutes	Daily-peak minutes
	(1)	(2)	(3)
Log (Involvement)	70.180***	78.285***	71.127***
	(10.286)	(11.508)	(10.428)
Pair FE	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes
KP F-stat	454.8	446.8	453.6
Observations	645,990	645,990	645,990

Table IA.9. Robustness test: alternative focal windows and visit-duration filters. This table re-estimates Table 2, column (1), varying two construction choices. In the baseline, a focal deal stays in the sample for the 24 months after the investment date, ending early if the portfolio company raises a qualifying next round. A stay counts as a visit only if it lasts between 30 and 300 minutes. Column (1) repeats this baseline. Columns (2) to (4) vary the first choice, the focal window, to 6, 12, and 18 months. Columns (5) to (7) vary the second choice, the visit-duration filter, to 10–300, 60–300, and 30–600 minutes. Specification, instrument, fixed effects, and standard errors are as in Table 2. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Dependent variable:	New pitches (t+1)						
		Focal window			Visit-duration filter (minutes)		
	Baseline	6 mo	12 mo	18 mo	10–300	60–300	30–600
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log (Involvement)	70.180***	53.182***	57.713***	66.076***	66.903***	77.932***	61.603***
	(10.286)	(16.159)	(11.813)	(10.299)	(9.870)	(11.452)	(9.043)
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Startup city × Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VC city × Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KP F-stat	454.8	187.7	313.1	420.3	490.8	416.4	462.1
Observations	645,990	269,692	448,766	571,567	645,990	645,990	645,990

Table IA.10. Robustness test: pitch measure including coffee-shop meetings. This table re-estimates Table 2 with a broader pitch measure that also counts first meetings at coffee shops. I identify coffee shops from SafeGraph's points-of-interest data, keeping places with NAICS code 722515 whose category tags contain Cafe, Coffee, Hong Kong Style Cafe, Internet Cafe, or Tea House. Each VC and each startup is assigned the coffee shop nearest its headquarters, if one exists within one mile. A coffee meeting is a same-day co-location of a VC-side device and a startup-side device, with overlapping stays, at the assigned shop of exactly one of the two sides. Coffee meetings enter the pitch panel under the same rules as office visits: meetings on or after the pair's first deal month are excluded, and each pair enters once, at the quarter of its first meeting, taking the earlier of the office-based and coffee-based dates. Column (1) repeats the baseline office-based measure. Column (2) requires the day's co-location overlap to last 30 to 300 minutes, the same duration window the office visit measure uses. Columns (3) and (4) use looser meeting definitions: any overlapping co-location, and co-location with a shared clock hour. Panels A to C report horizons of one to three quarters. Specification, instrument, fixed effects, and standard errors are as in Table 2. The Kleibergen–Paap F statistic and the number of observations are identical across the four columns within a panel: 454.8 and 645,990 in Panel A; 479.3 and 602,623 in Panel B; 439.9 and 555,636 in Panel C. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Pitch measure:	Baseline (office visits)	+ coffee, 30–300 min	+ coffee, any overlap	+ coffee, same hour
	(1)	(2)	(3)	(4)
Panel A. Dependent variable: Number of new pitches ($t+1$)				
Log (Involvement)	70.180***	69.771***	69.824***	72.549***
	(10.286)	(10.711)	(11.198)	(11.128)
Panel B. Dependent variable: Number of new pitches ($t+2$)				
Log (Involvement)	24.817***	21.600***	20.746**	22.559***
	(7.865)	(8.036)	(8.145)	(8.092)
Panel C. Dependent variable: Number of new pitches ($t+3$)				
Log (Involvement)	19.562***	16.128***	14.747**	15.086**
	(5.025)	(5.798)	(6.202)	(6.151)