

Financial Matchmaking through Third Party Networks: The Facilitation of Venture Capital Financings by Law Firm Connections

Sahil Raina and Linfeng Tian*

October 8, 2026

Abstract

Do networks unintentionally formed between capital suppliers and demanders through third parties uninvolved in matchmaking facilitate the flow of capital between them? We study this question here by examining the effect on venture finance deal formation of prior indirect connections through law firms between startups and venture capitalists (VCs). We find that the links formed through shared law firms increase the likelihood of deal formation when a VC has little experience in the startup's industry and have no effect on it when they are in the same industry. When we consider geographic distance and recency of information, we find a greater effect at greater distances and with more recent information. Overall, our findings imply that these shared law firms act as information intermediaries in settings where the startup and VC are less likely to have information about each other.

Keywords: capital allocation; venture capital; networks; law firms; startups

*Raina: Alberta School of Business, University of Alberta: sraina@ualberta.ca; Tian: Alberta School of Business, University of Alberta: linfeng.tian@ualberta.ca. We are very thankful to Vikas Agarwal for his useful comments. We acknowledge the use of generative artificial intelligence in performing the research and drafting the manuscript for this project.

1 Introduction

Finance studies how capital flows from those who supply it to those who demand it. Across several literatures in finance, papers have studied networks formed by *direct* links among suppliers and demanders and shown that the networks connecting these parties play an important role in the flow of capital.¹ However, far less attention has been paid to the networks unintentionally created by *indirect* links among them.² While networks based on direct linkages may be expected or intended to facilitate capital flows, the impact of unintentionally created, indirect networks on capital allocation is far less obvious. In this paper, we examine how unintentionally created networks may facilitate matching between capital suppliers and demanders by studying the effects on venture capital financing deal formation of networks formed between startups and venture capitalists (VCs) through their prior engagement of the same law firms.

Networks formed through law firms in venture capital are an ideal setting in which to study the effects of unintentionally created networks on matching between capital suppliers and demanders. Law firms are engaged by both VCs, capital suppliers, and startups, capital demanders, in essentially every VC financing. A VC financing is a negotiated exchange of funds for a bespoke security. As a result, the terms of that security are non-standard, effectively requiring both the startup and the VCs to retain their own legal counsel. As a result, every startup that has raised venture capital and every VC that has supplied it has an engagement history with law firms. Moreover, because the market for VC legal work is fairly concentrated, the same law firms routinely appear in the histories of startups and VCs that have never transacted with each other. Given their universal reach, such “connector” law firms could play an important matchmaking role in the VC ecosystem. Further, unlike a venture capital or private equity placement agent or an M&A advisor, a law firm is not hired to find its client a counterparty but to offer legal counsel. Any effect of a connector law firm on deal formation is an unintended by-product of a service purchased for another purpose.

Studying deal formation is difficult as it requires comparison to financings that did not happen.

¹Some of these papers study links connecting capital suppliers to one another (Hochberg et al., 2007; Sufi, 2007). Others study direct links connecting capital demanders to one another (Fracassi and Tate, 2012; Fracassi, 2017). A third group studies links connecting a supplier directly to a demander (Cohen et al., 2008; Engelberg et al., 2012; Hegde and Tumlinson, 2014; Bengtsson and Hsu, 2015).

²The papers that do study the effects of such indirect links focus on linkages formed purposefully to aid capital flow (Rider, 2009; Bao and Edmans, 2011; Cain et al., 2020).

A straightforward analysis studying this question would require studying the ratio of financings that did occur to financings that did not occur among VC-startup pairs that have a shared law firm connection to pairs that do not have such a connection. However, financings that did not occur (non-financings) are not recorded anywhere. This unobservability of non-financings has limited the study of financing deal formation in venture capital, as in other literatures. To overcome it, we use a case-control design developed in epidemiology (Breslow and Day, 1980) and adopted in the social sciences to study rare events such as the outbreak of war (Manski and Lerman, 1977; King and Zeng, 2001).³ We apply this design to PitchBook data on VC financings raised by US startups between 2001 and 2024. For each financing, PitchBook provides details such as the financed startup, the participating VCs, the law firms engaged, and which party each law firm represented. This allows us to observe every startup’s and every VC’s history of legal engagements. For each (observable) financing in a given year between a VC-startup pair (a *case*) that we observe in PitchBook, we sample VCs that could have financed the same startup in that year but did not and form *control* observations from those VC-startup pairs for that year. To improve comparability between our cases and controls, we modify this basic procedure in two ways. First, we create three strata of control observations based on whether the VC invested in the startup’s industry group (though not necessarily in the specific industry) in the past five years (*insider* stratum), it invested in the startup’s sector but not industry group in that time period (*near outsider* stratum), or it did not invest in the startup’s sector at all in that time period (*far outsider* stratum). This makes the industry distance-based distribution of the case and control observations more similar.⁴ Second, within each stratum, we draw controls using coarsened exact matching (Iacus et al., 2011) based on the case VC’s activity, type, age, and location, so that case and control observations are not too dissimilar on those dimensions.⁵ Using our case-control design with three industry distance-based strata and coarsened exact matching refining our control sampling procedure, we create a panel with 301,930 cases and 2,465,086 controls. We use this panel to study our research question.

³This design has been employed in some studies of venture capital and mergers, as well (Palepu, 1986; Sorenson and Stuart, 2001; Bena and Li, 2014; Gompers et al., 2016).

⁴If we did not do this, our dataset would have a far higher prevalence of insider case observations than insider control observations, as VCs with prior experience in an industry are far more likely to invest again in that industry.

⁵We use coarsened exact matching rather than propensity matching as it provides greater control over potential differences between case and control observations than propensity score matching, which can be misspecified (King and Nielsen, 2019).

We implement linear regression analysis on this panel to determine whether and how a prior connection through a shared law firm affects the likelihood of a financing deal between a startup and VC. We find that the deal facilitation effect of a shared law firm connection increases with the VC's industry distance from the startup. For the insider stratum, a shared law firm connection does not facilitate deal formation. In our preferred specification, the estimated effect of a shared law firm connection on the likelihood of a financing deal for the insider stratum is a precisely measured zero (a coefficient of -0.0003 with standard error of 0.008). For the near outsider stratum, a shared law firm connection increases the likelihood of a financing deal by 0.065 , an increase of 12.4% over the 52.5% base in our case-control panel. For the far outsider stratum, the connection increases the likelihood by 0.124 , an increase of 23.1% . Expressed as odds ratios, which carry over from our matched sample to the population, a shared law firm connection multiplies the odds of a financing by 1.30 for near outsiders and by 1.68 for far outsiders. Against a population base rate of approximately one financing per $10,000$ candidate VC-startup pairs, the far outsider odds ratio raises the probability to one in $5,900$. These estimates indicate that the effect of a shared law firm connection is nonexistent when the VC already invests in the startup industry and meaningful when the VC does not.

We study the deal facilitation effect of a shared law firm connection across two dimensions to corroborate our baseline findings. First, we confirm that geographic distance works much like industry distance. A shared law firm connection raises the probability of a financing by 0.068 when the startup and the VC are in different states and lowers it by 0.022 when they are in the same state. This difference is negative and statistically significant across all three strata. Second, the shared law firm connection has the greatest impact on financings when it is recent. Measured over a three-year window, a recent connection raises the probability of a financing for far outsiders by 0.141 and a stale one by 0.090 , and for insiders a recent connection has a small positive effect of 0.015 while a stale one seems to have a negative effect of -0.031 . A connector law firm facilitates a financing when the VC is far from the startup, in industry or in geography, and when the law firm's knowledge of the two parties is current.

We also run two placebo tests and a horse race to further support our facilitation interpretation of the findings. A shared law firm connection formed *after* the financing year has a much smaller association with the financing than one formed before it. This implies that our findings are not

driven by some persistent characteristic of the VC-startup pairs that share law firms. We also find that a connection through a law firm that was merely present on a deal the VC joined, rather than one that acted as the VC's own counsel, has half the effect for far outsiders, indicating that the shared law firm connection operates through the law firm's own client relationship with the VC. In addition to these two placebo-like tests, we also run a horse race between law firms and co-investors. VCs may meet startups through their co-investors, and we find that co-investor overlap increases the likelihood of financings strongly. While controlling for such co-investor networks attenuates the shared law firm effect across all three strata, it does not eliminate it or alter its relative magnitude across strata.

We interpret these findings as showing that the law firm network substitutes for startups' and VCs' own networks. A VC and startup that are active in the same industry, are located in the same state, or share co-investors have pre-existing ways for learning about each other. A connector law firm has little margin to add to them. But a VC-startup pair who are not in the same industry, are not located in the same state, and do not share co-investors have far less opportunity to learn about each other and, in such cases, a connector law firm supplies a crucial route for them to do so. The facilitation effect is therefore largest precisely for the financings that their own networks could not have produced. This is quite consistent with a shared law firm acting as an intermediary that resolves an information problem.⁶

Our paper contributes to three literatures. The first is the literature on social capital and brokerage. Burt (1992) argues that the value of a network position comes from bridging parties who are otherwise disconnected, and that bridges decay as the contacts behind them age (Burt, 2002). Our findings are a direct test of both claims in a capital market: the connector law firm's effect is largest where the gap between startup and VC is widest and it fades as the connection grows stale. The second is the literature on networks and the matching of capital. That literature shows that direct links of three kinds shape who finances whom. Links among capital suppliers matter: VCs that syndicate deals together have better access to future deal flow (Hochberg et al., 2007), and lead banks that have co-lent with a syndicate before retain smaller shares of the loans they

⁶It also explains why the effect is easy to miss. Insiders make up nearly three quarters of realized financings, so the pooled deal facilitation effect is only 0.025, and a study that did not stratify by industry distance would conclude that law firm connections barely matter.

arrange (Sufi, 2007). Links among capital demanders matter: firms whose boards interlock adopt similar financial policies and governance (Fracassi and Tate, 2012; Fracassi, 2017). And direct links between a supplier and a demander matter: a fund manager overweights companies whose directors share his alma mater (Cohen et al., 2008), a borrower whose executives went to school with a bank’s lending officers pays less for credit (Engelberg et al., 2012), and a VC is more likely to finance a founder of the same ethnicity (Hegde and Tumlinson, 2014; Bengtsson and Hsu, 2015). Hochberg et al. (2007) find that a VC’s ability to broker introductions among other VCs contributes less to fund performance than the size of its network or its access to well-connected VCs. Our findings suggest why intermediation may matter little within the VC network and a great deal outside it: VCs already reach one another through syndication, whereas an outsider VC and a startup often have no route to each other except through a third party such as a law firm. We show that a network belonging to neither side, maintained by intermediaries with no matchmaking role and no stake in the capital, shapes the match as well, and that it does so precisely where the parties’ own networks fall short. The third is the relatively small literature on law firms in financing transactions. Krishnan and Masulis (2013) and Moran and Pandes (2019) show that the expertise and reputation of a law firm are associated with better outcomes for the firms it represents in mergers and at IPO. We show that law firms matter at an earlier and more fundamental stage, in determining whether a startup and a VC transact at all, and that what matters for this is not the law firm’s reputation but the connection it forms between startups and VCs.

The rest of the paper proceeds as follows. Section 2 describes the role of law firms in VC financings and the PitchBook data we use. Section 3 describes our case-control design and our analysis data. Section 4 presents our empirical specification and findings. Section 5 concludes.

2 Institutional Setting

This section describes the role law firms play in the VC ecosystem and the PitchBook data we use to measure it. We first explain why law firms are present on essentially every venture investment and how the two-sided role of law firms creates an indirect network linking startups to VCs. We then describe the PitchBook data we use, presenting details on law firms’ characteristics, their advising coverage in VC, and the frequency with which they advise startups and VCs.

2.1 Law Firms in the VC Ecosystem

Given the complex and idiosyncratic nature of VC financings, law firms are essentially required for each financing deal. A VC financing round is effectively a negotiated exchange of cash for a bespoke claim on the startup. The VC's claim on the startup is crucial, both for its control over the startup and its share of the cash flows the startup generates. The provisions of the security establishing the claim are numerous and complex: liquidation preferences, participation rights, anti-dilution protection, board composition, protective provisions, and the conversion terms on any bridge or convertible instrument that precedes the round. None of these terms is standardized in the way a public equity issue is, and each allocates cash flow and control rights between parties who will remain in a long-term incomplete contract with one another. To navigate this complex environment carefully, both sides retain counsel, i.e., employ a law firm. The startup engages *sell-side* counsel, which, among other things, represents the startup in negotiation. And the VC engages *buy-side* counsel, which, among other things, represents the VC in negotiation. The same division of labor recurs at every round, so a startup and a VC each accumulate a history of legal engagements over time.

The two-sided role of law firms creates a network that links startups and VCs indirectly. Because both sides retain counsel on every deal, and because the market for VC legal services is concentrated in a small number of specialist firms, the same law firm may appear in the engagement histories of a startup and a VC that have never transacted with each other. Consider a VC that retains a law firm as buy-side counsel on an investment in one startup. A second, unrelated startup then retains that same law firm as sell-side counsel on a financing of its own. The VC and the second startup are now linked through the law firm they share. The link between the VC and the second startup is indirect: neither party chose it, the two need not have communicated, and it arises as a by-product of decisions each made about its own legal representation. It is this indirect link, formed before and outside of any transaction between the two parties, that we study. We call a law firm occupying this position a *connector*, and we count a link only when the law firm's engagement with the startup and its engagement with the VC occurred on different deals.

A connector law firm can plausibly affect deal formation through two channels. The first channel is *introduction*. A law firm that represents both startups raising capital and VCs investing capital

observes deal flow on both sides of the market, and it can direct one side to the other. Leading VC law firms market this capability explicitly, advertising their clients’ access to their investor networks alongside their technical legal services.⁷ Introductions should matter most where the two parties would otherwise have difficulty finding each other, i.e., in low-information environments. For instance, introductions should be more valuable when the VC has no prior experience in the startup’s sector, when the two are headquartered distantly from each other, or when neither shares a co-investor with the other. The second channel is *certification*. A law firm that has represented a startup observes details about the startup: its incorporation documents, its capitalization table, and its conduct in negotiation. The law firm can also vouch for the startup’s quality to a VC client. Certification of this kind should be stronger for more highly reputed law firms, whose endorsements carry more weight. Moreover, if the link conveys information about quality rather than merely reducing search costs, startups matched through a connector law firm should perform no worse, and possibly better, than startups matched otherwise. The two channels differ in what they imply about the period after the deal. But both imply the same first-order prediction, and it is this prediction we test: a prior shared law firm makes a financing between a given startup and a given VC more likely to occur.

2.2 Data on Law Firms in VC

We use the PitchBook database to study our research question. Our sample is the set of VC financings raised by startups headquartered in the United States between 2001 and 2024. We define a VC financing as a seed round, an early-stage venture round, a later-stage venture round, or an individual angel round. We begin in 2001 because PitchBook’s coverage before 2000 is backfilled and incomplete and because our sampling design requires a one-year look-back (see Section 3 for more details). The sample contains 205,658 financings raised by 94,946 startups from 88,614 distinct investors, and it accounts for \$2.1 trillion of reported capital. Deal size is reported for 84.6% of financings, and the median financing raises \$2.7 million. For each financing, PitchBook records the startup, the round date and type, the amount raised, the identity of every participating investor,

⁷Cooley LLP (a top VC law firm) explicitly markets this: “Our 7,000+ private company clients, combined with our unparalleled access to the largest and most influential network of investors, allows us to make introductions where appropriate.” See <https://www.cooley.com/services/practice/venture-capital/in-depth>.

and the identity of every service provider engaged on the deal, including law firms. Critically for our purposes, the service provider records identify which party the service provider represented and whether the engagement was on the buy side or the sell side. That is, we observe not only that a law firm worked on a financing, but also for whom it worked. Panel A of Table 1 reports the sample composition.⁸

Law firms are ubiquitous in VC financings, but they are quite unequal in scale. 1,935 distinct law firms advise on at least one US VC financing over our sample period, and they generate 85,895 law firm-deal engagements. If we include non-VC financing deals, 4,109 law firms appear in PitchBook from 2000 onward. This wider universe matters for our link construction because a law firm’s relationship with a startup or a VC need not begin on a venture round. PitchBook also records mergers and acquisitions, buyouts, IPOs, debt issues, and public offerings, and law firms advise on all of them. For instance, on the buy-side, for engagements that qualify a law firm as a connector, 30.2% come from merger and acquisition deals or buyouts rather than from venture rounds, and 36.9% come from other deal classes. A VC’s relationship with a law firm may begin on an exit or on an acquisition by a portfolio company rather than a financing. We therefore build law firm relationships from deals of every type rather than from venture rounds alone.

The distribution in terms of advising across law firms is extremely skewed. The median observed law firm advises two financings and serves one client, while the law firm at the ninetieth percentile advises twenty financings and serves twelve clients. Panel A of Table 2 reports the full distribution. Concentration at the top is similarly high: the ten (fifty) largest law firms account for 64.6% (84.7%) of all engagements and 66.7% (90.8%) of advised deal value. The largest law firms, as reported in Panel B of Table 2, are recognized VC specialists, i.e., Gunderson Dettmer (12,090 financings), Cooley (10,059), and Wilson Sonsini Goodrich & Rosati (8,083). We note also that law firm relationships tend to persist. Of the 36,362 distinct law firm-client pairs in our sample, 42.7% span two or more financings, and 75.7% of all engagements involve a client the law firm advises more than once. A law firm link is therefore a repeated relationship, which is what makes it a plausible conduit for information between the parties on either side of it.

⁸We use “startup” for the firm raising capital and “VC” for the firm supplying it. PitchBook’s investor records also include angels, corporate investors, and other non-VC investors, and we retain them; we use “investor” where the distinction matters.

Law firms in the VC ecosystem differ in whether they offer buy-side counsel, sell-side counsel, or both, with the last group being important for our study. We classify each law firm by the roles it plays across its engagements. 72.3% of law firms appear only as sell-side counsel to startups, 6.7% appear only as buy-side counsel to VCs, and 19.1% appear on both sides. Panel C of Table 2 reports these shares. Only the last group can create a link in our sense, because a link requires the same law firm to have represented a startup on one deal and a VC on another. While this restriction may sound severe, it does not bind in practice, because the one-sided law firms are generally small and with few clients. The 369 both-sided law firms account for 94.3% of all engagements and advise 60,663 of the financings in our sample. The gap in scale is large: the median both-sided law firm advises sixteen financings, serves eleven clients, and advises on \$323 million of capital, while the median sell-side-only law firm advises a single financing and \$4.8 million of capital. Both-sided law firms tend to advise more startups than VCs. The median both-sided law firm serves six startup clients against three VC clients, and across these law firms startup engagements outnumber VC engagements by 3.6 to one. This asymmetry reflects the structure of a financing, which has one startup but often several VCs, and also reflects the data: buy-side counsel is reported less consistently than sell-side counsel.

Turning from law firms to financings, PitchBook reports at least one law firm on 31.3% of the financings in our sample. Conditional on any law firm being reported, the modal financing has exactly one: 24.0% of all financings report one law firm, 5.6% report two, and 1.7% report three or more. Figure 1 presents these shares. Reported coverage is not constant over time or across sectors. It rises steeply over the sample period, from 15.7% of financings in 2001 to a plateau near 33% from 2008 onward, before falling back in the final year as recent deals are still being updated (Figure 2). Across industry groups it ranges from 18.6% in Capital Markets to 43.3% in Semiconductors, with an interquartile range of 26.8% to 32.7% (Figure 3). Most important for how our tests should be read, coverage rises steeply with deal size and stage. Only 11.1% of financings in the smallest size quartile report a law firm, against 61.3% in the largest, and only 6.4% of angel rounds do so, against 44.1% of later-stage rounds. Panel B of Table 1 reports these cuts.

In PitchBook data, the legal counsel roles are recorded unevenly. Among all financings, 20.9% report sell-side counsel only, 4.2% report buy-side counsel only, and 5.6% report counsel on both sides. Buy-side representation is thus recorded far less often than sell-side representation. This

pattern holds across every service provider category in PitchBook. We take this as a data limitation and assume that VCs often retain counsel on deals for which PitchBook records none. While we acknowledge this limitation, what matters for our empirical study is whether the advising law firm appears on both sides somewhere in its engagement history, because that is what allows it to bridge parties across deals. Through that lens, our coverage is nearly as broad as law firm coverage itself. At least one both-sided law firm advises 29.5% of all financings, which is 94.3% of the financings with any law firm reported. The both-sided share also tracks overall coverage almost exactly over time and across industry groups, as Figures 2 and 3 show. The set of financings on which we can detect a link is therefore essentially the set on which we observe a law firm at all, and not a further selected subset of it.

The above mentioned incomplete coverage has a specific consequence for our analysis. We can observe a link only when PitchBook records the underlying engagements, so our link indicator is measured with one-sided error. A value of zero means either that no shared law firm exists or that one exists but was not reported. Measurement error of this form attenuates the estimated relationship toward zero, so the estimates we report below should be read as lower bounds on the role of law firms. The dependence of coverage on deal size and stage also means that our tests are better powered for larger and later financings. We report the corresponding heterogeneity in Section 4 and show that our conclusions do not depend on it.

A key variable for our analysis is a shared law firm link. For each financing we take every VC-startup pair in the round and ask whether some law firm had, before that year, both a sell-side engagement for the startup on a deal in which the VC did not participate and a buy-side engagement for the VC on a deal that did not target the startup. The two exclusions remove links that arise mechanically from the two parties having already transacted together, so every link we count is indirect. Of the 586,134 VC-startup pairs in our sample, 1.79% are bridged by a prior shared law firm, and 4.64% of financings contain at least one bridged pair: 3.13% contain exactly one, 0.95% contain two, and 0.55% contain three or more. Panel C of Table 1 reports these shares. Where a link exists it is almost always singular, in that 97.6% of bridged pairs are bridged by exactly one connector law firm. The link is therefore a specific relationship rather than a diffuse overlap of legal networks. The bridged share also grows over the sample period as engagement histories accumulate, rising from near zero in the early 2000s to a plateau of roughly 2.1% after 2015 (Figure 4). These

unconditional rates are small in absolute terms. This makes a study like ours, examining the facilitation of deals by such shared law firm connections, even harder to execute.

3 Analysis Data

In this section, we describe the sample we create to test whether a prior shared law firm makes a financing more likely. We first explain our case-control design, which we use to build our sample of VC financings of startups and “non-financings,” describing the stratification and matching we impose on it. We then describe the resulting panel and the raw incidence of shared law firm links among financings and non-financings in our data.

3.1 Case-Control Sampling

Testing whether a shared law firm connection facilitates financing requires comparison to financings that *did not happen*. The frequency of realized financings with a shared law firm connection is only informative when compared to the frequency of such connections among startups and VCs that could have transacted and did not (i.e., non-financings). Unfortunately, no VC datasets, including PitchBook, contain such non-financings, as they (reasonably) act as repositories of realized transactions. A VC that considered a startup and passed on investing in it does not show up in PitchBook, and neither do the VCs that never considered the startup at all. The non-financings our research question needs, therefore, have to be constructed rather than collected.

Case-Control Sampling in the Literature

Case-control sampling addresses this problem by keeping every financing and a sample of the generated non-financings. The justification for sampling is a trade-off in how a researcher spends a fixed budget. Non-financings are close to unlimited in number,⁹ but each one on its own says very little, so measuring a few of them carefully is worth more than counting many of them crudely (King and Zeng, 2001). Sampling also makes the regression’s output more useful. When an outcome occurs a handful of times in ten thousand, the estimated probability of success or failure approaches zero,

⁹For instance, the number of realized financing deals in a year is tiny relative to the number of potential financing deals, the cross product of the number of active startups and VCs.

and the fit of a regression is driven more by the rarity of the event rather than by the variable of interest. Keeping every financing and only a sample of the non-financings raises the frequency of the outcome in the estimation sample and avoids that problem. The practical take-away is that a modest number of non-financings per financing is enough to estimate the relationship precisely. One non-financing per financing delivers about half the precision that using every non-financing would give and four deliver about eighty percent, after which each additional one adds little (Breslow and Day, 1980). We draw more than five, as described below, allowing for additional subsample comparisons.

Designs akin to case-control sampling are quite prevalent in studies that explore the causes of transactions between agents. In venture capital, Sorenson and Stuart (2001) study the formation of startup-investor ties against a sampled set of pairs that did not form. The same matching problem is taken up with different machinery by Sørensen (2007) and Bottazzi et al. (2016). Elsewhere in finance, the approach is common in work on mergers, where Bena and Li (2014), Ahern et al. (2015), and Akkus et al. (2016) compare realized acquirer-target pairs to counterfactual pairs that could have formed and did not. It also appears in work on the choice of a financial intermediary, as in Ljungqvist et al. (2006) on securities underwriting and Alperovych et al. (2022) on syndicated lending. In economics more broadly, the same structure appears in international trade, where the country pairs that do not trade carry the information about the fixed costs of entering a market (Helpman et al., 2008).

Sample Construction

In the estimation sample that we build using a case-control design, the unit of observation is the (startup, VC, year) triplet, and our outcome is an indicator for whether that VC invested in that startup in that year. The cases are the realized financings. Because a round usually has more than one participating VC, we take every VC financing from 2000 onward, defined as in Section 2, and record one triplet for each VC that participated in it. A round raised from three VCs therefore contributes three cases. This gives 731,541 triplets where a financing occurred.

The triplets for which no financing occurred vastly outnumber those where a financing did occur. For instance, in 2015, 6,264 startups in our sample raised a financing and 26,691 VCs had made at least one investment in the preceding five years. The cross product of the two contains 167

million triplets for that year alone, of which 16,578 ended in a financing, a rate of about one in ten thousand. Enumerating the full cross product is impractical, and, by the argument above, unnecessary. Sampling four or more controls per case observation is enough to establish statistical precision (Breslow and Day, 1980) and, as we detail next, we choose control observations to be as similar as possible to the case observations.

Non-financings (i.e., controls) must be drawn from the VCs who could have participated in a round and did not. The natural VC candidates for these control triplets are VCs with recent deals in a startup’s industry. Sampling from that pool alone, however, imposes an invisible restriction on controls that cases do not face. A startup’s realized investors include VCs with no prior deal in its industry. Roughly a quarter of our cases, 26.5%, are of that kind. Controls drawn only from industry insider VCs therefore have no common support with a quarter of our cases. Those cases have no comparable non-financings in the sample, so any effect estimated for them rests on extrapolation from insider VC controls. The problem is also not merely one of coverage. Industry experience is correlated with law firm sharing, because industry-specialist VCs and the law firms that serve them cluster around the same startups, so a comparison of all cases to industry-insider-VC-only controls conflates the effect of the law firm link with the effect of being an industry insider. It is also the dimension along which the law firm channel is most likely to operate, since an introduction is worth most to a VC who has no independent route into the startup’s sector.

We restore common support between cases and controls by stratifying on industry distance. For each triplet, we classify VC j relative to startup i using j ’s deal history over the five years before year t . VC j is an *insider* if it invested in i ’s primary industry group. It is a *near outsider* if it invested somewhere in i ’s industry sector but not in i ’s group. It is a *far outsider* if it was VC-active but invested in neither the startup’s industry group nor its sector. The classification is specific to the startup rather than to the investor, so the same VC may be an insider for one startup-year and a far outsider for another in the same year. VCs with no financings in the five-year window have no defined position relative to any startup, so we drop such cases, 145,867 in total. This is also why the estimation panel begins in 2001 rather than 2000. We observe deals from 2000 onward, so in 2000 itself no investor has an observable five-year history and no triplet can be placed in a stratum.¹⁰

¹⁰This step also drops the small number of cases whose startup carries no primary industry classification in PitchBook, since the strata are defined relative to that classification.

We then sample controls separately within each stratum. The sampling unit is the (startup, year, stratum) cell. For each cell that contains at least one case, we draw up to twenty controls from the investors in that stratum who did not participate in the round, and for cells with no case, we draw no controls. Drawing the same number of cases in every stratum, rather than a number proportional to the case count, keeps the financing itself as the event we sample, rather than each VC’s individual participation in it. The case VCs in a given cell competed against the same field of non-participant VCs, so we assign them the same pool of control VCs, rather than drawing a separate pool for each and potentially counting one non-financing several times over.¹¹ The stratum-eligible pools are large in every stratum, with a median of 8,005 candidates for insider cells, 7,535 for near outsider cells, and 25,446 for far outsider cells, so the size of the pool is not a binding constraint on the sample draw.

Stratifying on industry distance makes cases and controls comparable along the industry dimension but allows significant differences in all other dimensions. This is problematic as, within a stratum, the VCs that invest are much more active than the VCs that do not. Among insiders, our randomly drawn controls differ from cases by 1.06 and 1.08 standard deviations in terms of number of financing deals and number of distinct portfolio companies in the last five years for VCs. A VC that finances more startups also engages more law firms and has more opportunities to share one of those law firms with any given startup. A comparison that ignored activity would credit the law firm link with what is really a difference in how active the two sides are. We therefore match controls to cases on observable investor characteristics before estimating anything, so that comparability is built into the sample rather than imposed by the functional form of a regression. We use coarsened exact matching (Iacus et al., 2011) instead of matching on an estimated propensity score, for two reasons. First, the coarsening of the variables that we match on fixes in advance how different a matched control can be from its case, so the imbalance that remains is a property of the design rather than an outcome of a first-stage estimation. Second, coarsened exact matching requires no model of the probability that a VC invests in a startup, so there is no first stage that could be misspecified (King and Nielsen, 2019).

We apply our matching technique within each cell on five VC characteristics, coarsened as

¹¹It also means that the pooled estimates we sometimes report weight the three strata by their realized shares of financings, which is the population-relevant weighting.

follows: the number of VC deals and the number of distinct portfolio companies over the previous five years, each coarsened into quintiles of its global distribution; VC type, coarsened to institutional VC, corporate VC, and other; VC age since its first observed deal, coarsened into five bins; and the census division of the VC’s headquarters, matched exactly. A candidate control is retained only if it has the same set of coarsened characteristics as a case in the same cell. Each retained control is then assigned a proportional weight: the number of cases divided by the number of retained controls sharing that same set of coarsened characteristics. Cases are assigned a weight of one. Coarsened matching fails to find any counterpart for 0.52% of case rows. It closes most of the activity gap that motivated it, with the standardized differences among insiders falling from 1.06 and 1.08 to 0.35 and 0.38 for number of VC deals and number of portfolio companies, respectively. Table 3 reports the outcome of our matching and weighting on all five matching variables. Though our matching and weighting procedure drastically reduces observable differences among the cases and controls, the two measures of VC activity remain the least well balanced.¹² We control for these characteristics in our regressions to account for any remaining observable differences between cases and controls.

We apply two additional filters to finalize our estimation dataset. First, we keep only triplets in which the startup and the VC are both headquartered in the United States, which matches the scope of the institutional description in Section 2 and avoids mixing legal systems. We then drop any pair with a prior direct deal between them, so that every remaining pair is one that had not yet transacted. The key regressor is measured on this panel exactly as the link was defined in Section 2. *Shared law firm* equals one for the triplet (i, j, t) if some law firm had, before year t , a sell-side engagement for startup i on a deal in which VC j did not participate, and a buy-side engagement for VC j on a deal that did not target startup i . Engagements are counted from deals of every type, not only venture rounds, for the reason given in Section 2.¹³

¹²The conventional benchmark for acceptable standardized mean differences is 0.1. Our SMDs for both number of VC deals and portfolio companies are above that for all three strata. The reason is that these two variables end up more coarsely binned than the others. Ties at low counts collapse the quintile cut points, leaving three bins for the number of VC deals (one to two, three to four, and five or more) and two for the number of portfolio companies (one to three, and four or more). CEM only sets the bin a control falls in but does not reduce differences within a bin.

¹³Note that a shared law firm plays no part in how the sample is built. It is not a stratum boundary and not one of the five characteristics we match on, so a case and its matched controls are equally free to have one, and the design cannot mechanically produce a difference between them.

3.2 Data Description

The estimation panel contains 301,930 cases and 2,465,086 controls, drawn from 65,988 startups, 39,754 VCs, and 121,181 startup-years between 2001 and 2024. Panel A of Table 4 reports the construction from the 731,541 raw triplets to these counts. As Panel B shows, the overall ratio of controls to cases is 8.2 to one, which is below the twenty drawn per cell for two reasons. First, a cell often contains more than one case, 2.1 on average, and, second, coarsened matching sometimes leaves fewer than twenty control candidates. The median cell still has ten controls per case, with 1.9 at the tenth percentile and twenty at the ninetieth. Four fifths of cells, holding 62.1% of case rows, are above the four-to-one efficiency plateau (Breslow and Day, 1980), and the cells that fall below it are those with large syndicates or thin matched pools. In 3.0% of cells, no control survives at all, either because coarsened matching found no counterpart or because the United States filter removed the controls that had been drawn. Those cells hold 5,480 case rows, or 1.8% of all cases.

The panel grows over the sample period along with the VC market. As Panel C of Table 4 reports, cases rise from about 3,000 a year in the early 2000s to 40,648 in 2021, the peak year, and fall back to 22,728 in 2023.¹⁴ The ratio of controls to cases moves in the opposite direction, from about twelve to one in the middle of the 2000s to between six and seven to one in 2021 and 2022. This is a syndicate size effect rather than a sampling failure. Because controls are drawn per cell and cases are not, a year in which rounds include more VCs mechanically lowers the ratio.

As Panel B of Table 4 reports, insider cases make up most of the realized financings, but the outsider strata are large in absolute terms. Insiders account for 222,076 cases, or 73.6% of the total, near outsiders for 38,868 cases, or 12.9%, and far outsiders for 40,986 cases, or 13.6%. The two outsider strata together have almost 80,000 realized deals, which is what allows us to estimate a separate law firm effect in each of them. Panel D reports some key observables for cases and controls. VCs with realized financings remain more active than their matched controls, with a mean of 62.3 prior deals against 32.1, although the gap at the median is much smaller, at 19 against 13. As mentioned earlier, coarsening our matching variables into quintiles narrows their difference but does not erase it. We therefore control for VC activity directly in every specification.

Turning to the shared law firm connection, our variable of interest, we find that cases are linked

¹⁴The 2024 figure of 6,091 reflects the May 2024 cut of the data rather than a collapse in activity.

by a prior shared law firm more often than the controls they are matched against. Panel A of Table 5 reports the two-by-two table. Of the 301,930 cases, 4,368 have a prior shared law firm, a rate of 1.45%. Of the 2,465,086 controls, 28,221 do, a rate of 1.14%. The case rate exceeds the control rate by a factor of 1.26. This is an early indicator of the facilitation effect of shared law firm connections that we are studying in this paper. It is admittedly small in absolute terms because, as established in Section 2, shared law firm links are rare. But, the 26% higher rate of shared law firms among realized financings suggests a meaningful influence of shared law firms in financings.

The pooled comparison conceals an interesting gradient across our strata. Panel B of Table 5 and Figure 5 repeat the comparison of shared law firm connections for cases and controls within each stratum. Among insiders, the case rate is 1.75% against 1.51% for controls, a ratio of 1.16. Among near outsiders, it is 0.84% against 0.58%, a ratio of 1.46. And, among far outsiders, it is 0.39% against 0.17%, a ratio of 2.24. The prevalence of shared law firm connections is lowest when cases and controls differ most sharply (far outsiders) and vice versa. Because insiders make up nearly three quarters of the cases, the pooled ratio of 1.26 is close to the insider ratio and understates what happens in the two outsider strata. The ordering is visible here without any regression adjustment, and, as we will show in Section 4, it survives the full set of controls and fixed effects we apply in our regression analyses.

Figure 6 shows that the incidence of shared law firm links rises over the sample period for both cases and controls. Both begin at zero, which is mechanical: a link requires prior engagements, we observe engagements only from 2000 onward, and the first years of the panel are therefore years in which almost no pair can be linked. The shared law firm rates become economically meaningful in 2005 and plateau after 2014, with cases averaging 1.66% and controls 1.32% over 2015-2024.

4 Results

This section presents our findings for how a prior shared law firm connection affects the likelihood that a startup and a VC transact in a financing deal. First, we present our regression specification and explain how its coefficients should be read given our case-control design. Then, we present the baseline findings, which show that the deal facilitation effect of a shared law firm link increases monotonically with the industry distance between a startup and a VC. The remainder of the section

studies whether that pattern is consistent with law firms acting as information intermediaries. We test whether the effect is larger when the two parties are geographically distant, whether it depends on how recently the law firm worked for each side, how controlling for the co-investor network that connects the same VC-startup pairs affects our baseline findings, which law firm relationship matters most for the deal facilitation effect, and perform a placebo test using post-financing law firm relationships.

4.1 Empirical Specification

We estimate a linear probability model in which the effect of a shared law firm connection is allowed to vary across the three strata. For startup i , VC j , and year t , we estimate

$$D_{ijt} = \beta \text{SharedLF}_{ijt} + \delta^N \text{SharedLF}_{ijt} \times \text{Near}_{ijt} + \delta^F \text{SharedLF}_{ijt} \times \text{Far}_{ijt} + \gamma'_{s(i,j,t)} X_{ijt} + \alpha_{k(i)t, s(i,j,t)} + \mu_{j, s(i,j,t)} + \varepsilon_{ijt}, \quad (1)$$

where D_{ijt} equals one if VC j financed startup i in year t and zero otherwise, and SharedLF_{ijt} is the shared law firm indicator defined in Section 3. Near_{ijt} and Far_{ijt} are indicators for the near outsider and far outsider strata, so the insider stratum is the omitted category. The vector X_{ijt} holds the controls described below, $\alpha_{k(i)t}$ is a fixed effect for startup i 's industry code interacted with the year, and μ_j is a VC fixed effect. Because the strata are defined relative to the startup, a given VC may appear in all three strata within the same year, which allows us to estimate δ^N and δ^F alongside VC fixed effects.

The subscript $s(i, j, t)$ on the control coefficients and on both fixed effects indicates that each is specific to the stratum the triplet falls in. Every control slope and fixed effect is estimated separately for insiders, near outsiders, and far outsiders. Estimating Equation (1) is therefore equivalent to estimating the model separately within each stratum.¹⁵ We run our analyses jointly on all three strata because doing so keeps a single variance matrix, which lets us test the difference between any two strata's deal facilitation effects. Allowing the control slopes to differ is important because the three strata are different populations with differing characteristics. For instance, the relationship between a startup's maturity and its probability of raising from a given VC is not the same for a

¹⁵We tested this by running the model separately for each stratum and found identical coefficients.

VC that already invests in a startup’s sector as for one that does not.

We weight every observation’s contribution to the regression estimates by the coarsened exact matching weights and cluster standard errors two-way, on the startup and on the VC. We cluster our standard errors at the startup level and at the VC level because decisions made by either entity affect the outcome we study, creating dependence between observations within each startup and VC.

Case-control sampling affects how to interpret reported coefficients. Our sample retains every financing and only a sample of the non-financings, so the frequency of financings in the estimation sample is far above its true frequency in the population of VC-startup pairs. Under the matching weights, the financing rate in the estimation sample is 47.7%, against a true rate on the order of one in ten thousand in the full cross product of active startups and active VCs described in Section 3. A coefficient from Equation (1) is therefore a probability difference within the case-control sample rather than a population probability. The odds ratio, however, is invariant to how cases and controls are sampled as long as sampling does not depend on the shared law firm indicator, which by construction it does not. We therefore report coefficients as estimated, and, when providing magnitudes, convert them into odds ratios and into effects relative to the financing rate in the estimation sample.¹⁶

4.2 Baseline Results

We study the deal facilitation effect of prior shared law firm connections across six increasingly saturated specifications and present our findings in Table 6. Column (1) of the table reports findings for a regression which includes only year fixed effects. The regression reported in Column (2) adds VC activity controls: log of the number of VC deals and log of the number of distinct portfolio startups over the previous five years.¹⁷ Column (3) adds startup controls: log of the number of prior financing rounds, log of the total capital raised to date, and log of the startup’s age. Column (4) adds an indicator for the startup and the VC being headquartered in the same state. Column (5)

¹⁶An alternative is to estimate a logit, whose slopes are invariant to case-control sampling and whose intercept alone shifts. We use the linear probability model because our preferred specification carries high-dimensional fixed effects, for which logit is subject to the incidental parameters problem. For robustness, we have run the specification as a logit where the fixed effects remain low-dimensional, and the implied odds ratios are similar to those that the linear probability model delivers.

¹⁷These were the two variables that our matching left least well-balanced. See Section 3 for more details.

replaces the year fixed effects with industry code by year fixed effects. And, Column (6) adds VC fixed effects and is our preferred specification. The regression reported in that column identifies the shared law firm effect from variation within a single VC across the startups it could have financed, which removes any time-invariant difference in how active, how well connected, or how well covered by PitchBook a given VC is. Note that, as we care about the total shared law firm effect for each stratum, we also report the effect implied for each stratum in the second panel of the table. The deal facilitation effect is β for insiders, $\beta + \delta^N$ for near outsiders, and $\beta + \delta^F$ for far outsiders.¹⁸ Finally, we separately report a pooled effect across all three strata, which is the average of the three stratum effects weighted by each stratum’s share of the realized financings, i.e., 0.736 for insiders, 0.129 for near outsiders, and 0.136 for far outsiders.

As Table 6 reports, we find that the deal facilitation effect of prior shared law firm connections is effectively zero for the insider VC stratum, positive and significant for the near outsider VC stratum, and larger still for the far outsider VC stratum. Moreover, the estimates of the shared law firm effect are stable once we include startup controls. Between columns (3) and (6), the insider effect stays between 0 and 0.015, the near outsider effect stays between 0.065 and 0.074, and the far outsider effect stays between 0.110 and 0.124. The estimated coefficients in the first two columns show considerably more movement. This is because Column (1) does not control for any (potentially correlated) VC or startup characteristics and Column (2) controls for how active a VC is without conditioning on the sector and year in which the financing takes place. Note that adding VC fixed effects in column (6) raises the R^2 from 0.073 to 0.276 without significantly altering the outsider estimates, which tells us that the pattern is not driven by differences between VCs in how often they transact or in how often they share law firms with anyone.

For insider VCs, our analyses show that a shared law firm does not increase the likelihood of a financing. The point estimate in our preferred specification, Column (6), is -0.0003 with a standard error of 0.008, which is a t -statistic of -0.04 . Given the large number of insider pairs with a shared law firm, the estimate is precise enough to rule out an effect larger than 0.013 in magnitude,¹⁹ which is approximately one-tenth of the far outsider effect. We further explore this absence of an effect

¹⁸We compute the standard error of each sum by the delta method from the estimated variance-covariance matrix, so the reported significance of a stratum effect accounts for the covariance between the baseline coefficient and the interaction.

¹⁹Appendix A reports this bound as well as the equivalence test used to derive it.

later in this paper and find that a shared law firm raises the probability of a financing for insider pairs in different states and lowers it for insider pairs in the same state. We also show that old connections to the law firm lower the probability of a financing among insider pairs.

For outsider VCs, we find that a shared law firm increases the likelihood of a financing. The implied effect in Column (6) of Table 6 is 0.065 for near outsiders and 0.124 for far outsiders, with t -statistics of 3.84 and 4.94. These magnitudes are both statistically significant and economically meaningful. In the estimation sample, the probability of a financing for a near outsider pair with no shared law firm is 52.5%, so the estimate is a relative increase of 12.4% ($0.065/0.525$). For far outsider pairs, the corresponding base is 53.7% and the relative increase is 23.1%. Expressed as odds ratios, which are the magnitudes that carry over to the population, a shared law firm multiplies the odds of a financing by 1.30 for near outsiders ($(0.590/0.410)/(0.525/0.475)$) and by 1.68 for far outsiders. Applied to a population base rate of one financing per ten thousand candidate pairs, an odds ratio of 1.68 raises the probability of a financing from roughly one in 10,000 to roughly one in 5,900.

In Column (6) of Table 6, averaging the three stratum effects by their shares of realized financings gives a slightly positive (0.025) overall deal facilitation effect for prior shared law firm connections. While the coefficient estimates for the outsider strata are much larger, note that insiders make up nearly three quarters of realized financings, so the pooled number is closer to the insider estimate. This is why the shared law firm deal facilitation effect may be economically important and still be nearly invisible in the broader ecosystem. Precisely in the situations where some form of intermediary is most valuable (for VCs and startups in distant industries), shared law firms raise the odds of a financing by two thirds.

The pattern documented in Table 6 is consistent with the shared law firm acting as an information intermediary. A VC that already invests in a startup’s industry group sees that startup’s sector deal flow through its own channels, so a law firm has little to add. A VC with no prior experience in the sector has no such channel, and the information a law firm holds about which startups in that sector are raising funds, and on what terms, is correspondingly more valuable. The deal facilitation effect should therefore increase with the information asymmetry between the two parties, which is what the monotonic increase in the effect across strata by industry distance shows. The rest of this section tests the information intermediary interpretation on four further margins.

4.3 Information Intermediation

If the law firm functions as an information intermediary, its value should fall when a startup and a VC have other ways of finding each other. Geographic proximity may increase the likelihood of the two parties finding each other as VCs and startups operating near each other may meet through other local networks such as shared board memberships, industry events, etc. Moreover, the entrepreneurship literature has already documented that VC investment falls sharply with distance (Sorenson and Stuart, 2001). Therefore, we assert that, if a shared law firm acts as an information intermediary, then its value for deal facilitation should be lower for VC-startup pairs in the same state than for pairs in different states. We test this by interacting Equation (1) fully with the same-state indicator, which yields a separate deal facilitation effect for same-state and cross-state pairs within each stratum, and an interaction term estimating the difference between them.

Table 7 reports the results of our same- versus cross-state analysis. We find that geographic proximity substitutes for the shared law firm connection in every stratum. Across all pairs, a shared law firm raises the probability of a financing by 0.068 when the startup and the VC are in different states, and lowers it by 0.022 when they are in the same state, a difference of -0.096 with a standard error of 0.011. The difference between the effect in same- versus cross-state pairs is negative in all three strata: -0.093 for insiders, -0.056 for near outsiders and -0.109 for far outsiders, and it is statistically significant in each, albeit only at the 10% level for near outsiders. Overall, the table shows that every stratum’s deal facilitation effect is positive once the pair is geographically separated (i.e., for cross-state pairs), and every stratum’s effect falls when it is not (i.e., for same-state pairs).

The insider row in Table 7 presents interesting findings about the insider stratum. While Table 6 reports an insider average effect of zero, the geographic split shows a positive effect of 0.046 for cross-state pairs and a negative effect of -0.047 for same-state pairs. Even for a VC that already invests in the startup’s sector, a law firm connection *increases* the likelihood of a financing when the two parties are geographically distant. However, for a VC that already invests in the sector and is also geographically near the startup, the connection *reduces* the likelihood of a financing. The negative estimate arises exactly where a law firm introduction is most redundant. A VC that already invests in the startup’s sector and is headquartered in the same state has, through its own sector deal flow and its local network, many of the channels a law firm would otherwise supply. The shared law firm

channel has a positive impact where information is scarce, and where the two parties already have the channels it would substitute for, the effect turns negative.²⁰

Another implication of the information intermediary role is that the value of the connection should decay. Information about a startup’s prospects or a VC’s appetite ages, and a law firm that has not worked for either side in years holds stale information and has little current contact through which to pass it on. We test this by splitting the shared law firm indicator into a recent and a stale component. For a window of W years, a pair is linked *recently* if some connector law firm was engaged by both the startup and the VC within the last W years, and it is linked *stalely* if it is linked at all but not recently. We estimate Equation (1) with both components interacted with the strata, for W from one to five years. Table 8 reports the implied stratum effects.

We find that the deal facilitation effect is much larger for recent connections and smaller or even negative for stale ones. In Table 8, regardless of where we split the sample, the implied effect for the recent row exceeds that for the stale row for insiders and far outsiders at every window, and for near outsiders at the one- and two-year windows. At a three-year window the insider effect is +0.015 when the connection is recent and -0.031 when it is stale, the near outsider effect is 0.064 against 0.069, and the far outsider effect is 0.141 against 0.090. Widening the window makes the gap wider for insiders, whose stale effect falls to -0.038 at five years, and for far outsiders, whose stale effect falls from 0.108 at one year to 0.035 at five and loses significance while the recent effect holds near 0.14. The near outsider column is the exception, in that its stale effect rises with the window rather than falling.²¹

To study the role of recency further, we explore whether the recency of the VC or the startup link to law firms has any impact on the deal facilitation effect. Table 9 splits VC-startup pairs by whether the law firm’s engagement by the startup and its engagement by the VC are each within three years, assessed independently, and reports all four combinations for each stratum. Among

²⁰From a straightforward perspective, redundancy should result in a zero effect rather than a negative one. However, there may be countervailing effects of sharing a law firm that become visible precisely when the information sharing channel is redundant. For instance, law firms may introduce startup clients seeking financing to their VC clients selectively and, when the value of information sharing is sufficiently low, the selectivity of the process may lead to a negative deal facilitation effect.

²¹A possible mechanical explanation for this reversal is that, as W widens, the recent category absorbs more pairs and the stale category becomes a smaller and more selected group. Of the 2,355 near outsider pairs with a shared law firm, the stale category holds 1,832 at a one-year window but only 422 at a five-year window, so the estimate is drawn from a progressively smaller group.

insiders, the effect is positive when the VC side is recent, at +0.015 and +0.016 for the two levels of startup-side recency, and negative when the VC side is stale, at -0.040 and -0.034 . This shows that startup-side recency does essentially nothing on its own and the relationship that matters in terms of recency is the one between the law firm and the VC. This is consistent with an introduction channel running from the law firm to its VC client, since a law firm passes a startup to a VC it is currently working for and not to one it last advised five years ago. The same ordering appears for far outsiders (and, to a lesser extent, for near outsiders), where a recent VC connection shows a 0.143 (0.064) effect and a stale one shows a 0.099 (0.059) effect, though the stale-startup cells in the outsider strata rely on small samples.

Taken together, the two analyses described here are consistent with the shared law firm acting as an information intermediary. Both analyses study what happens when the connection has less information to supply, and both answer in the same direction. Geographic proximity gives the two parties other ways of finding each other, and the deal facilitation effect falls in every stratum when the startup and the VC are in the same state. Time erodes both what the law firm knows and whether it is still in contact with either side, and, for insiders and far outsiders, the effect is much stronger for recent connections than stale ones, with the decay seemingly driven through the law firm’s relationship with the VC rather than the startup’s. The two analyses also account for the insider average of zero in the same way: it reflects not the absence of a relationship but the average of a positive effect where the law firm connection is likely the primary information channel and a negative one where the connection is either redundant or stale.

4.4 Law Firm Connections Proxying for Co-investor Networks

A reasonable alternative interpretation of our results would be that a shared law firm proxies for a co-investor connection. Startups and VCs are connected through the syndicates that the VCs belong to, those co-investment connections are known to expand the radius over which a VC invests, and a law firm connection may be nothing more than an artifact of an underlying network position. We study this alternative interpretation carefully and find evidence that a correlation exists in the facilitation effects of co-investor networks and shared law firms, but neither effect entirely eliminates the other. In this section, we detail this exploration of co-investor networks.

First, we define our co-investor overlap measures. For each VC-startup-year triplet, we define

overlap as the number of third-party investors that both financed startup i in the five years before t and co-invested with VC j in a VC deal over the same window. We use *overlap* both in logs and as an indicator for any overlap. In a simple preliminary test, we confirm that VC-startup pairs with a shared law firm are far more likely to share a co-investor: 60.5% of them have at least one, against 12.7% of pairs without a shared law firm, and the correlation between the law firm indicator and log overlap is 0.18.

Next, we run a horse race between shared law firm connections and co-investor overlap. Table 10 presents the results of the horse race. Column (1) repeats our preferred baseline specification as a comparison. Columns (2) and (3) add co-investor overlap as a single pooled slope, in logs and as an indicator. The co-investor channel is strong on its own, at 0.116 per log unit of overlap, and column (4) shows that this is essentially unchanged when the law firm terms are dropped. Adding the co-investor overlap weakens the shared law firm effect but does not change the pattern of the increased deal facilitation effect by industry distance. The implied far outsider effect falls only from 0.124 to 0.101, and the near outsider effect from 0.065 to 0.034, while the insider effect turns negative at -0.035 .

We run an even more demanding version of the test, in which we let the co-investor channel vary by stratum as well. While it weakens the shared law firm effect further, we find that it persists in the far outsider stratum, where information intermediation between startups and VCs is most valuable and the effect weakens as the industry distance between startups and VCs declines. In Column (5), we regress financings on co-investor overlap by stratum separately and find that the co-investor effect is also steeply graded by industry distance, at 0.104 for insiders, 0.188 for near outsiders, and 0.225 for far outsiders. This implies that a single pooled slope is dominated by insiders and understates the outsider co-investor effect, so controlling for it leaves too much of the outsider variation for the law firm term to absorb. Column (6) therefore stratifies both channels. It is the toughest specification in the paper, and it attenuates our estimates without eliminating them. The law firm effect is -0.032 for insiders, 0.015 for near outsiders, and 0.080 for far outsiders. The near outsider effect is no longer statistically significant. What remains significant is the far outsider effect and the difference between strata. The near outsider premium over insiders is 0.047 with a standard error of 0.018, and the far outsider premium is 0.111 with a standard error of 0.027, both significant at the one percent level. These findings show that, even saturating the regression with

a fully stratified co-investor channel, the shared law firm connection increases the likelihood of a financing more when the VC is further from the startup’s sector and, for far outsiders, for whom the information intermediation is most valuable, it raises the probability of a financing on its own.

4.5 Which Law Firm Relationship Matters Most for Deal Facilitation?

Our link requires the law firm to have advised the startup on the sell side and the VC on the buy side, which is the pairing we describe in Section 2. We can also construct three mismatched versions of the same link, in which the law firm advised an investor in the startup’s earlier round rather than the startup itself, or advised the startup on a deal the VC participated in rather than acting as the VC’s own counsel. If the effect were specific to the formal roles, the mismatched links should have weaker facilitation effects for financings. In Table 11, we report our findings for analyses of such mismatched links.

Our analyses tell us that the VC’s link to the law firm matters more than the startup’s link. In the far outsider stratum, the correctly specified link increases the likelihood of a financing by 0.124. Mismatching the law firm’s link for the startup connection leaves 0.096, mismatching the VC’s link leaves 0.062, and mismatching both leaves 0.040. The near outsider row behaves the same way for the VC’s link, at 0.065 correct against 0.048 when the VC’s link is mismatched, while mismatching the startup’s link raises the estimate to 0.082. In both outsider strata, a law firm that acted as the VC’s own counsel has a larger effect than a law firm that was merely present on a deal in which the VC participated, twice as large for far outsiders and about a third larger for near outsiders. On the other hand, it does not matter as much whether the law firm came to know the startup by representing it or by representing an investor in one of its earlier rounds.²²

This asymmetry between VC and startup links helps address another objection. The two indicators that mismatch the VC’s link are far more common than the correctly specified one, at 265,403 and 145,102 pairs against 32,589, while the indicator that mismatches only the startup’s link is not, at 25,136. The difference arises from the definition of a mismatched VC link: only that the law firm was sell-side counsel on a deal the VC joined, which attaches the firm to every

²²The insider row’s correctly specified link has a well-estimated deal facilitation effect of zero, so it is not particularly meaningful to test which link matters more for the effect.

investor in that syndicate, rather than that it acted as the VC’s own counsel.²³ If so loose a measure impacted financing likelihoods as well as the correct bilateral relationship, it would suggest that our variable captures a general proximity between the two parties in the market for legal services rather than a direct informational connection between them. However, as the looser measures impact the likelihood substantially less, and particularly where the VC side is mismatched,²⁴ we infer that the effect requires that the law firm worked for the VC, rather than merely working in its proximity.

4.6 Post-Financing Relationships Placebo

A remaining concern is that our indicator marks VC-startup pairs that are close in the legal services market for persistent reasons that are unrelated to the passage of information between them. Two parties in the same city, the same sector, and the same corner of the legal market will interact with the same set of law firms and be more likely to jointly participate in a financing, without the law firm acting as an information intermediary between them. Such a time-invariant property of a VC-startup pair does not depend on when the engagements occurred, so it should be as visible in the law firms that the two parties engage *after* a financing as in those they engaged before it. We test this by building the identical shared law firm connection in both directions. One measure requires a connector law firm to have engaged the startup on the sell side and the VC on the buy side in the three years before year t ; the other requires the same in the three years after. Both measures have the same anti-participation restrictions and differ only in the direction of the window.²⁵

A shared law firm connection that precedes the financing increases its likelihood far more than a connection that follows it. Table 12 reports results for both types of connections. The backward-looking connection is correlated with an increase in the probability of a financing of 0.161 for far outsiders and 0.076 for near outsiders, against 0.055 and 0.038 for the forward-looking one, ratios of about three to one and two to one. For insiders, the backward-looking connection is 0.020 and the forward-looking one is -0.008 and indistinguishable from zero.²⁶ We note that the forward-looking

²³Buy-side counsel is also recorded less often than sell-side counsel in PitchBook, as Section 2 reports.

²⁴We note that recency also mattered more on the VC side.

²⁵The sample runs from 2003 to 2021, the years for which a full three-year window is observable in both directions. Estimates therefore differ from those in Table 6, which uses the whole sample period and builds the connection from the connector’s entire engagement history rather than a three-year window.

²⁶The insider estimate of 0.020 also corroborates the recency results of Table 8, where the insider effect at a three-year window is 0.015, despite being built differently.

measure is biased upward, because a financing in year t gives a VC reason to engage the startup’s law firm again afterwards, which makes the reduced forward effect an even stronger rejection of this concern. Also, the backward-looking connection is the rarer of the two, at 16,150 pairs against 35,868, so its coefficient is unlikely to be larger based on measurement quality alone. Given these points as well as clear evidence that the two sets of coefficients are not equal, it is improbable that our findings are driven by some persistent property of VC-startup pairs.

5 Conclusion

Networks shape which suppliers of capital transact with which demanders of capital. The networks studied to date are formed by direct links among the parties themselves, and such links are typically formed with future transactions in mind. In this paper, we study a network that neither party forms on purpose. Startups and VCs each engage law firms for legal counsel, and because the market for VC legal work is concentrated, a startup and a VC that have never transacted often share a law firm in their engagement histories. We ask whether such an inadvertent connection through a shared law firm increases the likelihood that the startup and the VC transact in a financing deal.

Using a case-control design that compares realized financings to financings that could have occurred but did not, we find that the deal facilitation effect of a shared law firm connection increases with the VC’s industry distance from the startup. For a VC that already invests in the startup’s industry group, the effect is a precisely measured zero. For a VC that invests in the startup’s sector but not its industry group, a shared law firm connection multiplies the odds of a financing by 1.30. For a VC with no recent investment in the startup’s sector, it multiplies the odds by 1.68.

Two further findings corroborate this pattern. The deal facilitation effect is larger when the startup and the VC are headquartered in different states than when they are in the same state. The effect is also larger when the law firm’s engagement by the VC is recent than when it is stale. Placebo tests using post-financing connections and mismatched law firm roles, and a horse race against the co-investor network, indicate that the shared law firm connection is neither a proxy for the parties’ other networks nor a marker of some persistent similarity between them.

Taken together, our findings indicate that law firm networks facilitate financings precisely in the

settings where information is least likely to flow between a VC and a startup on its own. A VC that already invests in the startup's industry, is located nearby, or shares co-investors with the startup has its own routes to the startup, and the law firm adds little to them. When those routes are absent, the law firm supplies one. The broad implication is that the matching of capital suppliers to capital demanders runs not only along the networks the parties build for themselves but also along networks inadvertently formed by third parties. Such third parties are present in nearly every financing transaction, in venture capital and elsewhere, so the networks they form may shape the flow of capital more widely than the direct links studied so far.

Our findings raise two questions that we leave to future work. First, we do not study whether financings that a law firm connection helped form perform differently from other financings. Their subsequent follow-on rounds and exits would indicate whether a connector law firm merely introduces a startup to a VC or also vouches for it. Second, law firms are only one of many potential inadvertent information intermediaries within one capital market, venture capital. Whether networks formed by intermediaries in other ecosystems similarly facilitate matching and capital flow in other markets is an open question.

References

- Ahern, K. R., D. Daminelli, and C. Fracassi (2015). Lost in translation? The effect of cultural values on mergers around the world. *Journal of Financial Economics* 117(1), 165–189.
- Akkus, O., J. A. Cookson, and A. Hortaçsu (2016). The determinants of bank mergers: A revealed preference analysis. *Management Science* 62(8), 2241–2258.
- Alperovych, Y., A. Divakaruni, and S. Manigart (2022). Lending when relationships are scarce: The role of information spread via bank networks. *Journal of Corporate Finance* 73, 102181.
- Bao, J. and A. Edmans (2011). Do investment banks matter for M&A returns? *Review of Financial Studies* 24(7), 2286–2315.
- Bena, J. and K. Li (2014). Corporate innovations and mergers and acquisitions. *Journal of Finance* 69(5), 1923–1960.
- Bengtsson, O. and D. H. Hsu (2015). Ethnic matching in the US venture capital market. *Journal of Business Venturing* 30(2), 338–354.
- Bottazzi, L., M. Da Rin, and T. Hellmann (2016). The importance of trust for investment: Evidence from venture capital. *Review of Financial Studies* 29(9), 2283–2318.
- Breslow, N. E. and N. E. Day (1980). *Statistical Methods in Cancer Research, Volume 1: The Analysis of Case-Control Studies*. IARC Scientific Publications No. 32. Lyon: International Agency for Research on Cancer.
- Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.
- Burt, R. S. (2002). Bridge decay. *Social Networks* 24(4), 333–363.
- Cain, M. D., S. B. McKeon, and S. D. Solomon (2020). Intermediation in private equity: The role of placement agents. *Journal of Financial and Quantitative Analysis* 55(4), 1095–1116.
- Cohen, L., A. Frazzini, and C. Malloy (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy* 116(5), 951–979.

- Engelberg, J., P. Gao, and C. A. Parsons (2012). Friends with money. *Journal of Financial Economics* 103(1), 169–188.
- Fracassi, C. (2017). Corporate finance policies and social networks. *Management Science* 63(8), 2420–2438.
- Fracassi, C. and G. Tate (2012). External networking and internal firm governance. *Journal of Finance* 67(1), 153–194.
- Gompers, P. A., V. Mukharlyamov, and Y. Xuan (2016). The cost of friendship. *Journal of Financial Economics* 119(3), 626–644.
- Hegde, D. and J. Tumlinson (2014). Does social proximity enhance business partnerships? Theory and evidence from ethnicity’s role in US venture capital. *Management Science* 60(9), 2355–2380.
- Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics* 123(2), 441–487.
- Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2007). Whom you know matters: Venture capital networks and investment performance. *Journal of Finance* 62(1), 251–301.
- Iacus, S. M., G. King, and G. Porro (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association* 106(493), 345–361.
- King, G. and R. Nielsen (2019). Why propensity scores should not be used for matching. *Political Analysis* 27(4), 435–454.
- King, G. and L. Zeng (2001). Logistic regression in rare events data. *Political Analysis* 9(2), 137–163.
- Krishnan, C. N. V. and R. W. Masulis (2013). Law firm expertise and merger and acquisition outcomes. *Journal of Law and Economics* 56(1), 189–226.
- Ljungqvist, A., F. Marston, and W. J. Wilhelm (2006). Competing for securities underwriting mandates: Banking relationships and analyst recommendations. *Journal of Finance* 61(1), 301–340.

- Manski, C. F. and S. R. Lerman (1977). The estimation of choice probabilities from choice based samples. *Econometrica* 45(8), 1977–1988.
- Moran, P. and J. A. Pandes (2019). Elite law firms in the IPO market. *Journal of Banking and Finance* 107, 105612.
- Palepu, K. G. (1986). Predicting takeover targets: A methodological and empirical analysis. *Journal of Accounting and Economics* 8(1), 3–35.
- Rider, C. I. (2009). Constraints on the control benefits of brokerage: A study of placement agents in US venture capital fundraising. *Administrative Science Quarterly* 54(4), 575–601.
- Sørensen, M. (2007). How smart is smart money? A two-sided matching model of venture capital. *Journal of Finance* 62(6), 2725–2762.
- Sorenson, O. and T. E. Stuart (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology* 106(6), 1546–1588.
- Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *Journal of Finance* 62(2), 629–668.

Figures and Tables

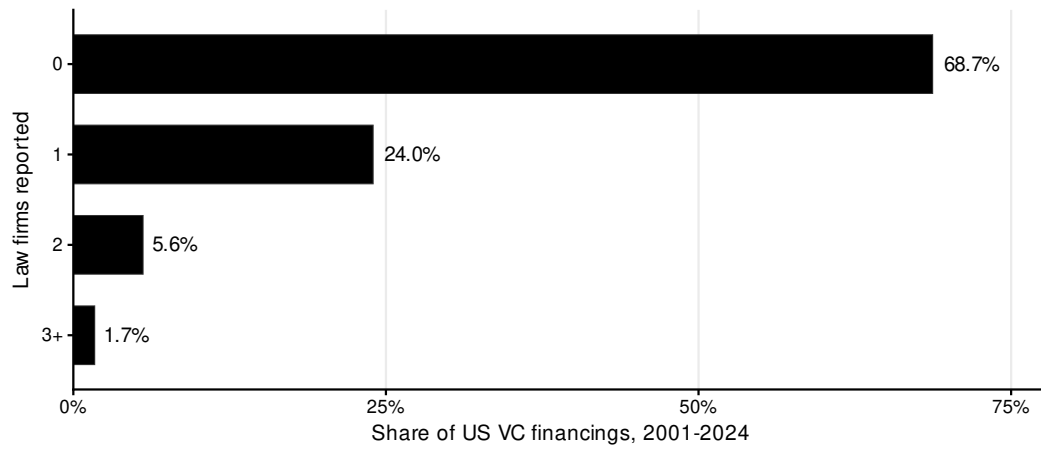


Figure 1: Number of law firms reported per financing

This figure plots the distribution of financings in the Table 1 sample by the number of distinct law firms PitchBook reports on each financing. Each bar is labeled with its share of financings.

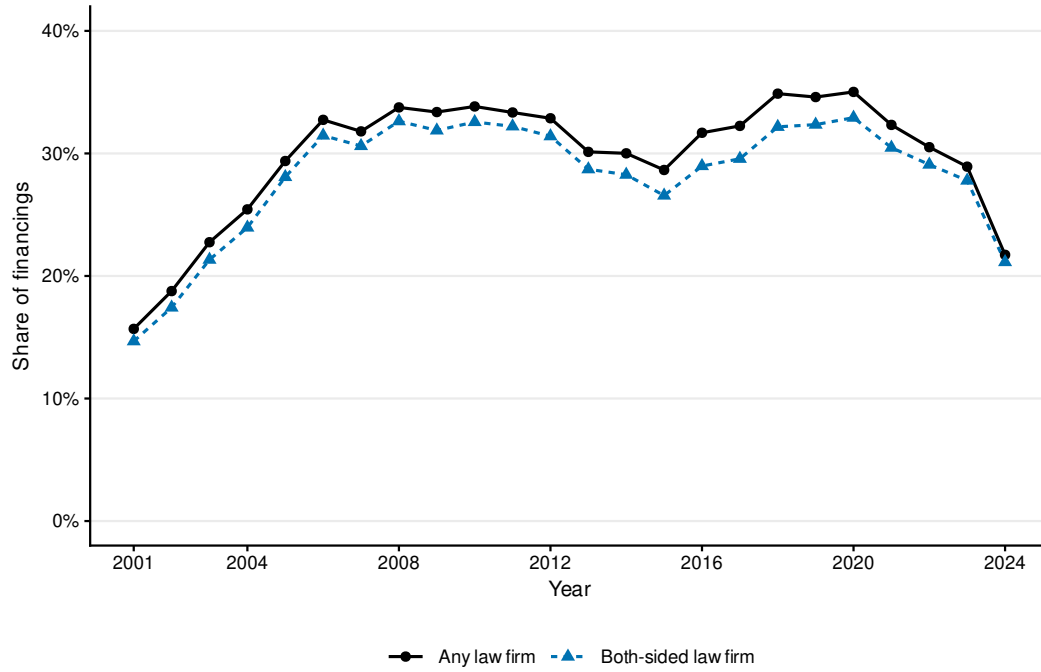


Figure 2: Law firm coverage of US VC financings by year

This figure plots, by financing year, the share of the financings in the sample of Table 1 for which PitchBook reports at least one law firm (solid line, circles) and the share advised by at least one both-sided law firm as defined in Section 2.2 (dashed line, triangles).

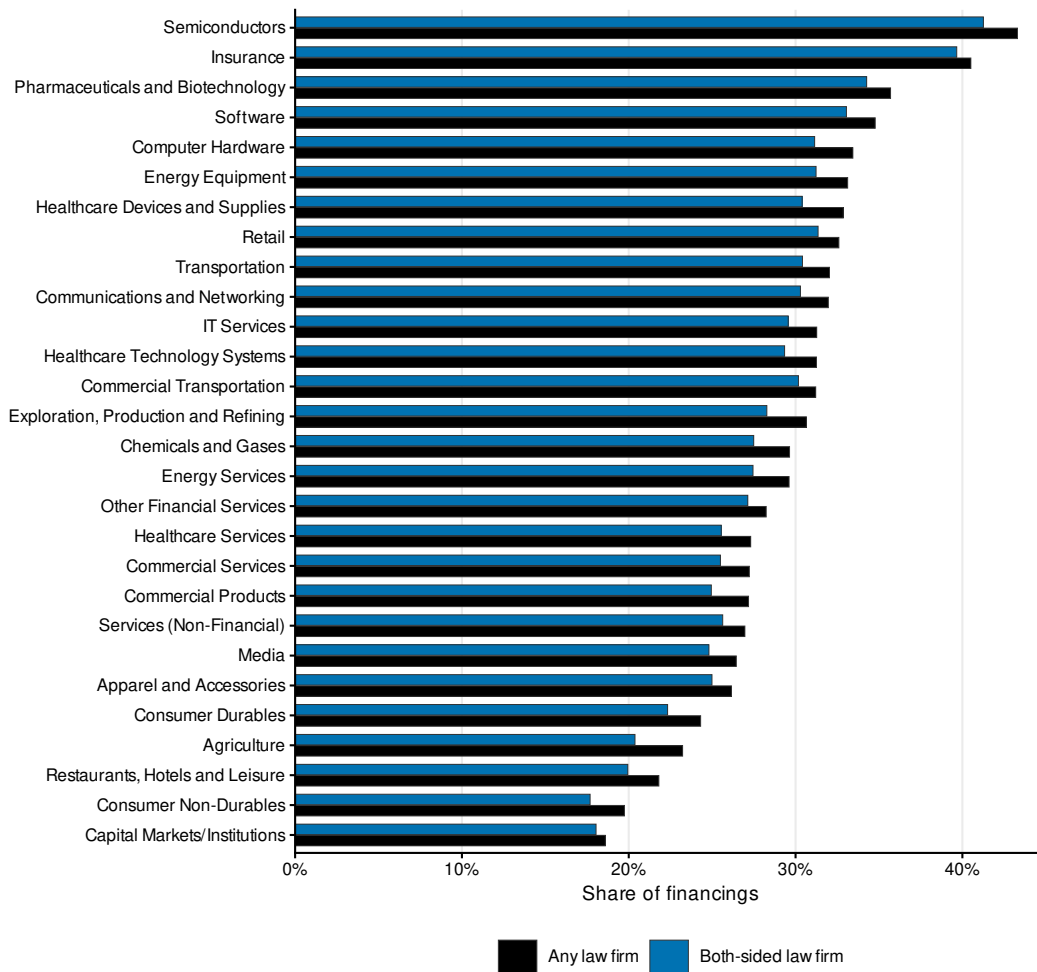


Figure 3: Law firm coverage of VC financings by industry group

This figure plots, by the startup’s primary industry group, the share of financings in the Table 1 sample for which PitchBook reports at least one law firm (black bars) and the share advised by at least one both-sided law firm as defined in Section 2.2 (blue bars). Only industry groups with at least 500 financings in the sample are shown, ordered by the share with any reported law firm.

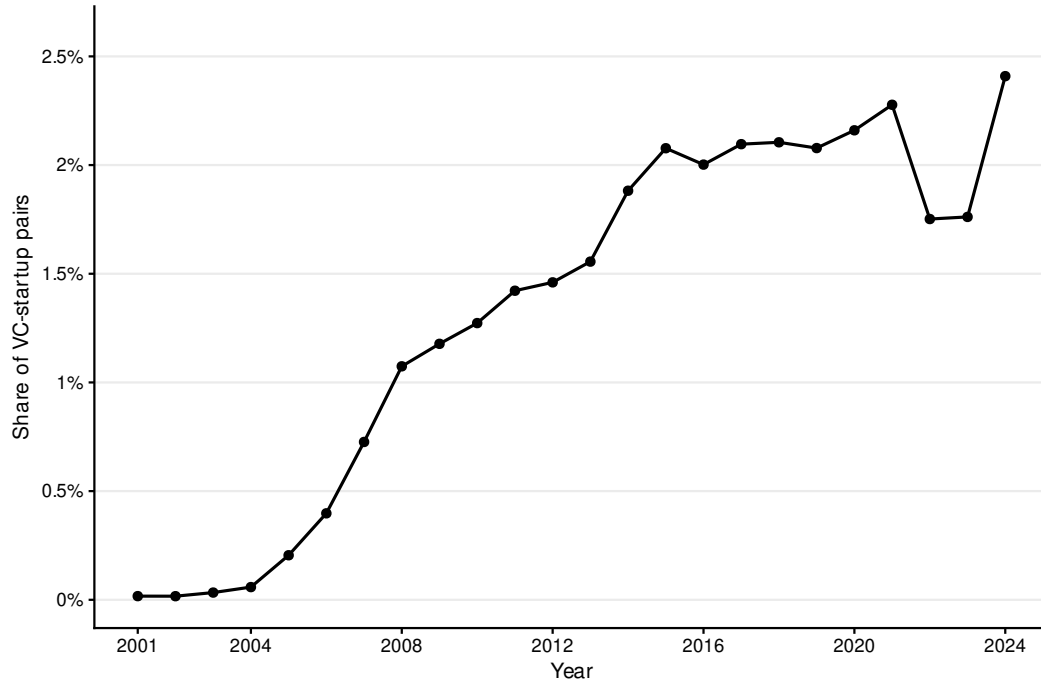


Figure 4: VC-startup pairs in realized financings bridged by a prior shared law firm, by year

This figure plots, by financing year, the share of VC-startup pairs in realized financings that are bridged by a prior shared law firm, as defined in Section 2.2. The pairs are those counted in Panel C of Table 1.

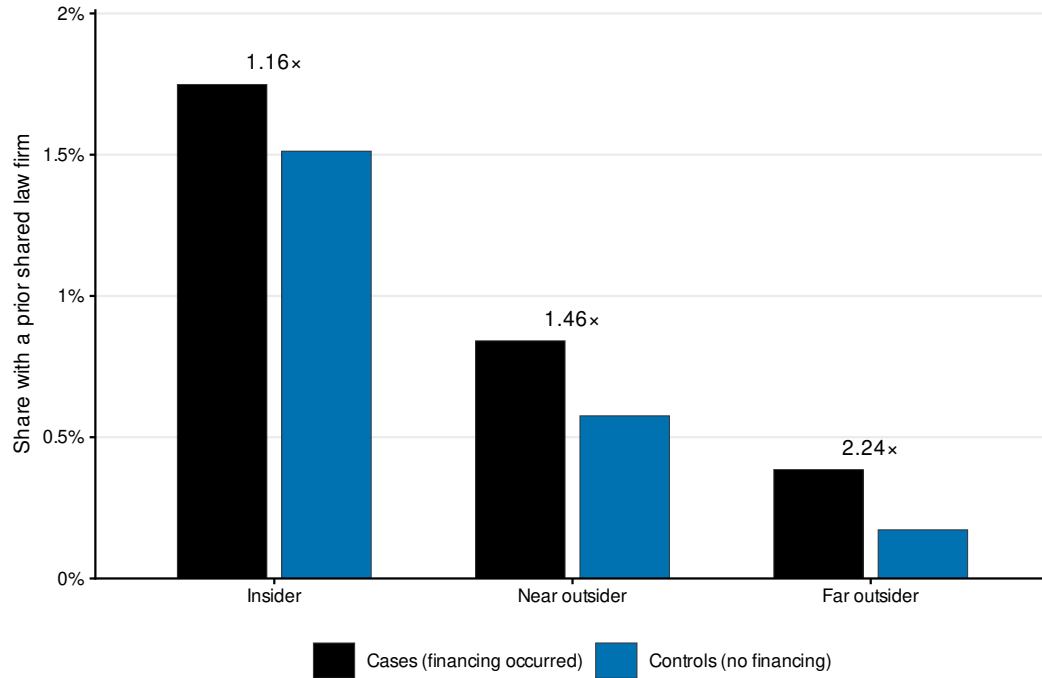


Figure 5: Share of cases and controls with a prior shared law firm, by stratum

This figure plots, within each stratum, the share of cases (black bars) and matched controls (blue bars) in the estimation panel that have a prior shared law firm as defined in Section 3.1. The label above each pair of bars is the ratio of the case rate to the control rate. Rates are unweighted and correspond to Panel B of Table 5.

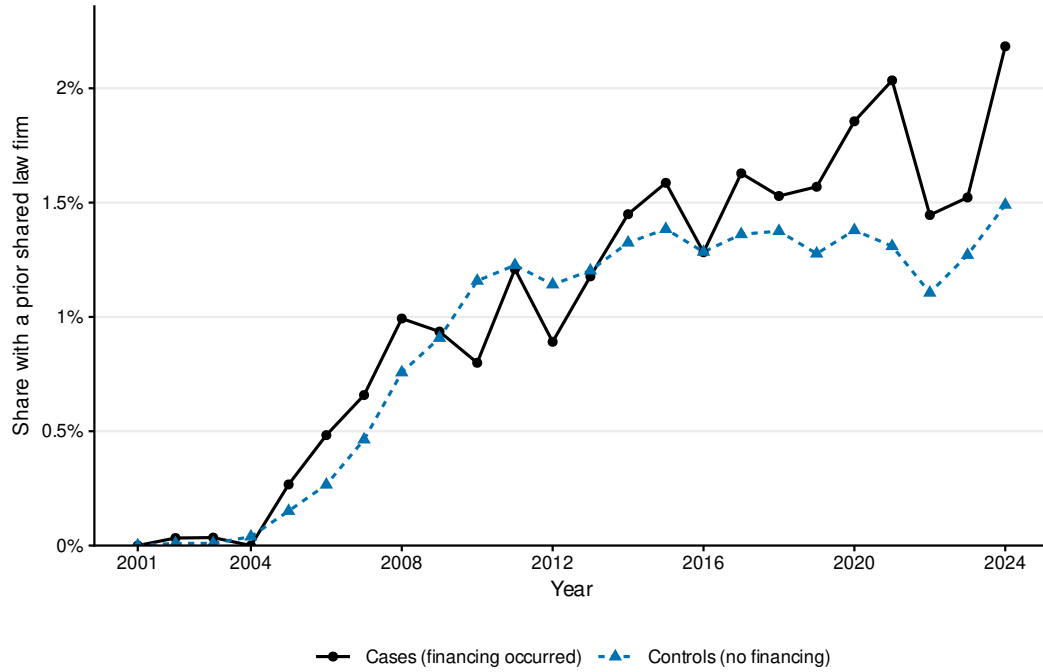


Figure 6: Share of cases and controls with a prior shared law firm, by year

This figure plots, by financing year, the share of cases (solid line, circles) and matched controls (dashed line, triangles) in the estimation panel that have a prior shared law firm as defined in Section 3.1. Rates are unweighted.

Table 1: Sample composition and law firm coverage

This table describes the sample of VC financings raised by US-headquartered startups between 2001 and 2024, drawn from PitchBook as described in Section 2.2. Panel A reports the number of financings by round type, the number of distinct startups, investors, and VC-startup pairs, the capital raised, and the median financing size. Panel B reports, for all financings and separately by deal size quartile and by round type, the share of financings for which PitchBook reports at least one law firm and the share advised by at least one both-sided law firm (Section 2.2). Size quartiles are computed over all financings with a reported deal size; the median size of each quartile is given in the row label. Panel C reports the number and share of VC-startup pairs in realized financings that are bridged by a prior shared law firm, as defined in Section 2.2, and the distribution of financings by the number of bridged pairs they contain. The financing counts in Panel C are restricted to financings with at least one identified investor.

Panel A: Sample

US VC financings, 2001–2024	205,658
Angel (individual)	22,675
Seed round	55,969
Early-stage venture	68,448
Later-stage venture	58,566
Distinct startups	94,946
Distinct investors	88,614
VC-startup pairs	586,134
Capital raised (\$ billions)	2,067
Median financing size (\$ millions)	2.7

Panel B: Share of financings with a law firm reported

	Financings	Any law firm	Both-sided law firm
All financings	205,658	31.3%	29.5%
<i>By size quartile</i>			
Q1 (median \$0.2m)	43,492	11.1%	9.1%
Q2 (median \$1.5m)	43,492	25.2%	22.8%
Q3 (median \$5.0m)	43,492	40.8%	38.8%
Q4 (median \$20.0m)	43,491	61.3%	60.0%
<i>By stage</i>			
Angel (individual)	22,675	6.4%	4.8%
Seed round	55,969	23.4%	21.6%
Early-stage venture	68,448	34.9%	33.3%
Later-stage venture	58,566	44.1%	42.1%

Panel C: Realized financing pairs bridged by a prior shared law firm

	Count	Share
VC-startup pairs	586,134	100.00%
Bridged by a prior shared law firm	10,485	1.79%
Financings with an identified investor	147,880	100.00%
No bridged pair	141,018	95.36%
One bridged pair	4,631	3.13%
Two bridged pairs	1,411	0.95%
Three or more bridged pairs	820	0.55%

Table 2: Law firms advising VC financings

This table describes the law firms that advise on at least one US financing in the sample of Table 1. A law firm's clients are the startups it represents on sell-side engagements and the investors it represents on buy-side engagements. Panel A reports the 10th, 25th, 50th, 75th, and 90th percentiles and the mean across law firms of the number of clients served, of startup clients and investor clients separately, of the number of financings advised, and of the deal value advised, which is the sum of the reported sizes of the distinct financings the law firm advised. Panel B lists the ten law firms advising the largest number of financings, with the number of financings, startup clients, and investor clients, and the deal value advised, in millions of dollars. Panel C classifies law firms by the roles they occupy across all of their engagements in the sample and reports, for each group, the number and share of law firms, the number and share of engagements, and the number of clients, financings, and deal value advised by the median law firm in the group.

Panel A: Scale of law firm practices

	p10	p25	p50	p75	p90	Mean
Clients served	1	1	1	3	12	19
Startup clients	1	1	1	3	9	15
Investor clients	0	0	0	1	3	4
Financings advised	1	1	2	5	20	43
Deal value advised (\$m)	0.3	1.6	10.0	59.8	421.4	1,460.5

Panel B: Ten largest law firms by financings advised

Law firm	Financings	Startup clients	Investor clients	Value (\$m)
Gunderson Dettmer	12,090	3,009	656	442,145
Cooley	10,059	3,084	796	351,703
Wilson Sonsini Goodrich & Rosati	8,083	2,984	419	211,935
DLA Piper	5,245	1,344	486	124,659
Goodwin Procter	5,124	1,507	516	227,744
Fenwick & West	4,864	1,684	235	191,943
Orrick	3,285	1,189	280	127,419
Latham & Watkins	1,901	644	260	117,978
Perkins Coie	1,892	725	107	37,405
Silicon Legal Strategy	1,665	284	99	37,550

Panel C: Law firms by the side of the market they advise

Roles observed	Law firms		Engagements		Median law firm		
	N	%	N	%	Clients	Deals	Value (\$m)
Sell-side only	1,399	72.3%	4,572	5.3%	1	1	4.8
Buy-side only	130	6.7%	273	0.3%	1	1	15.0
Both sides	369	19.1%	81,011	94.3%	11	16	322.9
Side not recorded	37	1.9%	39	0.0%	1	1	4.5
Total	1,935	100.0%	85,895	100.0%			

Table 3: Covariate balance on the matching variables

This table reports the balance between cases and controls on the five VC characteristics used in the coarsened exact matching of Section 3.1, within each stratum. The first two columns report unweighted means over the estimation panel. The remaining three columns report standardized differences, the difference in means between cases and controls divided by the pooled standard deviation. The unmatched column compares cases to a random draw of controls from the same stratum-eligible pools with no matching step, taken over 2,500 sampling cells per stratum and subject to the same US and first-time-pair filters as the estimation panel. The unweighted column compares cases to the matched controls with each control counted once, and the weighted column counts each control at its CEM weight. Census division is matched exactly on ten categories; its row reports the largest absolute difference in category share across the ten. Deal and portfolio counts are measured over the five years before the financing year, and VC age is years since the VC's first observed deal.

Variable	Unweighted mean		Standardized difference		
	Cases	Controls	Unmatched	Unweighted	Weighted
<i>Insider</i>					
Log VC deals, $[t - 5, t - 1]$	3.42	3.08	1.063	0.266	0.349
Log portfolio companies, $[t - 5, t - 1]$	3.29	2.92	1.082	0.304	0.380
VC age (years)	8.50	9.29	0.312	-0.140	-0.110
Institutional VC (share)	0.75	0.89	0.541	-0.355	-0.182
Corporate VC (share)	0.06	0.02	-0.155	0.217	0.111
Other VC type (share)	0.18	0.09	-0.477	0.265	0.136
Census division (largest gap)	—	—	0.230	0.142	0.089
<i>Near outsider</i>					
Log VC deals, $[t - 5, t - 1]$	2.54	2.41	0.797	0.121	0.198
Log portfolio companies, $[t - 5, t - 1]$	2.45	2.30	0.822	0.146	0.218
VC age (years)	6.61	7.03	0.087	-0.083	-0.043
Institutional VC (share)	0.70	0.79	0.481	-0.193	-0.059
Corporate VC (share)	0.06	0.03	-0.174	0.154	0.058
Other VC type (share)	0.23	0.18	-0.392	0.127	0.032
Census division (largest gap)	—	—	0.149	0.151	0.048
<i>Far outsider</i>					
Log VC deals, $[t - 5, t - 1]$	1.58	1.32	0.650	0.288	0.163
Log portfolio companies, $[t - 5, t - 1]$	1.53	1.27	0.684	0.312	0.186
VC age (years)	4.84	4.56	-0.038	0.063	-0.004
Institutional VC (share)	0.60	0.64	0.512	-0.078	-0.037
Corporate VC (share)	0.06	0.05	-0.310	0.063	0.028
Other VC type (share)	0.34	0.31	-0.313	0.050	0.024
Census division (largest gap)	—	—	0.161	0.072	0.030

Table 4: Construction and composition of the case-control panel

This table describes the estimation panel constructed in Section 3.1. Panel A reports the number of cases and controls at each step of the construction process, from all realized VC-startup-year triplets from 2000 onward to the estimation panel, followed by the number of distinct startups, VCs, startup-years, and sampling cells the estimation panel contains. Panel B reports, for each stratum, the number of cases, the number of controls, the number of controls per case, and the stratum’s share of cases. Panel C reports cases, controls, and the ratio of controls to cases by financing year. Panel D reports unweighted means and medians of the regression control variables for cases and for controls. VC deals and VC portfolio companies are counted over the five years before the financing year; startup age, prior rounds, and capital raised to date are measured at the financing; same state equals one when the startup and the VC are headquartered in the same US state.

Panel A: Sample construction

	Cases	Controls
Realized VC-startup-year triplets, 2000 onward	731,541	—
with an assignable stratum	585,674	4,023,723
US startup and US VC	397,294	2,469,770
first-time VC-startup pairs	301,930	2,465,086
Estimation panel (2001–2024)	301,930	2,465,086
distinct startups	65,988	
distinct VCs	39,754	
distinct startup-years	121,181	
sampling cells (startup-year-stratum)	160,161	

Panel B: Composition by stratum

Stratum	Cases	Controls	Controls per case	Share of cases
Insider	222,076	1,682,893	7.6	73.6%
Near outsider	38,868	352,313	9.1	12.9%
Far outsider	40,986	429,880	10.5	13.6%
Total	301,930	2,465,086	8.2	100.0%

Panel C: Cases and controls by year

Year	Cases	Controls	Ratio	Year	Cases	Controls	Ratio
2001	3,007	27,404	9.1	2013	12,570	105,853	8.4
2002	3,031	31,520	10.4	2014	15,663	119,799	7.6
2003	2,868	31,117	10.8	2015	16,578	129,253	7.8
2004	3,142	35,911	11.4	2016	13,952	121,532	8.7
2005	3,367	41,023	12.2	2017	15,910	138,227	8.7
2006	3,727	43,220	11.6	2018	18,903	155,175	8.2
2007	4,558	53,718	11.8	2019	21,538	176,620	8.2
2008	4,733	57,466	12.1	2020	24,142	182,649	7.6
2009	3,740	45,054	12.0	2021	40,648	263,252	6.5
2010	5,379	55,624	10.3	2022	36,866	247,037	6.7
2011	8,354	72,492	8.7	2023	22,728	188,541	8.3
2012	10,435	86,366	8.3	2024	6,091	56,233	9.2

Panel D: Analysis variables, cases versus controls

Variable	Cases		Controls	
	Mean	Median	Mean	Median
VC deals, [t-5, t-1]	62.32	19.00	32.12	13.00
VC portfolio companies, [t-5, t-1]	51.03	17.00	25.42	11.00
Startup age (years)	3.78	3.00	4.31	3.00
Startup prior rounds	1.14	1.00	1.25	1.00
Startup capital raised to date (\$m)	26.28	0.00	15.69	0.25
Same state (indicator)	0.44	0.00	0.37	0.00

Table 5: Shared law firm links among cases and controls

This table reports the incidence of the shared law firm indicator in the estimation panel. Panel A cross-tabulates cases and controls by whether the pair has a prior shared law firm and reports the share of each group. Panel B reports the share of cases and the share of controls with a prior shared law firm within each stratum as well as the ratio of the case rate to the control rate. All rates are unweighted.

Panel A: Shared law firm and financing occurrence

	No shared law firm	Shared law firm	Total	Shared law firm rate
Cases	297,562	4,368	301,930	1.45%
Controls	2,436,865	28,221	2,465,086	1.14%
Total	2,734,427	32,589	2,767,016	1.18%

Panel B: Shared law firm rates by stratum

Stratum	Cases	Controls	Case/control ratio
Insider	1.75%	1.51%	1.16×
Near outsider	0.84%	0.58%	1.46×
Far outsider	0.39%	0.17%	2.24×
All	1.45%	1.14%	1.26×

Table 6: Shared law firm effect on the likelihood of a financing, by stratum

This table reports estimates of Equation (1), a linear probability model in which the dependent variable equals one if VC j financed startup i in year t and zero otherwise. The sample is the estimation panel of Section 3, weighted by the CEM weights. The first block reports the coefficient on the shared law firm indicator, which is the effect for the insider stratum, and the coefficients on its interactions with the near outsider and far outsider indicators. The second block reports the implied effect within each stratum, the insider coefficient plus the relevant interaction, with standard errors computed by the delta method. The pooled row is the average of the three stratum effects weighted by each stratum's share of realized financings (0.736 insider, 0.129 near outsider, 0.136 far outsider). Column (1) includes year fixed effects only. Column (2) adds the VC controls and column (3) the startup controls described in Section 4.2, and column (4) adds an indicator for the startup and the VC being headquartered in the same state. Column (5) replaces the year fixed effects with industry code by year fixed effects, and column (6) adds VC fixed effects. Every control and fixed effect is interacted with the stratum indicators, as in Equation (1). Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Shared law firm	-0.027** (0.013)	-0.057*** (0.010)	0.015 (0.009)	0.011 (0.010)	0.007 (0.009)	0.000 (0.008)
Shared law firm $\times$ Near outsider	0.120*** (0.020)	0.135*** (0.018)	0.059*** (0.018)	0.055*** (0.018)	0.059*** (0.019)	0.066*** (0.018)
Shared law firm $\times$ Far outsider	0.170*** (0.025)	0.176*** (0.025)	0.098*** (0.025)	0.099*** (0.025)	0.111*** (0.026)	0.124*** (0.025)
<i>Implied shared law firm effect by stratum</i>						
Insider	-0.027** (0.013)	-0.057*** (0.010)	0.015 (0.009)	0.011 (0.010)	0.007 (0.009)	0.000 (0.008)
Near outsider	0.093*** (0.018)	0.078*** (0.017)	0.074*** (0.017)	0.066*** (0.017)	0.066*** (0.018)	0.065*** (0.017)
Far outsider	0.143*** (0.023)	0.119*** (0.024)	0.113*** (0.024)	0.110*** (0.024)	0.119*** (0.025)	0.124*** (0.025)
Pooled	0.011 (0.011)	-0.016* (0.009)	0.036*** (0.009)	0.032*** (0.009)	0.030*** (0.009)	0.025*** (0.007)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry $\times$ Year FE	No	No	No	No	Yes	Yes
VC FE	No	No	No	No	No	Yes
VC controls	No	Yes	Yes	Yes	Yes	Yes
Startup controls	No	No	Yes	Yes	Yes	Yes
Same-state control	No	No	No	Yes	Yes	Yes
Observations	2,767,016	2,767,016	2,767,016	2,767,016	2,766,840	2,751,193
R^2	0.007	0.048	0.060	0.063	0.073	0.276

Table 7: Shared law firm effect for cross-state and same-state pairs, by stratum

This table reports estimates from the specification of column (6) of Table 6 interacted fully with an indicator for the startup and the VC being headquartered in the same state. Each entry is an implied effect of a prior shared law firm on the likelihood of a financing, recovered from the estimated coefficients by the delta method. Within each stratum, the first column reports the effect for pairs headquartered in different states and the second for pairs in the same state. The third column reports the difference between the two, which is the coefficient on the interaction of the shared law firm indicator with the same-state indicator for that stratum. The pooled row weights the three strata by their shares of realized financings, as in Table 6. The sample, CEM weights, fixed effects, and controls are those of column (6) of Table 6. Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Stratum	Implied shared law firm effect		Difference (same – cross)
	Cross-state pairs	Same-state pairs	
Insider	0.046*** (0.009)	−0.047*** (0.010)	−0.093*** (0.011)
Near outsider	0.091*** (0.023)	0.035 (0.025)	−0.056* (0.033)
Far outsider	0.162*** (0.028)	0.054 (0.047)	−0.109** (0.053)
Pooled	0.068*** (0.009)	−0.022** (0.011)	−0.096*** (0.011)
Industry × Year FE	Yes	Yes	Yes
VC FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Observations		2,751,193	
R^2		0.276	

Table 8: Shared law firm effect by the recency of the connection

This table reports implied stratum-specific effects of a prior shared law firm on the likelihood of a financing when the shared law firm indicator is split into a recent and a stale component. For a window of W years, a pair is recently connected if some connector law firm engaged both the startup and the VC within the W years before year t , and stalely connected if it has a prior shared law firm that is not recent. Column (1) reports the unsplit indicator, which is column (6) of Table 6, so its stale rows are empty. Columns (2) to (6) each report one estimate of Equation (1) in which the two components, each interacted with the stratum indicators, replace the shared law firm indicator, for W from one to five years. Each entry is the insider coefficient plus the relevant interactions, with standard errors computed by the delta method. The sample, CEM weights, fixed effects, and controls are those of column (6) of Table 6. Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	(1) Unsplit	(2) $W=1$	(3) $W=2$	(4) $W=3$	(5) $W=4$	(6) $W=5$
<i>Insider</i>						
Recent	0.000 (0.008)	0.034*** (0.012)	0.024** (0.009)	0.015* (0.009)	0.009 (0.008)	0.007 (0.008)
Stale	—	−0.011 (0.008)	−0.025*** (0.009)	−0.031*** (0.010)	−0.029** (0.012)	−0.038*** (0.013)
<i>Near outsider</i>						
Recent	0.065*** (0.017)	0.106*** (0.033)	0.072*** (0.024)	0.064*** (0.021)	0.063*** (0.020)	0.061*** (0.019)
Stale	—	0.050*** (0.019)	0.057** (0.023)	0.069*** (0.027)	0.075** (0.031)	0.087** (0.037)
<i>Far outsider</i>						
Recent	0.124*** (0.025)	0.172*** (0.045)	0.146*** (0.032)	0.141*** (0.030)	0.144*** (0.028)	0.141*** (0.026)
Stale	—	0.108*** (0.030)	0.100*** (0.037)	0.090* (0.046)	0.058 (0.057)	0.035 (0.072)
Industry $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
VC FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations			2,751,193			
R^2	0.276	0.276	0.276	0.276	0.276	0.276

Table 9: Shared law firm effect by the recency of the startup-side and VC-side engagements

This table reports implied stratum-specific effects of a prior shared law firm on the likelihood of a financing from a single estimate of Equation (1) in which the shared law firm indicator is replaced by four mutually exclusive indicators, each interacted with the stratum indicators. The four indicators classify a connected pair by whether a connector law firm engaged the startup within the three years before year t and, separately, whether one engaged the VC within the three years before year t . Rows give the four combinations and columns the three strata. Each entry is the insider coefficient plus the relevant interactions, with standard errors computed by the delta method. The sample, CEM weights, fixed effects, and controls are those of column (6) of Table 6. Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Recency of each side	Insider	Near outsider	Far outsider
Startup recent, VC recent	0.015* (0.009)	0.064*** (0.021)	0.143*** (0.029)
Startup recent, VC stale	-0.040*** (0.011)	0.059* (0.033)	0.099* (0.052)
Startup stale, VC recent	0.016 (0.020)	0.064 (0.059)	0.155 (0.124)
Startup stale, VC stale	-0.034 (0.029)	0.156 (0.097)	-0.145 (0.177)
Industry $\times$ Year FE	Yes	Yes	Yes
VC FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Observations		2,751,193	
R^2		0.276	

Table 10: Shared law firm effect controlling for co-investor overlap

This table reports estimates of Equation (1) augmented with co-investor overlap, defined in Section 4.4 as the number of third-party investors that both financed the startup and co-invested with the VC over the five years before year t . Column (1) repeats column (6) of Table 6. Columns (2) and (3) add the log of one plus overlap and an indicator for any overlap, respectively, each with a single slope across strata. Column (4) retains the log overlap slope and drops the shared law firm terms. Column (5) enters log overlap interacted with the stratum indicators and omits the shared law firm terms, and column (6) includes both the shared law firm terms and the stratum-specific overlap slopes. The first block reports the shared law firm coefficients, the second and third blocks the overlap coefficients, and the fourth block the implied shared law firm effect within each stratum, which is the insider coefficient plus the relevant interactions, with standard errors computed by the delta method. The pooled row weights the three strata by their shares of realized financings, as in Table 6. The sample, CEM weights, fixed effects, and controls are those of column (6) of Table 6. Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Shared law firm (regression coefficients)</i>						
Shared law firm	0.000 (0.008)	-0.035*** (0.008)	-0.021*** (0.008)	—	—	-0.032*** (0.008)
Shared law firm × Near outsider	0.066*** (0.018)	0.069*** (0.018)	0.067*** (0.018)	—	—	0.047*** (0.018)
Shared law firm × Far outsider	0.124*** (0.025)	0.136*** (0.026)	0.127*** (0.026)	—	—	0.111*** (0.027)
<i>Co-investor overlap, single slope</i>						
Log co-investor overlap	—	0.116*** (0.004)	—	0.115*** (0.004)	—	—
Any co-investor overlap	—	—	0.129*** (0.004)	—	—	—
<i>Co-investor overlap (log, by stratum)</i>						
Insider	—	—	—	—	0.104*** (0.004)	0.105*** (0.004)
Near outsider	—	—	—	—	0.188*** (0.005)	0.187*** (0.005)
Far outsider	—	—	—	—	0.225*** (0.006)	0.225*** (0.006)
<i>Implied shared law firm effect by stratum</i>						
Insider	0.000 (0.008)	-0.035*** (0.008)	-0.021*** (0.008)	—	—	-0.032*** (0.008)
Near outsider	0.065*** (0.017)	0.034** (0.017)	0.046*** (0.017)	—	—	0.015 (0.017)
Far outsider	0.124*** (0.025)	0.101*** (0.025)	0.107*** (0.025)	—	—	0.080*** (0.026)
Pooled	0.025*** (0.007)	-0.008 (0.007)	0.005 (0.007)	—	—	-0.011 (0.007)
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
VC FE	Yes	Yes	Yes	Yes	Yes	Yes
VC controls	Yes	Yes	Yes	Yes	Yes	Yes
Startup controls	Yes	Yes	Yes	Yes	Yes	Yes
Same-state control	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,751,193					
R^2	0.276	0.283	0.282	0.283	0.284	0.284

Table 11: Shared law firm effect by the role of the law firm on each side of the connection

This table reports implied stratum-specific effects on the likelihood of a financing for four versions of the shared law firm indicator that differ in the role through which the connector law firm is linked to each side. On the startup side, the law firm either advised the startup itself on an earlier deal or advised an investor in one of the startup's earlier financings. On the VC side, the law firm either acted as the VC's own buy-side counsel or was sell-side counsel on a deal in which the VC participated. Column (1) uses the indicator of Section 3.1 and repeats column (6) of Table 6; columns (2) to (4) replace it with the indicator named in the column header. Each column is a separate estimate of Equation (1). Each entry is the insider coefficient plus the relevant interaction, with standard errors computed by the delta method, and the pooled row weights the three strata by their shares of realized financings, as in Table 6. The sample, CEM weights, fixed effects, and controls are those of column (6) of Table 6. Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
Startup side:	Advised startup	Advised its investor	Advised startup	Advised its investor
VC side:	Own counsel	Own counsel	Counsel on VC's deal	Counsel on VC's deal
Insider	0.000 (0.008)	0.020** (0.010)	0.003 (0.003)	0.007* (0.004)
Near outsider	0.065*** (0.017)	0.082*** (0.018)	0.048*** (0.005)	0.037*** (0.007)
Far outsider	0.124*** (0.025)	0.096*** (0.028)	0.062*** (0.007)	0.040*** (0.009)
Pooled	0.025*** (0.007)	0.038*** (0.009)	0.017*** (0.003)	0.015*** (0.004)
Industry $\times$ Year FE	Yes	Yes	Yes	Yes
VC FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations			2,751,193	
R^2	0.276	0.276	0.276	0.276

Table 12: Shared law firm connections formed before and after the financing

This table reports the coefficient on a shared law firm indicator built over the three years before the financing year, in the first column, and over the three years after it, in the second. The two indicators are constructed identically: each requires a connector law firm to have engaged the startup on the sell side and the VC on the buy side within the window, under the same participation exclusions as the indicator of Section 3.1, and they differ only in the direction of the window. Each entry comes from a separate estimate of the specification of column (6) of Table 6 within the stratum, with the CEM weights, fixed effects, and controls of that column. The sample is restricted to financing years 2003 to 2021, the years for which a full three-year window is observable in both directions. The number of observations and the R^2 are reported for each within-stratum estimate. Standard errors, clustered by startup and by VC, are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Stratum	Shared law firm connection formed in	
	$[t - 3, t - 1]$	$[t + 1, t + 3]$
Insider	0.020** (0.009)	-0.008 (0.007)
Observations	1,457,859	
R^2	0.254	0.253
Near outsider	0.076*** (0.023)	0.038*** (0.014)
Observations	299,431	
R^2	0.296	0.296
Far outsider	0.161*** (0.033)	0.055*** (0.018)
Observations	375,087	
R^2	0.417	0.417
Industry $\times$ Year FE	Yes	Yes
VC FE	Yes	Yes
Controls	Yes	Yes

Appendix

A Equivalence Test: Setting Bounds for Effect Sizes

Table 6 reports that a shared law firm has no effect on the probability of a financing when the VC already invests in the startup’s industry group. This appendix shows the precision with which that zero effect is estimated.

A conventional statistical test cannot confirm a zero effect; it can only fail to reject one. Given that the insider stratum holds nearly three quarters of the financings in our sample, failure to reject is a weak position. Here, we improve on that by inverting the null using equivalence testing. Equivalence testing fixes a bound Δ and tests the hypothesis that the effect is at least Δ in magnitude. Rejecting that hypothesis allows us to claim that the effect is smaller than Δ in magnitude.

Table A1 reports, for each stratum, the smallest bound at which a two one-sided test (TOST) rejects the hypothesis that the effect is at least Δ in magnitude at a five percent level. For insiders, this bound is 0.013. The data therefore rule out an insider effect larger than about a tenth of the far outsider effect, in either direction. This provides a more powerful restatement of the insider result: using our large sample, we can posit that a shared law firm does not meaningfully increase the likelihood of a financing between a startup and a VC if the VC already invests in the startup’s sector, rather than having to posit that we simply fail to find such an effect.²⁷ The same table reports the mirror image for the outsider strata: at the five percent level, the near outsider effect is at least 0.037 and the far outsider effect at least 0.083.

Table A1: Equivalence bounds and one-sided lower bounds on the shared law firm effect, by stratum

This table reports, for each stratum, the estimate and standard error of the shared law firm effect from the specification of column (6) of Table 6 estimated within the stratum, together with two bounds. It denotes the estimate by $\hat{\theta}$ and its standard error by SE. The equivalence bound Δ^* is the smallest value of Δ at which a two one-sided test rejects, at the 5% level, the hypothesis that the effect is at least Δ in absolute value; it equals $|\hat{\theta}| + 1.645 \times \text{SE}$. The one-sided lower bound is the 5% one-sided lower confidence bound on the effect, $\hat{\theta} - 1.645 \times \text{SE}$. The number of observations and the R^2 refer to the within-stratum estimate. Standard errors are clustered by startup and by VC.

Stratum	Estimate	SE	Equivalence bound Δ^*	One-sided lower bound	Observations	R^2
Insider	0.000	(0.008)	0.013	−0.013	1,898,915	0.251
Near outsider	0.065	(0.017)	0.093	0.037	384,980	0.283
Far outsider	0.124	(0.025)	0.165	0.083	467,298	0.404

²⁷The bound is the smallest Δ at which a two one-sided test rejects, $|\hat{\theta}| + 1.645 \times \text{SE}$, where $\hat{\theta}$ is the stratum’s estimated effect. It is by construction the weakest version of the argument, since it is based on the data rather than fixed in advance. Against a threshold of 0.012 fixed in advance as a tenth of the far outsider effect, the two one-sided test gives $p = 0.063$ and falls just short at the five percent level; at 0.015 it gives $p = 0.027$. We report both rather than choosing the threshold that passes.