NBER WORKING PAPER SERIES

ONLINE SYNDICATES AND STARTUP INVESTMENT

Christian Catalini
Xiang Hui

Working Paper 24777
http://www.nber.org/papers/w24777

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018, Revised January 2019

We thank Avi Goldfarb, Kevin Laws, Meng Liu, Hong Luo, Ramana Nanda, Scott Stern and Jane Wu for helpful discussions and comments. The researchers acknowledge the support of the Junior Faculty Research Assistance Program at the MIT Sloan School of Management, and the MIT Initiative on the Digital Economy (http://ide.mit.edu/). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Online Syndicates and Startup Investment
Christian Catalini and Xiang Hui
NBER Working Paper No. 24777
June 2018, Revised January 2019
JEL No. G24,L26,O31,O33

ABSTRACT

Early crowdfunding platforms were based on a premise of disintermediation from professional investors, and relied on the 'wisdom of the crowd' to screen high quality projects. This becomes problematic when equity is involved, as the degree of asymmetric information between entrepreneurs looking for funding and the crowd is higher than in reward-based crowdfunding. As a result, platforms later experimented with incentives for professional investors to curate deals for crowd. We study how the introduction of such incentives influenced the allocation of capital on the leading US platform, finding that the changes led to a sizable 33% increase in capital flows to new regions. Professional investors use their reputation to vouch for high potential startups that would otherwise be misclassified because of information asymmetry. This allows them to arbitrage opportunities across regions and shift capital flows to startups that are 37% more likely to generate above median returns. At the same time, this 'democratization effect' relies on the presence of intermediaries with professional networks that bridge these new regions to California. Using a large-scale field experiment with over 26,000 investors we further unpack the frictions to online investment, and show that social networks constitute a key barrier to additional democratization, since they influence how the crowd evaluates intermediaries in the first place.

Christian Catalini
MIT Sloan School of Management
100 Main Street, E62-480
Cambridge, MA 02142
and NBER
catalini@mit.edu

Xiang Hui
Washington University in St Louis
and MIT
hui@wustl.edu

1 Introduction

Early crowdfunding platforms (e.g. Kickstarter, Indiegogo) were based on a premise of complete disintermediation from professional investors: the crowd would directly fund projects based on the information shared online by the entrepreneurs, bypassing in the process traditional gatekeepers. This approach becomes problematic when equity is involved because of asymmetric information between entrepreneurs and investors. Moreover, it favors geographic regions that already attract a disproportionate share of capital offline, as online investors free ride on others and rely on highly visible (but imperfect) proxies for quality such as accumulated capital, formal education, affiliation with top accelerator programs, prominent advisors, etc. Online investors have also little incentive to perform costly due diligence because they invest small amounts and all receive the same investment terms. Traditional solutions to this problem used in other online markets are ineffective, as entrepreneurs may lack an established track record, and transactions are too rare, idiosyncratic and uncertain to support a robust reputation or feedback system.

While platforms could invest in curation and become third-party certifiers, this approach does not scale, as information about investment opportunities is dispersed across different domains of expertise and regions, and due diligence is labor-intensive and time-consuming. This has led to the emergence of an alternative market design approach which borrows ideas from the venture capital model: the creation of a market for curation. By allowing experienced, professional investors to screen startups and invite the crowd to participate in their deal flow in exchange for a share of future profits (the 'carry'),[1] platforms can avoid the unraveling of the market, and align incentives between the crowd, the startups and the experts. This model resembles the relationship between general and limited partners in a venture capital firm, with the crowd replacing limited partners as the source of funding. It also has the advantage of turning what would otherwise be a one-time relationship between online investors and startups into a repeated

[1] In reward-based equity crowdfunding (e.g. Kickstarter, Indiegogo), platforms typically earn a fee on the total amount successfully raised through the website, irrespective of long run outcomes. Such a model is particularly ineffective when equity is involved, as it does not incentivize platforms to surface and match only high quality deals, but encourages them to increase the total volume of transactions, irrespective of quality. The use of a carry to incentivize curation is borrowed from the venture capital model. In addition to carried interest, venture capital firms also charge management fees on the capital they invest on behalf of limited partners.

one between online investors and the intermediaries syndicating the deals. Intermediaries can then leverage their reputation – often established offline over multiple years of activity as angel investors – to arbitrage investment opportunities and vouch for startups that have high potential, but that would otherwise struggle to raise capital online.

The arrival of these intermediaries, however, may reinforce the agglomeration of capital flows, as most angel investors are based in top entrepreneurial hubs such as California[2] and primarily invest in their home region.

The objective of this paper is to explore the trade-offs the introduction of intermediaries in equity crowdfunding markets entails, and identify some of the key frictions that prevent high quality entrepreneurs from accessing capital online. We start by developing a simple theoretical model to compare how investors and startups in top entrepreneurial hub regions versus non-hub regions are affected by intermediaries, and then use novel data on capital flows and deal performance from the leading US equity crowdfunding platform to test our predictions.

When investment opportunities are directly posted online and the crowd makes its own investment decisions (as in reward-based crowdfunding), we find that non-hub regions severely underperform in terms of capital attracted relative to their share of startup activity. Consistent with information asymmetry being the culprit, this performance gap is substantially smaller for non-hub startups that have high, observable signals of quality.[3]

The introduction of intermediaries (called syndicate leads on the platform) reverses this trend, increasing capital flows to non-hub regions by 33%. While our main specification relies on an event study around the months immediately preceding and following the launch of the new feature (which was not pre-announced and can be therefore considered a natural experiment on the existing crowd of investors), the result is robust to multiple empirical approaches and alternative definitions of the relevant control group.[4]

[2]In the paper, hubs refer to geographic regions with a high concentration of startup activity. Empirically, in the high tech software sector we study, California is the main hub region. The next region by startup and investment activity is New York, although this region is more similar in activity levels to the third one, Massachusetts, than to California. In the data, we adopt a conservative approach and only classify California as a hub. Adding New York (or even Massachusetts) to our definition of hub regions does not change our results.

[3]We build measures of observable startup quality based on the startups' profiles on the platform. See Section 4 for more details.

[4]Our main specification controls for investor heterogeneity by introducing investor fixed-effects and focusing

This 'democratization effect' however, relies on the presence of a particular type of intermediary that, independent of location, has a 'hybrid' professional network on the platform that includes individuals from both hub and non-hub regions. These intermediaries can be either located in California but have extensive professional connections[5] to individuals located in a non-hub region, or can be located in a non-hub region but have extensive professional connections to individuals based in California.

Furthermore, consistent with intermediaries identifying high potential startups that would otherwise be misclassified and using their reputation to vouch for them on the platform, their role in shaping capital flows is particularly pronounced for startups that have ex-ante low, observable quality. When using an intermediary, online investors are also less constrained by geography or their networks when making their decisions, a result that is particularly pronounced when the intermediary is of high observable quality.[6]

To better understand how investors evaluate intermediaries on the platform, we also ran a large-scale field experiment. In the experiment, we contacted over 26,000 potential investors and presented them with different information about a random triplet of intermediaries.[7] In addition to a control group that only featured the intermediary's name, picture and profile link, we included treatments highlighting the intermediary's past performance, endorsements, professional network, and direct connections to the focal investor.[8] The experiment led to two key findings: 1) online investors do not pay attention to intermediaries' past performance and endorsements, and are more likely to explore the intermediaries' profiles when shown information

on within-investor variation. Results are robust to including investors that join or abandon the platform after intermediaries are introduced, to using propensity score matching, and a difference-in-differences estimation combined with matching.

[5]We define a professional connection as a reciprocal connection between two individuals on the platform's social network (similar to a reciprocal follow on Twitter, or a connection on Facebook or LinkedIn).

[6]Intermediaries with low, observable quality are instead mostly chosen by investors located in, or with extensive connections to their home region. Intermediaries are evaluated by the platform through an extensive review process. We use the scores resulting from the platform's assessment to separate intermediaries by quality. See Section 4 for more details.

[7]Messages were sent to individuals who had not invested before in two batches: 40% of potential investors received the email in October 2017, and the remaining 60% the following month. Each message featured a random triplet selected from a total pool of 21 intermediaries.

[8]Following a common approach in online platforms, we listed the number of connections of an intermediary as a summary measure for their professional network (connections treatment); featured past investments with their outcomes (portfolio treatment); showed a reference from an industry insider (reference treatment); and displayed how far the intermediary was on the platform's social graph from the focal investor (degree treatment).

about the size of the intermediary's professional network on the platform; 2) if the investor and the intermediary's professional networks do not overlap on the platform (1st or 2nd degree connection), investors ignore the information they are presented with. Taken together, these results suggest that investors may see the intermediary's professional network as a valuable proxy for their ability to access and screen opportunities. Social networks may also constitute a key obstacle for expanding access to capital to entrepreneurs and regions that are otherwise not well-connected. By paying more attention to intermediaries that have large professional networks and that are a 1st or 2nd degree connection, online investors may reinforce path-dependency in capital flows.

Last, to see if intermediaries are arbitraging opportunities across regions by matching capital to areas where its marginal return is higher, we introduce startup valuation data. These are typically difficult to obtain in a systematic and reliable way, as they capture private information about a startup's progress after funding has occurred. We take advantage of a thorough auditing process the platform conducted for its shareholders to classify deals based on their returns.[9] The valuation data confirms that intermediaries are indeed finding 'gems' in new regions: startups from non-hub regions that are featured by an intermediary are 36.9% more likely to have above median returns relative to regular deals on the platform, and this increases to 68.5% for startups of ex-ante low, observable signals of quality. Expanding access to capital to new regions may not only allow new types of high quality entrepreneurs to develop their ideas, but also represents a profitable endeavor because of the less saturated nature of these markets.

Taken together, our results suggest that when equity crowdfunding platforms simply adopt the market design of reward-based platforms like Kickstarter or Indiegogo, they are unlikely to meaningfully expand access to capital to new types of entrepreneurs and regions. Because of information asymmetry and the lack of incentives for investors from the crowd to perform due diligence, the matches they facilitate will predominantly involve startups that have high, observable signals of quality. Whereas intermediaries can counterbalance this trend, only those with professional networks that bridge hub and non-hub regions may be able to effectively do

[9] We categorize returns by cohort, since younger deals have a shorter period of time to update their valuation (both positively and negatively).

so. Social networks influence how online investors decide which intermediaries to pay attention to and how they evaluate them. This can reinforce pre-existing agglomeration, limiting the role intermediaries can play in arbitraging opportunities across regions.

The paper proceeds as follows: In the next section we review the relevant literature. In Section 3 we introduce a simple theoretical framework to guide our empirical predictions. Section 4 describes the data and empirical strategy. Section 5 discusses our results, and Section 6 concludes.

2 Related Literature

The diffusion of the internet, by reducing the frictions that communication, transportation and search costs impose on economic transactions (Forman et al., 2018), has been generally associated with improvements in the access to opportunities, products and services. At the same time, the impact of these changes has been often uneven, as some regions have benefited disproportionately from them while others have lagged behind, increasing inequality. In particular, the internet, despite its lowering of transaction costs across distance, seems still to be largely constrained by 'gravity'[10] and geographic frictions (Blum and Goldfarb, 2006).

Crowdfunding platforms[11] have become increasingly important for entrepreneurial endeavors from the arts to technology (Agrawal et al., 2014; Mollick and Nanda, 2016; Belleflamme et al., 2014). Despite this growth, it is not clear if they have been able to overcome geographic frictions. As platforms have scaled, researchers have found mixed evidence about their ability to actually

[10] Gravity is a seminal concept from international trade where economic transactions are largely concentrated between similar, neighboring economies (Tinbergen, 1962). Whereas in the context of trading of physical goods, this may be driven by distance-related transaction costs, Blum and Goldfarb (2006) find that geographic distance still plays a significant role in shaping online transactions where the cost of transportation, time and distribution are zero. A closely related concept is home bias, where investors hold less than optimal amounts of foreign equity (French and Poterba, 1991; Coval and Moskowitz, 1999).

[11] The literature on crowdfunding has been rapidly expanding in the last years. Mollick (2014) provides an early overview of the funding dynamics on crowdfunding platforms, and Agrawal et al. (2014) cover some of the economic trade-offs entrepreneurs and investors face when moving deals online. A core finding of the literature is that funding accelerates with accumulated capital, since the crowd uses it as a proxy for the quality of projects, leading to herding (Agrawal et al., 2015; Zhang and Liu, 2012; Kuppuswamy and Bayus, 2013). Herding effects can be counteracted by public goods-concerns (Burtch et al., 2013). Mollick and Nanda (2016) examine the "wisdom of the crowd" and show that in the arts the crowd makes similar decisions to traditional experts. Also related to information, Burtch et al. (2015) show that privacy concerns influence funding decisions.

expand access to capital and substitute for traditional sources of early stage funding, with empirical papers finding both support for crowdfunding reducing pre-existing agglomeration (Sorenson et al., 2016; Vulkan et al., 2016; Agrawal et al., 2018, 2016), as well as reinforcing agglomeration (Kim and Viswanathan, 2014; Lin and Viswanathan, 2015). Sorenson et al. (2016) show that crowdfunding not only enables more reward-based projects in US counties that are typically not known for their inventive activity, but that this is also positively correlated with follow on activity from traditional VCs. Similarly, Agrawal et al. (2018) rely on college breaks to show that a large share of high quality crowdfunding projects on Kickstarter are led by college students - a demographic that has high human capital, but that in non-hub regions may otherwise be excluded from traditional sources of funding. However, Lin and Viswanathan (2015) find that on peer-to-peer lending platform Prosper capital is largely concentrated between borrowers and lenders from the same state.

Overall, this raises the question of whether and how the market design of crowdfunding platforms influences capital flows and which types of entrepreneurs and regions can benefit from it. Recent work has found that the quality and quantity of crowdfunded investments can be increased through different information mechanisms, such as: providing product certifications together with signals of customer traction or accumulated investment (Bapna, 2017); revealing investors' past fundraising and investment activity on the platform (Kim and Viswanathan, 2018); emphasizing founding team information (Bernstein et al., 2017); and providing equity and financial information (Ahlers et al., 2015).[12] We build on this emerging literature by studying a novel change in market design: the introduction of intermediaries with high powered incentives to perform due diligence and reduce information asymmetry on the platform. We also explore the trade-offs the change entails in terms of capital flows across regions, and for different types of entrepreneurs with high versus low observable signals of quality.

The paper also contributes to the extensive literature on the market design and development of trust mechanisms in online, peer-to-peer (P2P) platforms (see Einav et al. (2016) for a compre-

[12] Relatedly, in the peer-to peer lending context, Lin et al. (2013) find that online friendships between borrowers can be used by lenders to infer quality, and Liu et al. (2015) find that offline friendships between lenders influence lenders' investment decisions.

hensive overview). P2P platforms typically rely on three types of mechanisms to establish trust between the different sides of their markets: 1) past reputation (see the large body of research surveyed in Tadelis (2016)); 2) third-party certification by trusted institutions (e.g., Dranove and Jin (2010); Elfenbein et al. (2015); Hui et al. (2017)); and 3) warranties (Grossman (1981); Roberts (2011); Hui et al. (2016)). On crowdfunding platforms, however, these mechanisms are less likely to be effective. Entrepreneurs, have limited past reputation, and investment is often a one-shot game with a long lag between the investment decision and a liquidity event.[13]

Under such conditions, we show that online platforms can bootstrap their reputation systems by borrowing from the offline markets they are trying to complement or replace. In particular, we find that the introduction of intermediaries in the form of online syndicates can serve as another trust-building mechanism for at least two reasons. First, it allows startups without an established reputation to leverage the reputation of an intermediary that can interact repeatedly with the crowd to raise capital in a situation where the crowd faces an extremely high information asymmetry problem. Second, it allows investors to use information they can access through their professional networks and through shared contacts to evaluate intermediaries in the first place. This mechanism is similar to the one described by Holtz et al. (2017). Although rarely seen on P2P platforms, the use of syndication is common in offline markets for early stage capital, such as the syndication of venture capital investments (e.g., Gompers (1995); Brander et al. (2002); Kaplan and Strömberg (2003), and the syndication of financial products (e.g. Peek and Rosengren (1997); Ivashina and Scharfstein (2010); Giannetti and Laeven (2012)).

Last, our paper also relates to a large body of literature that explores how geographic distance and social networks affect the intensity of economic activity in different settings. Distance between angel investors (or venture capitalists) and entrepreneurs has been repeatedly shown to constitute an obstacle to early stage investment (e.g., Lerner (1995); Cumming and Dai (2010); Chen et al. (2010); Lin and Viswanathan (2015)). Home bias and geographic frictions also influence investment decisions in more mature markets such as stock and equity markets (e.g., Coval

[13]Certification and warranties are also difficult to provide because startup investments are inherently uncertain, and establishing a reputation as a certification agent requires a long period of time: among venture-funded startups, which are typically in a later stage of the funnel relative to the angel-funded ones we study, it takes on average 7 years to reach an acquisition and 8.25 years for an IPO.

and Moskowitz (1999); Cooper and Kaplanis (1994); Huberman (2001); Van Nieuwerburgh and Veldkamp (2009)), as well as overall trade flows even in the presence of digital technology (e.g., Blum and Goldfarb (2006); Hortaçsu et al. (2009); Hui (2018)). Interestingly, and consistent with some of our findings, social connections of various forms have been shown to counterbalance geographic frictions, whether it is through networks of family and friends (Agrawal et al. (2015)), professional connections between investors and entrepreneurs (Hochberg et al. (2007); Hsu (2007); Cumming and Johan (2013)), diaspora networks (Nanda and Khanna (2010)), online communities (Mollick (2014)), or professional, offline syndication networks (Sorenson and Stuart (2008); Lerner (1994)).

3 Theoretical Framework

The objective of this section is to provide a simple framework for describing how the introduction of intermediaries on equity crowdfunding platforms may influence the allocation of capital across regions. We start by characterizing direct online investments by investors from the crowd, and by comparing the optimal investment decisions of different investors based on their ability to access investment opportunities and to perform due diligence on startups located in entrepreneurial hub versus non-hub regions. We then allow for intermediaries in the form of syndicate leads that scout and curate deals on behalf of the crowd in exchange for a share of future returns. The theoretical framework guides our empirical predictions and helps us identify some of the key mechanisms that may drive changes in the investment behavior of the crowd when intermediaries are introduced.

Throughout the paper, we refer to investors from the crowd as investors, and to deals surfaced by intermediaries as syndicated deals. In syndicated deals, investors only select intermediaries, who then select which startups to invest in on behalf of the investors.

3.1 Direct Investment by the Crowd

We model individual investment decisions using a static, single-agent optimization framework. Investors indexed by i are either based in a top entrepreneurial region (hub, or $L_i = H$) or in a peripheral region (non-hub, or $L_i = NH$). As a result of agglomeration, hub regions are assumed to have, on average, higher quality startups: i.e. if we were to randomly draw a startup from a hub versus a non-hub region, the quality of the former is likely to be higher because of Marshallian agglomeration economies (economies of scale, labor market pooling, knowledge spillovers). Empirically, in the high tech software sector we study, only California is clearly a hub region. The next region by startup and investment activity is New York, although this region is more similar in activity levels to the third one, Massachusetts, than to California.[14] In the data, we adopt a conservative approach and only classify California as a hub. Adding New York (or even Massachusetts) to our definition of hub regions does not change our results.

Investors are profit maximizers, and their returns depend on three factors: their access to deals, ability to perform due diligence, mentoring and monitoring costs. Every period, investors observe their investment parameters and decide to either invest in a startup or not to invest. Conditional on investment, their return from investing directly (i.e. without an intermediary) is given by:

$$\Pi_i^D = max\{\gamma(n_i^H), \gamma(n_i^{NH}), \rho\} - \kappa_1 d_{id}$$

The first term captures investors' returns from investing in high quality startups thanks to their access to deal flow through their professional networks (e.g. a referral from a connection), or from their ability to screen startups and perform due diligence more effectively. Note that we do not distinguish between the two because empirically the size of one's professional network is likely to be highly correlated with one's ability (more on this below). Investors select between the following options: 1) they can leverage their professional connections to a hub region, select a startup in a hub and obtain $\gamma(n_i^H)$; 2) they can leverage their connections to a non-hub region and obtain $\gamma(n_i^{NH})$; or 3) they can randomly select a startup from a hub region and get ρ. ρ

[14] This is consistent with the findings in Chen et al. (2010).

captures the difference between the average return from startups located in hubs versus non-hubs. $\gamma(n_i^H)$ and $\gamma(n_i^{NH})$ are respectively the highest investment return that investors can make given their degree of connectedness in hub versus non-hub regions. The professional network strength measures – n_i^H and n_i^{NH} – are represented empirically by the number of professional connections investors have that are located in a specific region. We assume that $\gamma(n)$ increases in n, i.e. investors with more connections in the focal region have access to better investment opportunities in the same area (e.g. because of lower search costs, more connections to other local investors, higher degree of specialization in the area). For tractability, we assume that $n_i^H \sim U[0,1]$ and $n_i^{NH} \sim U[0,1]$. We allow for an arbitrary correlation between investors' locations and their connections in different regions: $n_i^H|(L_i = H) \sim U[\Delta_1, 1]$ and $n_i^{NH}|(L_i = NH) \sim U[\Delta_2, 1]$, where Δ_1 and Δ_2 are both between 0 and 1. The *max* operator captures the idea that investors choose the option that yields the highest profit.

The second term $\kappa_1 d_{id}$ captures the cost of mentoring and monitoring a startup conditional on investing. Of course, investors may decide to not apply any effort to any of these activities. d_{id} is the geographic distance between investors and startups. We assume that mentoring and monitoring costs increase in d_{id}, governed by a marginal effect of $\kappa_1 > 0$. For example, meeting with a distant startup could incur additional travel or information acquisition costs.

It is important to note that the size of an investor's pre-existing professional network is likely to not only be positively correlated with their access to deals in a region, but also with their ability to conduct due diligence, to attract interest from entrepreneurs, and to monitor and mentor startups (i.e. n_i is not orthogonal to overall ability and talent). Investors of higher ability will attract more inbound requests for investment from local and distant startups both because of their broader professional network, and because they are better at mentoring startups and extracting meaningful information about them during the due diligence process. In the model, we do not separate investors' ability from the reach of their professional network and access to deal flow, although in Table 6, we provide evidence that both mechanisms are at work.

Proposition 1 *(Geography of Direct Investments) Direct online investments are subject to a local premium (LP^1) and a hub premium (HP) if $\rho > \gamma\left(\frac{\Delta_2+2\eta(\kappa_1)(1-\Delta_2)}{1+\Delta_2}\right)$. The probability of* *investing outside of a hub weakly increases with the share of an investor's professional connections to individuals outside of hubs.*

Proof. The local premium (LP^1) results from two sources: a disproportionate share of an individual's professional network is typically local, and the mentoring and monitoring cost term $\kappa_1 d_{id}$. HP comes from the fact that investors with low n_i prefer a randomly drawn startup from a hub region – provided the quality differential ρ between startups in hub regions and non-hub regions is high enough – over investing in a startup they can source and evaluate through their professional network.[15] $\gamma'(\cdot) > 0$ implies that the probability of investing in non-hubs strictly increases with an investor's connections to individuals located in non-hubs if $n_i^{NH} \geq \gamma^{-1}(\rho)$; this probability is constant with respect to n_i^{NH} if $n_i^{NH} < \gamma^{-1}(\rho)$. ■

3.2 Introduction of Intermediaries on the Platform

We extend the model by allowing for intermediaries on the platform in the form of syndicate leads. Intermediaries source deals, and perform due diligence, mentoring and monitoring on behalf of investors (i.e. the crowd) in exchange for a share of future returns (the 'carry'). This approach borrows from the venture capital model, where general partners scout and curate deals (for the limited partners providing the capital) in exchange for a share of the returns that are realized when a startup is acquired or has an IPO. Intermediaries are indexed by s, and differ on the same dimensions as investors in terms of their location and the size of their professional network. Since the aim of our paper is to study the changes in capital flows induced by the introduction of intermediaries, we define the investor's return function under intermediated deals as:

$$\Pi_i^S = (1-\tau)\left[max\{\widetilde{\gamma}(n_i^H), \widetilde{\gamma}(n_i^{NH}), \rho\} - \kappa_2 d_{sd}\right]$$

Intermediaries charge carry τ for their services, and investors get $(1-\tau)$ of the overall investment return. Like investors, intermediaries take the profit-maximizing option among the three presented before: investing in hub regions through their professional network, investing in

[15]Formal proofs are in Appendix Section 8.1.

non-hubs through their professional network, or taking a random draw from a hub. We assume that the syndicated investment return $\widetilde{\gamma}(\cdot)$ – which increases with an intermediary's network $\widetilde{n_s}$ – also increases with the network of the investor n_i. This could be driven by better-connected investors having access to better intermediaries (e.g. if a deal syndicated by an intermediary is oversubscribed, only the best investors will be allocated a share in it), or by better-connected investors being able to screen intermediaries more effectively in the first place (e.g. because they not only rely on the intermediaries' public reputation, but also on information coming from their network). $\kappa_2 d_{sd}$ captures the intermediaries' cost of mentoring and monitoring startups.

Proposition 2 *(Geography of Investments Syndicated by Intermediaries) When investors select intermediaries they will exhibit a local premium (LP^2). Similar to direct investments, intermediaries exhibit a local premium in their selection of startups (LP^3). Under syndication by intermediaries, the overall hub premium (HP) decreases if $\widetilde{\gamma}(n) \geq \gamma(n)$ for all n.*

Proof. *(Informal proof)* We have LP^2 and LP^3 because professional connections are disproportionately local, and mentoring and monitoring costs increase with distance (both for screening intermediaries and startups). If $\widetilde{\gamma}(n) \geq \gamma(n)$ for all n, under syndication investors derive more value from their professional network. HP decreases because some investors who used to take a random draw from a hub region now switch to investments syndicated by intermediaries. Some of these investment funds may end up in non-hub regions depending on the geography of the intermediary's network. If instead $\kappa_2 < \tilde{\gamma}(\gamma^{-1}(\rho + \kappa_1)) - \rho$, investors with extensive networks will still prefer direct investment to using an intermediary.[16] ■

3.3 Model Extension: Heterogeneous Startup and Intermediary Quality

In Appendix Section 8.1.3, we extend the model by introducing heterogeneous startup and intermediary quality. In particular, we discuss both quality that is observable on the platform, as well as unobservable without a direct interaction with the startup (investors can obtain

[16] Formal proofs are in Appendix Section 8.1.

an estimate of true quality through face-to-face due diligence). This leads to two additional predictions. First, under direct investment, startups from non-hubs with weak quality signals will struggle the most to raise capital. Second, if intermediaries can extract a more precise signal of a startup's true quality through due diligence, then the difference in local premium between intermediaries (LP^3) and investors (LP^1) should be larger for startups that have ex-ante low, observable signals of quality. Hence, intermediaries will be more likely to surface startups that would otherwise look less promising on observables.

4 Data and Empirical Strategy

We use online investment reservation[17] data from AngelList from October 2012 to October 2016. AngelList is the leading equity crowdfunding platform operating under Title II of the US JOBS Act in terms of the number of accredited investors,[18] startups and deals that are performed online. According to the platform, the majority of the deals are seed stage investments (56%), although follow on investments at the Series A (17.7% of the activity) and Series B (15.5%) level are increasing. More than 300 intermediaries are on the platform, spanning multiple geographic locations and industry sectors: from IT, software and e-commerce to health care, fintech, hardware, logistics and analytics. As of 2016, over $440M have been invested online through the platform in over 1,000 startups. These online, early stage investments are often followed by larger, traditional venture capital rounds led by some of the top US VC firms.[19] Venture capital firms are also increasingly co-investing with the crowd to have the option to lead larger, follow-on rounds later. Although online investment by accredited investors was only introduced in 2013, by 2016 the platform had two 'unicorn' exits from its early investments:

[17]We focus on reservations since they reflect an investor's commitment to invest in a startup. If a deal is oversubscribed, some investors may be excluded from it and a reservation may not convert into a finalized investment. Our results do not change if we use use finalized investments instead of investment reservations, see Table A-4.

[18]For an investor to be considered accredited by the SEC, the individual must have a net worth of $1M (not including the primary residence), or an income of $200K in the last two years with the expectation of similar income going forward, see https://www.sec.gov/files/ib_accreditedinvestors.pdf

[19]E.g., Accel Partners, Andreessen Horowitz, Bessmer Venture Partners, First Round Capital, Greylock Partners, Khosla Ventures, Sequoia Capital, Union Square Ventures.

Dollar Shave Club, acquired by Unilever for $1B, and Cruise Automation, acquired by General Motors for $1B.[20]

The data we use in the paper consists of detailed information on investors, intermediaries, startups, and capital flows. One key feature of our dataset is that it allows us to observe investors' locations, professional networks, investment amounts, and whether they use an intermediary or decide to invest directly. We also have access to 2017 valuation data directly from the platform, and use this information to compare the unrealized returns of different types of deals over time.

Table 1 presents descriptive statistics for our sample. In Panel A, we see that during our observation period (2013–2016), most investment reservations come from investors located in California (65%). The share of investments syndicated by intermediaries has been steadily increasing on the platform since the introduction of the feature, and by the end of our sample the majority of online deals are syndicated. On average, 43% of reservations come from investors located in the same US state as the startup. This percentage is substantially higher within California (61%). As of 2017, 56% of investments syndicated by intermediaries have achieved above the median performance (*'Share High Performance'*) relative to their cohort, compared to 48% for direct investments.

The geographic distribution of startups (Panel B) and intermediaries (Panel D) is similar to the one presented in Panel A for reservations. Investors (Panel C) instead, are more geographically dispersed: 52% of investors are from California, 15% from New York, 4% from Massachusetts and the remaining 30% are from other regions. The total amount of capital reserved is approximately $870M, with a per startup average of $730K.

Observable startup quality is a measure which we build using information from the startups' profiles, and ranges from 0 to 10. The measure is the sum of a series of dummy variables that capture if the startup has: a video description, a description of products or services, a summary of what the startup does, information on customer traction, information on distribution, information on the technology behind the startup, any existing, pending, or granted patents,

[20] The internal, unrealized rate of return for all 2013 investments conducted on the platform has been higher than comparable metrics for top quartile venture capital funds of the same vintage, which is also consistent with these investments representing earlier stage, riskier deals relative to VC rounds.

information on hiring and job postings.

Comparing Panels C and D, we see that the average number of professional connections is 179 for investors, and 451 for intermediaries. On average, intermediaries charge 13% carry (share of future returns) on a per-deal basis on the investments they syndicate.

4.1 Empirical Strategy

Intermediaries (syndicate leads) were introduced as an experiment on the platform to see if offering a share of future returns (i.e. a carry) to angel investors could incentivize them to share their expertise and deal flow with the online crowd.[21] Their success progressively turned these intermediaries into the most used investment mode on the platform. Since the feature was not pre-announced, the introduction can be considered an exogenous event for the pre-existing online crowd of investors. Our main identification strategy takes advantage of this change in market design using an event-study approach, and introduces investor fixed effects to control for unobservable, time-invariant investor heterogeneity (e.g. investors' attitude towards risk, preferences for different sectors or investment types etc.). We estimate variations of:

$$Y_{rit} = \beta Post_t + \mu_i + \psi_t + \epsilon_{rit}, \tag{1}$$

where Y_{rit} is a dummy equal to one if investment reservation r is in state Y; $Post_t$ is a dummy for the introduction of syndication; μ_i and ψ_t are respectively investor and month fixed effects; and ϵ_{rit} is an idiosyncratic error term. The estimated $\widehat{\beta}$ captures the correlation between the introduction of syndication and the likelihood that capital will end up in the focal region.[22] In some of the regressions, we also control for the location and geographic clustering of the

[21] On AngelList, intermediaries create a syndicate profile on the platform, which contains information on how many deals they expect to syndicate each year and their typical investment size. As a syndicate lead, the intermediary commits to providing a written investment thesis for each investment and to disclosing potential conflicts of interest.

[22] Note that 'Post' is identified using variation over time in the overall share of investment across regions, and that direct investments are still available after intermediaries are introduced. Results are unchanged if we drop all direct investments in the post-period.

professional network of intermediaries:

$$Y_{rit} = \beta Post_t + \gamma_1 SameState\_s + \gamma_2 MajorityNetwork\_s + \mu_i + \psi_t + \epsilon_{rit}, \tag{2}$$

where suffix $\_s$ indicates an intermediary (syndicate lead). $SameState\_s$ is a dummy equal to one if the intermediary is co-located with the startup, and $MajorityNetwork\_s$ is a dummy variable for whether the majority of an intermediary's professional connections are in the same region as the startup.[23]

To test the robustness of our preferred specification based on the event-study approach – which exploits within-investor variation around the introduction of syndication – we also implement two matching estimators: a propensity score matching (PSM), and a matching, difference-in-differences estimator (MDID). The idea behind both methods is to compare capital flows for intermediated versus direct investments done by investors that constitute a credible control group for each other. Hence, the key identifying assumption for both approaches is that conditional on the matched observables, the groups of investors that select direct versus intermediated deals have otherwise a similar propensity to allocate capital to hub versus non-hub regions. To perform the matching, we take advantage of all available variables and match investors on: whether they are based in California or in a non-hub region, whether they have more than 500 connections on LinkedIn, have investment experience, had an investment that resulted in an IPO or acquisition, and whether they graduated from a Top 25 MBA program according to US News. This rich set of observables, which is based on key dimensions highlighted by the entrepreneurial finance literature, should proxy for both the sophistication of the investors as well as their preferences.

The propensity score matching (PSM) estimation is done in two steps: 1) we parametrically estimate the propensity of using intermediaries through a probit model; and 2) we estimate the changes in capital flows induced by intermediaries through a comparison between matched[24]

[23]Results do not change if we use the share of connections in the region of the startup instead of a dummy variable. To avoid issues related to reverse causality, connections formed within the last 6 months are dropped from the network measures.

[24]We rely on multiple matching methods: nearest neighbour matching, radius matching, kernel matching,

investors with similar estimated propensity scores. The matching, difference-in-differences estimator (MDID) follows instead the form:

$$\widehat{\beta}^{MDID} = \frac{1}{N} \sum_{i \in \{I^1 \cap S^*\}} \left\{ \Delta Y_{it} - \sum_{j \in \{I^0 \cap S^*\}} W_{ij} \Delta Y_{jt} \right\}$$

where I^1 is the set of "treated" investors who use syndication at least once; I^0 is the set of "control" investors who never use syndication in the year after its introduction; S^* is the region of common support revealed by the propensity score; W is the weight placed when comparing control unit j with treatment unit i (which depends on the matching method). Therefore, the MDID estimation compares temporal changes in the treated units against those of the matched, non-treated units. The key advantage of MDID over the PSM estimation is that MDID also allows for time-invariant, unobservable characteristics of the investors to affect selection into treatment. The identification assumption of the MDID approach is that there are no time-varying, unobservable effects that are correlated both with selection into using intermediaries and the decision of where to invest.

5 Results

We start by presenting descriptive results on the geography of capital flows under direct investment (Section 5.1). In the absence of intermediaries, i.e. when investment opportunities are directly posted online by startup founders, non-hub regions underperform entrepreneurial hubs relative to what one would expect based on their share of overall startup activity.

Next, using an event-study approach, we show that the introduction of intermediaries through online syndicates counterbalances this trend by expanding access to capital to non-hubs (Section 5.2). However, this is conditional on the presence of a particular type of intermediary located in a non-hub region or with substantial professional connections to a non-hub region. Results are unchanged when we use two different matching estimators to create a control group for investors

stratification matching, and inverse probability weights.

that use intermediaries, and when we include investors that were not present on the platform before the introduction of syndicated deals. We further unpack this result by splitting our sample by startups with ex-ante high versus low observable quality: our main "democratization effect" is largely driven by non-hub startups that had, before funding, weaker observable quality signals. This is consistent with intermediaries using their reputation to substitute for the more noisy reputation of non-hub startups.

But how do investors select an intermediary in the first place? To answer this question, in Section 5.3 we take advantage of heterogeneity in the observable quality of intermediaries. As the observable quality of intermediaries increases, we see that they are able to attract funding from everywhere. Moreover, the local bias of investors selecting intermediaries decreases monotonically with quality in the same way for intermediaries from California and from non-hub regions, suggesting that geographic frictions in the selection of intermediaries can be effectively overcome if the reputation of the intermediary is strong enough. Interestingly, whereas the local bias disappears, investors with more extensive professional connections to a non-hub region are still more likely to select an intermediary from their home region (a result that we do not find in California). Hence, information that travels through professional networks may still be important for identifying intermediaries from non-hubs that may appear as low quality to outsiders, but that are actually talented.

To understand the decision making process of online investors and how they evaluate intermediaries, in Section 5.3.2 we ran a large-scale field experiment. In the experiment, over 26,000 potential investors are presented with different information about three randomly selected intermediaries. We randomly show investors information about their past investment portfolio and performance (e.g. exits and acquisitions), endorsements by industry participants, the size of their professional network on the platform, and their direct overlap in professional networks (e.g. 1st degree, 2nd degree connection etc). Surprisingly, endorsement and past performance are quite ineffective in this market, and investors mostly care about the size of an intermediary's professional network, possibly because they see it as a difficult-to-game proxy for their access to deals and ability. Furthermore, investors mostly ignore the information presented in the emails if it is not about someone in their immediate professional network (1st or 2nd degree connection), suggesting that social networks may constitute a key, remaining friction to a broader democratization of access to capital because of the way they are used to weight information on the platform and screen information in a context where attention is limited.

Last, to assess if intermediaries actually reduce information asymmetry and identify startups that would otherwise be misclassified, in Section 5.4 we compare how intermediaries select startups relative to online investors. Geographic proximity plays a bigger role in the selection decisions of intermediaries relative to investors from the crowd, and this is driven by startups that have ex-ante low observable signals of quality. This suggests that intermediaries, through offline due diligence and private information, may be able to arbitrage investment opportunities across regions and identify mispriced deals. To test if this is the case, we introduce data on ex-post startup valuations: non-hub deals syndicated by intermediaries are more likely to have higher returns than other deals on the platform, and this is predominantly coming from startups that had ex-ante, low observable quality.

5.1 Capital Flows Under Direct Investment

In this section, we explore capital flows under the direct investment model, and use reservation data[25] for all deals that do not involve an intermediary. AngelList started as a website for listing angel investors and only later added startup and founder profiles. Direct investment was launched on the platform in 2013 after AngelList received a 'no-action letter' from the SEC,[26] which allowed it to accept online investments from accredited investors.[27]

As can be seen in Panel A of Table 2, under direct investment, startups from California receive 48% more investments than startups from other regions, although California's premium

[25] We focus on reservations since they reflect an investor's commitment to invest in a startup. If a deal is oversubscribed, some investors may be excluded from it and a reservation may not convert into a finalized investment. Our results do not change if we use use finalized investments instead of investment reservations, see Table A-4.

[26] https://www.sec.gov/divisions/marketreg/mr-noaction/2013/angellist-15a1.pdf

[27] For an investor to be considered accredited by the SEC, the individual must have a net worth of $1M (not including the primary residence), or an income of $200K in the last two years with the expectation of similar income going forward, see https://www.sec.gov/files/ib_accreditedinvestors.pdf

in terms of startup activity is only 30% (see Panel B, 65% − 35%). As hypothesized in the theoretical framework, this is consistent with the presence of a hub premium (HP): in the absence of intermediaries, online investors are more likely to invest in startups from regions with higher average startup quality, and may discount non-hub startups. Whereas investors from California exhibit a local premium (LP^1) and are more likely to invest in their own home region, because of the hub premium, this is not true for anyone else.[28]

Ironically, while the online channel expands the choice set of investors – and theoretically allows them to diversify their portfolios to include regions that otherwise face higher frictions in accessing capital – in the absence of complementary changes in market design and platform curation, investors default to areas that already have easier access to capital offline, reinforcing agglomeration.

One potential explanation for the inability of the platform to expand access to capital to new regions and for the observed hub premium (HP) is that online markets often rely on offline signals of quality to build trust, and that startups from California simply have access to better signals (e.g. association with notable, early investors or a top accelerator program). This would explain why investors seem more comfortable investing in them also over distance.

We test this hypothesis by splitting the sample by startups with low versus high observable signals of quality on the platform:[29] Interestingly, California has a comparable share of both types of startups (respectively 65% versus 63%), suggesting that observables may not drive the gap. At the same time, California's premium in direct investments on the platform is even larger (62%) for startups that have low, observable quality, and decreases (44%) when more information is available about the startup to begin with. This points to information asymmetry as a potential mechanism: whereas startups from hubs versus non-hubs may look similar on

[28]The hub premium is 52% for investors from California, and 44% for other investors. Consistent with Proposition 1, as the share of professional connections an investor has to a non-hub region increases, their hub premium decreases. Note that we aggregate all other states into the same category because they fit similar investment patterns. Results are the same if we further break down the analysis by state, or if we include New York or Massachusetts among hub regions.

[29]This is based on a measure of profile integrity which ranges from 0 to 10. Low and high signal are defined as above/below the median (5). The measure is the sum of a series of dummy variables that capture if the startup has: video description, description of products or services, a summary of what the startup does, information on customer traction, information on distribution, information on the technology behind the startup, any existing, pending, or granted patents, information on hiring and job postings.

observables, they may still differ on dimensions that are unobservable to us, but that investors may be able to extract from their online profiles, from asking for opinions from their professional network, or through offline due diligence. Alternatively, investors may rationally believe that hub startups have a higher chance of success because of the ecosystem they are embedded in (which, for example, may make it easier to raise follow on funding). This leads them to be more selective when investing outside of California, which would explain why the hub premium is lower for startups with high signals of quality.

Taken together, results from this section highlight how online equity crowdfunding platforms are unable to democratize access to capital to new regions in the absence of mechanisms designed to reduce information asymmetry.

5.2 Introduction of Intermediaries

Intermediaries (syndicate leads) were introduced as an experiment to see if incentivizing angel investors through a carry model would motivate them to share their expertise and conduct investments online. We rely on the fact that the change was not pre-announced – and can therefore be considered an exogenous event for pre-existing investors – to conduct an event-study around the months immediately preceding and following the change on the platform. To account for unobservable, investor heterogeneity (e.g. expertise, different preferences for risk, for types of investment etc.) that may drive the decision to use the new investment mode, we use investor fixed effects.[30] In particular, using our main specification (1), in Table 3 we rely on reservation data from the 10 months before and after the introduction of intermediaries[31] to see how the change in market design shaped capital allocation. In addition to investor fixed effects, all regressions include month fixed effects to control non-parametrically for the changing propensity to invest across regions over time. The analysis is performed at the reservation level (i.e. investor-startup pairs), and the dependent variable is equal to 1 if the startup is located in

[30] We later show robustness to not limiting the analysis to within-investor variation and build multiple control groups for the investors that use syndication.

[31] There are only 10 months of data between the introduction of intermediaries and the date when the SEC allowed platforms like AngelList to enable online investments.

California (Panel A), or in any other region (Panel B). Standard errors are clustered at the deal level.

Results are consistent with a 'democratization effect': the use of syndication is associated with a large 33% decrease in investment in Californian startups, and a corresponding increase in investment in non-hub regions (Column 1 of Table 3).[32] Decomposing this change in the hub premium by the location of the investors involved (see Table A-1, Column 1) shows that it results from both Californian investors diversifying outside of their home region, and from non-hub investors now favoring their home region over California (this second effect is 8% larger in magnitude than the first).

In Column 2 of Table 3, when we control for the intermediary being co-located with the startup (*'Intermediary in Same State'*) according to specification (2), coefficients are positive and comparable across the two panels. This is consistent with intermediaries, like investors, exhibiting a local premium (LP^3) in their startup selection decisions (we will explore this in more detail in Section 5.4). The role of geography is unchanged when we additionally control for the degree of geographical specialization of an intermediary's professional network: in Column 3, coefficients for *'Majority Network in Same State'*[33] are positive in both panels, and do not influence the estimates for *'Intermediary in Same State'*. This suggests that both the location of the intermediary and the composition of their professional network play a role in shaping capital flows. Intermediaries with more professional connections in a region may have lived there before, may regularly travel to the region, or may have a sizable share of contacts that have moved to that area. This may allow them to source deals and acquire information about local startups more effectively even if not co-located.[34]

Interestingly, whereas the *'Intermediary in Same State'* coefficients are of similar magnitude in Panels A and B, the professional network plays a substantially larger role for capital flows

[32]Results are the same if we use finalized investments instead of investment reservations, see Table A-4.

[33]The dummy variable captures whether more than half of intermediary's professional connections are in the startup's region. Results are the same if we use the share of an intermediary's connections (or the number of connections) in the region instead of the dummy variable.

[34]Since intermediaries of greater ability will also have broader networks, we focus on the relative degree of geographic specialization of their network rather than the absolute number of professional connections to an area.

to non-hub regions. Decomposing the main democratization effect by intermediary type (see Column 2 of Table A-1) sheds additional light on this difference. The largest changes in capital allocation are driven by: 1) intermediaries from California with professional networks specialized in non-hub regions; followed by 2) intermediaries located in non-hub regions with a disproportionate share of their network in California. In both cases, intermediaries seem to be able to take advantage of their professional network to overcome geographic distance: Californian intermediaries to identify startups in non-hub regions; non-hub intermediaries to credibly feature startups from their home region to Californian investors. When intermediaries do not have a professional network that allows them to bridge between hubs and non-hubs instead, capital flows are predominantly local, and non-hub regions are unable to tap into additional sources of funding.

In Table 4, we test the robustness of these results to including investors that join (or leave) the platform after the introduction of syndication. These regressions are similar to specifications (1) and (2), except that they do not include investor fixed effects. This allows us to introduce the location of the investors and the geographic specialization of their professional network in the regression. Similar to what we saw in Table 3, the use of syndication is associated with a 38.5% decrease in capital flows to California, and a corresponding increase in investments in other regions (Column 1). Positive coefficients for *'Investor in Same State'* are consistent with the presence of a local premium (LP^1) for investors (Column 2). The local premium is substantially larger in non-hub regions (Panel B), which is a reflection of non-hub investors investing in deals from California, but Californian investors not diversifying outside of their home region. Similar to what we saw for the networks of intermediaries, investors that have a disproportionate share of professional connections to a region are more likely to invest in it (Column 3), and this effect is independent of where they are located. When we add controls for the location and network of intermediaries in Columns 4 and 5, the weights on the investor variables are greatly diminished: the investors' local premiums are approximately halved (from 0.085 to 0.053 in California, and from 0.629 to 0.319 in other regions), and the network premiums are reduced by approximately one third in both regions. This is consistent with the idea that conditional on using an intermediary, the location of investors and the geographic specialization of their network play a lesser role in shaping capital flows. At the same time, they are still relevant, an issue that we will further unpack in the next section by looking at how investors select intermediaries.

Comparing Panels A and B of Table 4 also shows that within hub regions, geography and professional networks of both investors and intermediaries are less of a constraint for attracting capital from distant investors. To explore if this is the result of non-hub startups having weaker quality signals, in Table A-5 we repeat the analysis in Panel B for startups with high versus low signals of quality on the platform. This leads to three additional insights. First, consistent with information asymmetry constituting a key obstacle to online investment, the vast majority of the local and network premium for both investors and intermediaries is coming from startups with low, observable signals of quality. When information asymmetry is high, investors may be able to obtain additional information about a startup through offline due diligence and by asking for the opinion of individuals they trust, such as members of their professional network. Second, in line with the idea that intermediaries can identify and use their reputation to attract funding for high potential startups that would otherwise be ignored, the democratization effect is largest for non-hub startups with ex-ante low signal. As we have seen in the direct investment data (Section 5.1), in the absence of an intermediary filling the information gap, these startups are more likely to be misclassified relative to comparable ones located in California. Third, for high signal startups, once the intermediary's location and network are accounted for, the location and network of investors do not matter.

Overall, this section shows that the arrival of intermediaries is associated with investors from all regions allocating more capital to non-hub regions. At the same time, this 'democratization effect' is far from universal. In particular, it relies on the presence of intermediaries that are either based in California and have extensive connections to the target non-hub regions, or are based in non-hub regions but are plugged into the professional networks of California. The role of intermediaries is particularly critical when information asymmetry is likely to be high, and intermediaries can leverage their reputation to vouch for high potential, non-hub startups the crowd would otherwise classify as low potential ones. The data also hints at investor characteristics and professional networks mattering less when intermediaries are involved. However, they still influence how investors select intermediaries to begin with, an issue we delve into in the next section.

5.3 How do Investors Select Intermediaries?

Having provided robust evidence that the introduction of intermediaries is associated with an increase in capital flows to non-hub regions, we now try to isolate the underlying mechanism, and explore how investors select intermediaries. In investments syndicated by intermediaries, investors first select intermediaries, who then select which startups to invest in on behalf of the investors. In the next sub-sections, we rely on reservation data as well as a large-scale field experiment to shed light on this 'first stage'.

5.3.1 The Role of Intermediaries' Reputation

In the first part of the analysis, we focus on whether a selected intermediary is located in a hub or a non-hub region. We control for co-location between the investor and the intermediary (Table 5), as well as the geographic specialization of the investor's network (Table A-6). As predicted in the theoretical framework, in both tables, investors exhibit a local premium when selecting intermediaries (LP^2). This is not surprising, as co-location may proxy for the investors having additional information about the intermediary (e.g. from offline interactions and reputation), or simply feeling more comfortable investing through someone from their home region. What we are interested in testing is if this local premium is universal, or can be overcome through other sources of information. In particular, we want to test under which conditions the observable reputation of an intermediary can compensate, if at all, for geographic distance.

In Table 5, we explore how the local premium of investors changes with the observable reputation of intermediaries. The dependent variable is equal to one if the selected intermediary is based in California (Panel A), or is based in a non-hub region (Panel B). Intermediaries are evaluated by the platform through an extensive, internal review process before they can endorse any startup. The platform needs to ensure that these intermediaries are trustworthy, and will not take advantage of the association with the platform to endorse low quality deals. We use the scores (which range from 1 to 10) resulting from the platform's assessment to separate intermediaries by quality.[35] In Columns 1 to 4, we rely on this measure to split the data into quartiles of the observable intermediary's quality (1 being the lowest bin, 4 the highest): as the reputation of intermediaries increases – irrespective of whether they are located in a hub versus a non-hub region – they are able to attract investors from everywhere. Moving across bins, the local premium (*'Investor in Same State'*) almost completely disappears, from 21% in the first bin to 0.6% in the last one.

We observe a similar drop in the investors' network premium in Table A-6, although in this case, non-hub regions follow a slightly different pattern. Whereas the network premium (*'Majority Network in Same State'*) becomes insignificant for intermediaries from California starting from the 3rd bin, in non-hub regions it is still present for intermediaries of the highest, observable quality. Similar to non-hub startups, intermediaries from non-hub regions are more likely to be selected by investors that have professional connections to their region, and may be able to complement the information available on the platform with information coming from their professional network. We now turn to a large-scale field experiment to causally identify what type of information drives investors' interest in intermediaries on the platform.

5.3.2 A Large-Scale Field Experiment: How Do Online Investors Evaluate Intermediaries?

The previous analysis shows that the reputation of intermediaries – which is often established offline over multiple years of investments as an angel investor – can help overcome geographic frictions in online investment. But what aspects of an intermediary's reputation do investors really care about? And what type of information may make investors more likely to invest through intermediaries that are geographically distant or that have non-overlapping professional networks with them?

[35] Results do not change if we build more simple measures based on the profile pages of the intermediary. We prefer the platform's metrics because they are based on a more detailed evaluation of the intermediary.

To address these questions we designed a field experiment in which we randomly presented over 26,000 users on the platform with different information about three intermediaries through an email campaign. Messages were sent to individuals who had not invested before in two batches: 40% of potential investors received the email in October 2017, and the remaining 60% the following month. Each message featured a random triplet of intermediaries selected from a total pool of 21 syndicate leads.[36] In addition to a control group where only the name and picture of the intermediaries were shown together with a link to their full profile, we have five treatment groups (see Figure A-1): 1) in the *'connections'* treatment, we showed information about the size of an intermediary's professional network; 2) in the *'degree'* treatment, we highlighted if the intermediary belonged to the investor's 1st degree, 2nd-degree or 3rd-degree-or-above network (similar to how LinkedIn displays social proximity on its pages); 3) in the *'reference'* treatment, we displayed a short reference about the intermediary written by someone in the field (e.g. an entrepreneur that received investment from them); 4) in the *'portfolio'* treatment, we listed two of the intermediary's past investments, and included a short headline about their performance (e.g. in terms of follow-on funding, acquisition etc); 5) last, we included a treatment which covered all the information presented in the other four treatments.

We first check that our randomization is valid. In Appendix Table A-3, we report the differences (together with t-statistics) for multiple variables linked to our interventions between treated and control groups.[37] Reassuringly, there are no statistically significant differences between our groups at the 10% level (or lower). The evaluation of the experiment follows a simple intent-to-treat estimation. In particular, we are interested in assessing which pieces of information are more likely to lead users to click on an intermediary's profile and learn more about them. Results, which are based on regressions performed at the investor-intermediary level, are summarized in Figure 1. In this figure we focus on users that had a 1st or 2nd degree connection

[36]The list was formed by selecting intermediaries that had closed at least one deal in the last 6 months, and had completed deals and endorsements that could be used in the email. The 21 intermediaries were randomly allocated to 7 triplets, and then each message was randomly assigned to one of those triplets.

[37]In particular, we analyze whether there are differences in the following variables for each of the three intermediaries featured in an email: first-degree connection with the potential investor, number of professional connections, intermediary observable quality, the notoriety of the startups in the intermediaries' past investment set, the notoriety of intermediaries' references.

to the intermediary featured in the email. Appendix Figure A-2 covers the remaining part of the sample (i.e. users that had a 3rd degree connection or higher, including no connection at all). We divide the sample between investor-intermediary pairs that are socially connected (Figure 1) versus not (Figure A-2) because they respond differently to the interventions.

As can be seen in Figure 1, among users that were already connected or that shared a common connection with the intermediary (1st or 2nd degree connection), providing information about the size of the intermediary's professional network was the most effective way to get them to explore the full profile. Compared to a baseline click rate of 0.14% for the control group, this treatment almost doubles the likelihood that an investor will visit the intermediary's profile page. The other treatments, including the one where all information was presented at once, were either insignificant or did not increase click rates relative to the more simple control email (which only showed the intermediary's name and picture).

Whereas many online markets heavily rely on reviews and reputation scores similar to our 'reference' treatment (which was essentially an endorsement) and 'portfolio' treatment (which displayed past performance) to build trust, in this setting both approaches appear to be ineffective. One possible explanation for this behavior is that in markets for early stage capital the degree of information asymmetry is extremely high, and investors may not know how to interpret an endorsement or past performance of an intermediary in the absence of additional context (e.g. at what stage did the intermediary invest in the growth of the featured startup, who was leading the round, how many other deals did the intermediary invest in that did not have positive performance etc).

Furthermore, given the high degree of uncertainty and noise in this space, they may discount any piece of information that does not come from a source they already trust. This is consistent with an additional insight from the experiment: as can be seen in Appendix Figure A-2, which only includes investor-intermediary pairs that are not connected through their network (3rd degree or higher), in the absence of overlap in professional networks, the positive effect of displaying information about an intermediary's network is greatly diminished (and all other treatments are noisy and ineffective).

Taken together, the results from the randomized field experiment highlight that social networks may constitute a key obstacle for increasing capital flows to new regions and entrepreneurs. Online investors, both in the investment and in the experimental data, seem to heavily rely on their own professional networks to filter and weight the information available on the platform. Moreover, when screening intermediaries, they seem to prioritize information about their professional network, possibly because they see it as difficult to game, credible proxy for the intermediary's ability to access high quality deals and make good investment decisions, or simply because it is an easy to interpret, summary measure that individuals pay attention to in other contexts (e.g. LinkedIn connections, Twitter followers etc.).

By paying more attention to intermediaries that already overlap with them professionally and that have larger networks, investors may inadvertently reinforce path-dependency in capital flows and make it more difficult for intermediaries and entrepreneurs that are not already connected to hub regions to access capital.

5.4 How do Intermediaries Select Startups? Are They Arbitraging Opportunities Across Regions?

Having explored how investors select intermediaries, we now turn to the complementary question of how intermediaries select startups, and how this later relates to the performance of the deals they intermediate on the platform. One of the reasons investors may be interested in intermediaries with larger professional networks is that extensive professional connections may proxy for an intermediary's ability to access and secure deals that a regular investor may not be privy to. Those same connections may allow them to later introduce the startup to venture capitalists, reducing the follow on financing risk. A different, but not necessarily mutually exclusive explanation is that the size of intermediaries' networks may also be a proxy for their ability to effectively screen deals, perform due diligence and identify high potential startups that others may miss. In such a scenario, the network is a result of them being higher ability to begin with, and of having accumulated connections over a successful career in the industry.

But what would ability look like in this setting? Intermediaries earn a carry if the startups they endorse are successful, and face a reputational cost if they turn out to be of low quality. Relative to investors, they therefore have high powered incentives to reduce information asymmetry and invest time and resources in offline due diligence. By discovering 'gems' that are currently underpriced, they can also appropriate a sizable part of the returns from the startup's success. According to the theoretical framework, this should translate into intermediaries having a stronger local premium than regular investors, and such premium should increase with the degree of information asymmetry between the startup and its potential investors.

Indeed, in both panels of Table 6 we see that intermediaries exhibit a higher local premium in their selection of startups (*"Syndicated × Same State'*) than direct investors,[38] and that this premium is more visible for startups that have ex-ante low, observable signals of quality (Quartile 1). As the observable quality of startups increases and information asymmetry is less likely to be a concern, the local premium disappears. This is consistent with investors being able to evaluate startups with high signal on the platform irrespective of where they are located. Similar to what we have seen in the previous sections, asymmetric information is more of a concern in non-hub regions, where the local premium of intermediaries is present in all quartiles except for the highest one (Quartile 4).

But are these startups that appeared to be of lower quality on observables actual 'gems'? If intermediaries use their ability and network to access and identify high quality startups that would otherwise be misclassified, then their deals, especially in the presence of information asymmetry, should outperform regular investments on the platform. To test this, we use unique data on startup valuations, and calculate if an investment has above versus below the median (unrealized) returns within its cohort.[39] As can be seen in Column 1 of Table 7, deals syndicated

[38] *"Same State'* is a dummy equal to one if the startup and the agent making the startup selection decision (an investor in direct investment, an intermediary in the case of a syndicated investment) are co-located. Hence, the coefficient on this variable is meant to capture the "average" local premium of screening agents, and is driven both by investors in direct investments and intermediaries in syndicated investments. *"Syndicated × Same State'* instead, is the interaction between a dummy equal to 1 for deals syndicated by an intermediary and the *"Same State'* dummy. This second term measures the additional local premium we observe when intermediaries select startups relative to the baseline local premium we see when investors choose startups.

[39] This helps us account for the fact that older deals had more time to mature and experience both positive and negative updates in valuation. All valuations updated in 2017.

by intermediaries in startups based in non-hub regions (*'Startup in Other* $\times$ *Syndicated'*) are 36.9% more likely to have above the median returns. Consistent with intermediaries arbitraging an information asymmetry gap, the result is mostly driven by startups of extremely low observable quality (Column 2) – which are 68.5% more likely to have above the median returns when syndicated – and decays in the subsequent quartiles before becoming insignificant in the top one (Column 5).[40]

6 Conclusions

The impact of the internet and of digital platforms on economic outcomes has often been uneven (Forman et al., 2018). This has also been true within crowdfunding, with evidence supporting both an expansion of access to capital to projects and entrepreneurs that would otherwise not be funded, as well as the reinforcement of agglomeration and path-dependency in capital flows to hub regions.

We argue that one of the reasons why equity platforms have failed to substantially lower barriers to entry is because they have borrowed the market design of reward-based platforms like Kickstarter, and have ignored the idiosyncratic needs of a market where information asymmetry between buyers and sellers is much higher. In the absence of a mechanism for surfacing reliable information about startups through offline due diligence, and for building trust between investors and founders, the market unravels and only startups that already have very high signals of quality beforehand receive funding. This disproportionately favors startups that are located in top entrepreneurial hubs, limiting the potential of these platforms and the returns investors can enjoy on them.

In a context where traditional methods for establishing a reputation system are ineffective, we show that the introduction of intermediaries with domain expertise and an established offline reputation can help bootstrap new types of online transactions. Using novel data we show that

[40]Interestingly, within California, the contribution of intermediaries seems concentrated in the middle of the distribution (Columns 2 and 3), and turns negative in the top quartile, possibly because deals featuring high signal startups from California are already recognized and priced correctly on the platform without an intermediary, and adding one only increases costs.

intermediaries allow capital to be allocated outside of traditional startup hubs, expanding the set of founders that can rely on this new source of early-stage capital to fund their startup.[41]

At the same time, this ‘democratization effect’ is far from universal, and relies on intermediaries that have professional networks that span hub and non-hub regions. Consistent with information asymmetry constituting a first key obstacle in equity crowdfunding markets, the changes induced by intermediaries are most visible among startups in non-hub regions that have ex-ante low, observable signals of quality. These same startups are also responsible for above median investment returns, suggesting that intermediaries are indeed able to arbitrage opportunities between regions and make early stage capital markets more efficient.

The second key obstacle we identify relates to the social networks of the individuals and intermediaries involved. In the absence of reliable proxies for the quality of intermediaries, investors default to the size of their professional networks (possibly a ‘popularity’ measure), and rely on their social graph to weight the information they receive from the platform. This may again introduce path-dependency, as regions and entrepreneurs that are already connected to California and to influential, notable investors are more likely to receive attention on the platform.

Given the increasing importance of this new source of early-stage capital for high growth startups, we believe our findings are useful for understanding the matching process taking place on these platforms, and for assessing how intermediaries can contribute through their access to deal flow and expertise to the functioning and scaling of these online markets. Our geography results also have the potential to inform policy, as they highlight the conditions under which online capital is more versus less likely to flow to regions that do not have a strong angel and venture capital presence. On the one hand, we find that the platform is able to expand access to capital to high quality entrepreneurs from non-hub regions that otherwise might not have been funded due to their weak, observable signals of quality. On the other hand, we also show that this effect rests on the availability of intermediaries that have hybrid professional networks to

[41] This result is consistent with the finding that under the right market design, peer-to-peer platforms increase welfare in disadvantaged regions (see Lam and Liu (2017) in the case of Uber, Farronato and Fradkin (2018) in the case of Airbnb, and Lendle et al. (2016); Hui (2018) in the case of eBay).

begin with.

In his overview of innovation policy in a networked world, Sorenson (2018) highlights a key challenge we face in a context where geography is becoming less relevant because of the internet, but social networks are often used for filtering information and deciding who gets access to opportunities. Our paper shows that for equity crowdfunding platforms to further expand their ability to identify undervalued 'gems' and remove frictions in early stage capital markets, new market design mechanisms are required. In particular, platforms need to experiment with how to balance the value of information that travels through professional networks with the cost of excluding high potential entrepreneurs (and intermediaries) that do not already have access to such networks.

References

Agrawal, A., C. Catalini, and A. Goldfarb (2014). Some simple economics of crowdfunding. *Innovation Policy and the Economy 14*(1), 63–97.

Agrawal, A., C. Catalini, and A. Goldfarb (2015). Crowdfunding: Geography, social networks, and the timing of investment decisions. *Journal of Economics & Management Strategy 24*(2), 253–274.

Agrawal, A., C. Catalini, and A. Goldfarb (2016). Are syndicates the killer app of equity crowdfunding? *California management review 58*(2), 111–124.

Agrawal, A., C. Catalini, A. Goldfarb, and H. Luo (2018). Slack time and innovation. *Organization Science*.

Ahlers, G. K., D. Cumming, C. Günther, and D. Schweizer (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice 39*(4), 955–980.

Bapna, S. (2017). Complementarity of signals in early-stage equity investment decisions: Evidence from a randomized field experiment. *Management Science*.

Belleflamme, P., T. Lambert, and A. Schwienbacher (2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing 29*(5), 585–609.

Bernstein, S., A. Korteweg, and K. Laws (2017). Attracting early-stage investors: Evidence from a randomized field experiment. *The Journal of Finance 72*(2), 509–538.

Blum, B. S. and A. Goldfarb (2006). Does the internet defy the law of gravity? *Journal of International Economics 70*(2), 384–405.

Brander, J. A., R. Amit, and W. Antweiler (2002). Venture-capital syndication: Improved venture selection vs. the value-added hypothesis. *Journal of Economics & Management Strategy 11*(3), 423–452.

Burtch, G., A. Ghose, and S. Wattal (2013). An empirical examination of the antecedents and consequences of contribution patterns in crowd-funded markets. *Information Systems Research 24*(3), 499–519.

Burtch, G., A. Ghose, and S. Wattal (2015). The hidden cost of accommodating crowdfunder privacy preferences: A randomized field experiment. *Management Science 61*(5), 949–962.

Chen, H., P. Gompers, A. Kovner, and J. Lerner (2010). Buy local? the geography of venture capital. *Journal of Urban Economics 67*(1), 90–102.

Cooper, I. and E. Kaplanis (1994). Home bias in equity portfolios, inflation hedging, and international capital market equilibrium. *Review of Financial Studies 7*(1), 45–60.

Coval, J. D. and T. J. Moskowitz (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance 54*(6), 2045–2073.

Cumming, D. and N. Dai (2010). Local bias in venture capital investments. *Journal of Empirical Finance 17*(3), 362–380.

Cumming, D. J. and S. A. Johan (2013). *Venture capital and private equity contracting: An international perspective*. Academic Press.

Dranove, D. and G. Z. Jin (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature 48*(4), 935–63.

Einav, L., C. Farronato, and J. Levin (2016). Peer-to-peer markets. *Annual Review of Economics 8*, 615–635.

Elfenbein, D. W., R. Fisman, and B. McManus (2015). Market structure, reputation, and the value of quality certification. *American Economic Journal: Microeconomics 7*(4), 83–108.

Farronato, C. and A. Fradkin (2018). The welfare effects of peer entry in the accommodation market: The case of airbnb. Technical report, National Bureau of Economic Research.

Forman, C., A. Goldfarb, and S. Greenstein (2018). How geography shapes—and is shaped by—the internet. *The New Oxford Handbook of Economic Geography*, 269.

French, K. R. and J. M. Poterba (1991). Investor diversification and international equity markets. *American Economic Review 81*(2), 222–226.

Giannetti, M. and L. Laeven (2012). The flight home effect: Evidence from the syndicated loan market during financial crises. *Journal of Financial Economics 104*(1), 23–43.

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

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics 24*(3), 461–483.

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

Holtz, D., D. L. MacLean, and S. Aral (2017). Social structure and trust in massive digital markets. *Working Paper*.

Hortaçsu, A., F. Martínez-Jerez, and J. Douglas (2009). The geography of trade in online transactions: Evidence from ebay and mercadolibre. *American Economic Journal: Microeconomics 1*(1), 53–74.

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

Huberman, G. (2001). Familiarity breeds investment. *Review of Financial Studies 14*(3), 659–680.

Hui, X. (2018). Facilitating inclusive global trade: Evidence from a field experiment. *Working Paper*.

Hui, X., M. Saeedi, Z. Shen, and N. Sundaresan (2016). Reputation and regulations: Evidence from ebay. *Management Science*.

Hui, X., M. Saeedi, G. Spagnolo, and S. Tadelis (2017). Certification, reputation and entry: An empirical analysis. *Working Paper*.

Ivashina, V. and D. Scharfstein (2010). Loan syndication and credit cycles. *American Economic Review 100*(2), 57–61.

Kaplan, S. N. and P. Strömberg (2003). Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *The Review of Economic Studies 70*(2), 281–315.

Kim, K. and S. Viswanathan (2014). The experts in the crowd: The role of reputable investors in a crowdfunding market. In *The 41st Research Conference on Communication, Information and Internet Policy*.

Kim, K. and S. Viswanathan (2018). The'experts' in the crowd: The role of experienced investors in a crowdfunding market. *MIS Quarterly*, forthcoming.

Kuppuswamy, V. and B. L. Bayus (2013). Crowdfunding creative ideas: the dynamics of projects backers in kickstarter. *Working Paper*.

Lam, C. T. and M. Liu (2017). Demand and consumer surplus in the on-demand economy: the case of ride sharing. *Working Paper*.

Lendle, A., M. Olarreaga, S. Schropp, and P.-L. Vézina (2016). There goes gravity: ebay and the death of distance. *The Economic Journal 126*(591), 406–441.

Lerner, J. (1994). The syndication of venture capital investments. *Financial Management*, 16–27.

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

Lin, M., N. R. Prabhala, and S. Viswanathan (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science 59*(1), 17–35.

Lin, M. and S. Viswanathan (2015). Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science 62*(5), 1393–1414.

Liu, D., D. J. Brass, Y. Lu, and D. Chen (2015). Friendships in online peer-to-peer lending: Pipes, prisms, and relational herding. *MIS Quarterly 39*(3), 729–742.

Mollick, E. R. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing 29*(1), 1–16.

Mollick, E. R. and R. Nanda (2016). Wisdom or madness? comparing crowds with expert evaluation in funding the arts. *Management Science 62*, 1533–1553.

Nanda, R. and T. Khanna (2010). Diasporas and domestic entrepreneurs: Evidence from the indian software industry. *Journal of Economics & Management Strategy 19*(4), 991–1012.

Peek, J. and E. S. Rosengren (1997). The international transmission of financial shocks: The case of japan. *The American Economic Review*, 495–505.

Roberts, J. W. (2011). Can warranties substitute for reputations? *American Economic Journal: Microeconomics 3*(3), 69–85.

Sorenson, O. (2018). Innovation policy in a networked world. *Innovation Policy and the Economy 18*(1), 53–77.

Sorenson, O., V. Assenova, G.-C. Li, J. Boada, and L. Fleming (2016). Expand innovation finance via crowdfunding. *Science 354*(6319), 1526–1528.

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

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics 8*, 321–340.

Tinbergen, J. (1962). An analysis of world trade flows. *Shaping the World Economy*, 1–117.

Van Nieuwerburgh, S. and L. Veldkamp (2009). Information immobility and the home bias puzzle. *The Journal of Finance 64*(3), 1187–1215.

Vulkan, N., T. Åstebro, and M. F. Sierra (2016). Equity crowdfunding: A new phenomena. *Journal of Business Venturing Insights 5*, 37–49.

Zhang, J. and P. Liu (2012). Rational herding in microloan markets. *Management Science 58*(5), 892–912.

7 Tables and Figures

Table 1: Summary Statistics

Panel A. Investment Reservations Characteristics					
	All	CA	NY	MA	Other States
Share of Investment Reservations	100%	65%	16%	3%	16%
Mean Amount per Reservation	7,757	8,050	8,305	10,286	5,519
Share of Syndicated Reservations	95%	94%	95%	99%	97%
Share of Same State Reservations	43%	61%	14%	9%	6%
Share High Perf. (Direct Investment)	48%	46%	76%	N/A	25%
Share High Perf. (Syndicated by Intermediary)	56%	52%	65%	63%	61%
Panel B. Startup Characteristics					
	All	CA	NY	MA	Other States
Share of Startups ($\geq$1 res.)	100%	64%	16%	6%	15%
Share of Syndicated Startups	79%	76%	94%	81%	82%
Total Dollars Received	870,878,876	586,457,027	151,191,112	34,899,734	98,331,003
Mean Dollars Received	733,063	774,712	804,208	505,793	565,121
Median Observable Startup Quality	5	5	5	4	5
Panel C. Investor Characteristics					
	All	CA	NY	MA	Other States
Share of Investors ($\geq$1 res.)	100%	52%	15%	4%	30%
Mean Number of Professional Connections	179	117	24	6	33
Panel D. Intermediary Characteristics					
	All	CA	NY	MA	Other States
Share of Intermediaries ($\geq$1 res.)	100%	68%	13%	9%	10%
Mean Number of Professional Connections	451	298	63	15	74
Median Observable Intermediary Quality	7	7	7	8	7
Intermediary Share of Profits ('carry')	13	14	13	12	13
Minimum Investment Amount	6,709	6,128	5,825	9,650	5,231
Number of Annual Deals	6	6	7	3	6

Notes: The data covers all investment reservation data from AngelList from October 2012 to October 2016. A reservation has high performance if its investment return is above the median for its vintage.

Table 2: Under Direct Investment, Non-Hub Regions Underperform

Panel A. California's Premium in the Share of Direct Investments

	All Startups	Low Signal	High Signal
All Investments	+48%	+60%	+44%
Investments from California	+52%	+62%	+46%
Investments from Other	+44%	+52%	+42%

Panel B. Geography of Startups

	All Startups		Low Signal		High Signal	
	CA	Other	CA	Other	CA	Other
Share of Startups in	65%	35%	65%	35%	63%	37%

Notes: The data covers all direct investment reservation data from AngelList from October 2012 to October 2016. California's premium is defined as the difference between the measure of interest for California and the same measure calculated for other regions. Quality signals are based on the profiles of the startups across ten dimensions (see Section 4 for additional details), and therefore range from 0 to 10 (low and high signal are defined as above/below the median, which is 5).

Table 3: Intermediaries Expand Capital to Non-Hub Regions (Event Study Around the Introduction of Syndication with Investor and Month Fixed Effects)

Panel A. Dependent Variable: Startup from California (1/0)

	(1)	(2)	(3)
Post	-0.330**	-0.405***	-0.349**
	(0.131)	(0.133)	(0.141)
Intermediary in Same State		0.220***	0.235***
		(0.064)	(0.063)
Majority Network in Same State			0.361**
			(0.162)
User FE	✓	✓	✓
Month FE	✓	✓	✓
Adj R-squared	0.067	0.098	0.129
N	3959	3959	3959

Panel B. Dependent Variable: Startup from Other Region (1/0)

	(1)	(2)	(3)
Post	0.330**	0.409***	0.404***
	(0.131)	(0.133)	(0.132)
Intermediary in Same State		0.204***	0.200***
		(0.065)	(0.064)
Majority Network in Same State			0.716***
			(0.070)
User FE	✓	✓	✓
Month FE	✓	✓	✓
Adj R-squared	0.067	0.093	0.135
N	3959	3959	3959

Notes: For the estimation, we use data from the ten months before and after the introduction of syndication (August 2013) using specifications (1) and (2). *'Post'* is a dummy for the introduction of syndication. *'Intermediary in Same State'* is equal to one if the intermediary and the startup are co-located. *'Majority Network in Same State'* is equal to one if the majority of the intermediary's professional network is located in the startup's home region. Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Robustness to Including Investors that Join the Platform After the Market Design Change

Panel A. Dependent Variable: Startup from California (1/0)

	(1)	(2)	(3)	(4)	(5)
Post	-0.385***	-0.362***	-0.357***	-0.484***	-0.421***
	(0.103)	(0.103)	(0.103)	(0.110)	(0.121)
Investor in Same State		0.090***	0.085***	0.056***	0.053***
		(0.016)	(0.015)	(0.011)	(0.011)
Investor Majority Network In Same State			0.046***	0.029*	0.030*
			(0.017)	(0.015)	(0.015)
Intermediary in Same State				0.248***	0.263***
				(0.083)	(0.084)
Intermediary Majority Network In Same State					0.357**
					(0.162)
Month FE	✓	✓	✓	✓	✓
Adj R-squared	0.030	0.039	0.040	0.091	0.129
N	19098	19098	19098	19098	19098

Panel B. Dependent Variable: Startup from Other Region (1/0)

	(1)	(2)	(3)	(4)	(5)
Post	0.385***	0.373***	0.372***	0.343***	0.340***
	(0.103)	(0.101)	(0.101)	(0.100)	(0.100)
Investor in Same State		0.649***	0.629***	0.330***	0.319***
		(0.049)	(0.048)	(0.075)	(0.074)
Investor Majority Network in Same State			0.379***	0.298***	0.261***
			(0.071)	(0.088)	(0.090)
Intermediary in Same State				0.714***	0.734***
				(0.061)	(0.059)
Intermediary Majority Network in Same State					0.721***
					(0.072)
Month FE	✓	✓	✓	✓	✓
Adj R-squared	0.030	0.056	0.058	0.134	0.191
N	19098	19098	19098	19098	19098

Notes: For the estimation, we use data from the ten months before and after the introduction of syndication (August 2013). *'Post'* is a dummy for the introduction of syndication. *'Investor in Same State'* is equal to one if the investor and the startup are co-located. *'Investor Majority Network in Same State'* is equal to one if the majority of the investor's professional network is located in the startup's home region. *'Intermediary in Same State'* is equal to one if the intermediary and the startup are co-located. *'Majority Network in Same State'* is equal to one if the majority of the intermediary's professional network is located in the startup's home region. Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Investors Selecting Intermediaries by Intermediaries' Observable Quality

Panel A. Dependent Variable: Intermediary from California (1/0)

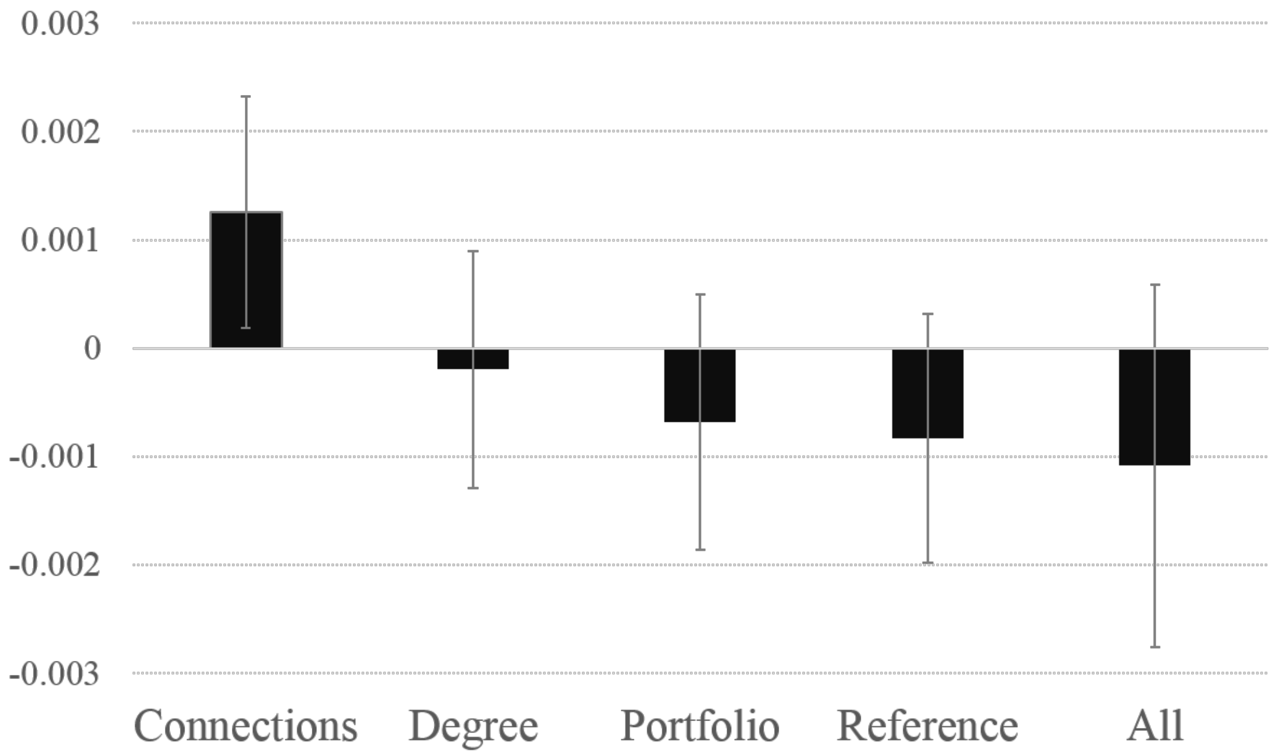
	(1) Intermediary Quality Quartile 1	(2) Intermediary Quality Quartile 2	(3) Intermediary Quality Quartile 3	(4) Intermediary Quality Quartile 4
Investor in Same State	0.208***	0.085***	0.082***	0.008***
	(0.008)	(0.018)	(0.020)	(0.002)
Adj R-squared	0.044	0.014	0.007	0.000
N	45280	5225	3867	58157

Panel B. Dependent Variable: Intermediary from Other Regions (1/0)

	(1) Intermediary Quality Quartile 1	(2) Intermediary Quality Quartile 2	(3) Intermediary Quality Quartile 3	(4) Intermediary Quality Quartile 4
Investor in Same State	0.206***	0.137***	0.076***	0.006***
	(0.013)	(0.025)	(0.028)	(0.002)
Adj R-squared	0.045	0.031	0.014	0.001
N	45280	5225	3867	58157

Notes: For the estimation, we use all data from the period after the introduction of syndication (August 2013–October 2016). *'Investor in Same State'* is equal to one if the investor is located in the same region as the intermediary. Quartiles are based on the intermediaries' observable quality, which ranges from 0 to 10 (see Section 5.3 for more details). Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1: Randomized Field Experiment: Investors Pay Attention to the Intermediary's Connections



Notes: The outcome variable is a dummy equal to one if the investor visited the intermediary's profile page. This figure focuses on investors that had either a 1st or 2nd degree connection with the intermediary (results for the remaining part of the sample are in Figure A-2). The treatment effect of 'Connection' is statistically different from 0 and from the other treatments at the 5% level. In the *'connections'* treatment, we showed information about the size of an intermediary's professional network. In the *'degree'* treatment, we highlighted if the intermediary belonged to the investor's 1st degree, 2nd-degree or 3rd-degree-or-above network (similar to how LinkedIn displays social proximity on its pages). In the *'reference'* treatment, we displayed a short reference about the intermediary written by someone in the field (e.g. an entrepreneur that received investment from them). In the *'portfolio'* treatment, we listed two of the intermediary's past investments, and included a short headline about their performance (e.g. in terms of follow-on funding, acquisition etc). Last, we included a treatment which covered all the information presented in the other four treatments ('All'). For an overview of the different treatments, see Figure A-1.

Table 6: Investors Selecting Startups versus Intermediaries Selecting Startups: Heterogeneity by Observable Startup Quality

Panel A. Dependent Variable: Startup from California (1/0)

	Startup Quality Quartile 1 (1)	Startup Quality Quartile 2 (2)	Startup Quality Quartile 3 (3)	Startup Quality Quartile 4 (4)
Same State	0.062**	0.027	-0.003	0.043
	(0.029)	(0.035)	(0.033)	(0.035)
Syndicated × Same State	0.262***	0.014	-0.013	0.030
	(0.088)	(0.073)	(0.060)	(0.093)
Adj R-squared	0.391	0.425	0.509	0.594
N	30109	31260	33826	17334

Panel B. Dependent Variable: Startup from Other Regions (1/0)

	Startup Quality Quartile 1 (1)	Startup Quality Quartile 2 (2)	Startup Quality Quartile 3 (3)	Startup Quality Quartile 4 (4)
Same State	0.111	0.120	0.101	0.143
	(0.087)	(0.091)	(0.064)	(0.166)
Syndicated × Same State	0.497***	0.616***	0.735***	0.244
	(0.148)	(0.139)	(0.100)	(0.232)
Adj R-squared	0.587	0.718	0.821	0.549
N	30109	31260	33826	17334

Notes: We use all data from the period after the introduction of intermediaries (August 2013–October 2016) for the estimation. 'Same State' is a dummy equal to one if the startup and the agent making the startup selection decision (an investor in direct investment, an intermediary in the case of a syndicated investment) are co-located. Standard errors are clustered at the reservation level. Quartiles are based on the observable quality of the startups involved, which ranges from 0 to 10 (see Section 4 for more details). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Investment Performance by Investment Mode and Startup Observable Quality

	(1) All Startups	(2) Startup Quality Quartile 1	(3) Startup Quality Quartile 2	(4) Startup Quality Quartile 3	(5) Startup Quality Quartile 4
Startup in California × Syndicated	0.065	-0.023	0.173	0.380**	-0.328**
	(0.082)	(0.157)	(0.121)	(0.164)	(0.137)
Startup in Other × Syndicated	0.369***	0.685***	0.393*	0.409*	0.242
	(0.138)	(0.115)	(0.205)	(0.237)	(0.222)
Startup in California	0.210	0.653***	0.144	0.114	0.227
	(0.146)	(0.145)	(0.209)	(0.248)	(0.233)
Adj R-squared	0.009	0.005	0.012	0.023	0.099
N	71276	17394	26995	12632	14255

Notes: We use data from the period before and after the introduction of syndication for which we have startup valuation data on (October 2016–October 2016). Quartiles are based on the observable quality of the startups involved, which ranges from 0 to 10 (see Section 4 for more details). Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

8 Online Appendix (Not for Publication)

Table A-1: Decomposing the 'Democratization Effect'

Dependent Variable: Startup in California

	(1)		(2)
Californian Investors (Baseline)	-0.340*** (0.106)	Int. in California w/Majority Network in California (Baseline)	-0.214* (0.108)
Other Investors	-0.081** (0.037)	Int. in Other w/Majority Network in California	-0.256*** (0.067)
		Int. in California w/Majority Network in Other	-0.727*** (0.081)
		Int. in Other w/Majority Network in Other	0.073 (0.212)
Month FE	✓		✓
Adj R-squared	0.041		0.083
N	19098		19098

Notes: For the estimation, we use data from ten months before and after the introduction of syndication (August 2013). The table decomposes the 'democratization effect' by the location of investors and the type of intermediary. In column (1), we see that the decrease in capital flows to California is 34% for Californian investors, and 8% larger for investors from non-hub regions. In column (2), we see that the change in capital flows is only significant at the 10% level for Californian intermediaries that have a majority of their professional connections to individuals located in California. It is instead large and significant for both intermediaries that are based in California, but have extensive professional connections to other regions (73% more intense than the baseline), as well as for non-hub intermediaries with professional connections to individuals in California (26% more intense than the baseline). Last, there is no change in capital flows for intermediaries that reside in non-hub regions, but that do not have a majority of professional connections outside of their home area. Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-2: Propensity Score and Matching with Difference-in-Differences Estimation

Panel A. Matching Quality									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	CA	NY	Other	Follower	Invested	Entrep.	Success	Exp.	MBA25
Block 1 ×	0.064	-0.033	-0.031	-0.038	0.039	0.071*	0.006	0.054	0.054*
Syndicated	(0.044)	(0.028)	(0.041)	(0.036)	(0.031)	(0.041)	(0.026)	(0.050)	(0.028)
Block 2 ×	-0.010	-0.019	0.029	-0.004	-0.026*	-0.011	-0.001	-0.021	-0.020
Syndicated	(0.022)	(0.014)	(0.021)	(0.018)	(0.016)	(0.021)	(0.013)	(0.025)	(0.013)
Block 3 ×	-0.029	0.210**	-0.181	0.031	-0.014	0.013	0.061	-0.069	-0.172
Syndicated	(0.154)	(0.097)	(0.143)	(0.125)	(0.107)	(0.142)	(0.091)	(0.175)	(0.129)
Panel B. Dependent Variable: Startup from California (1/0)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	NN	Radius	Kernel	Strata	IPW	IPW	MDID	MDID
ATT	-0.149***	-0.143**	-0.174***	-0.178***	-0.110***	-0.118**	-0.058*	-0.142**	-0.075*
	(0.049)	(0.056)	(0.045)	(0.041)	(0.020)	(0.049)	(0.035)	(0.070)	(0.045)
Same State							✓		✓
Month FEs							✓		✓
Panel C. Dependent Variable: Startup from Other Region (1/0)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	NN	Radius	Kernel	Strata	IPW	IPW	MDID	MDID
Syndicated	0.152***	0.102***	0.166***	0.150***	0.120***	0.100***	0.059**	0.094**	0.056**
	(0.025)	(0.031)	(0.025)	(0.023)	(0.014)	(0.028)	(0.027)	(0.038)	(0.027)
Same State							✓		✓
Month FEs							✓		✓

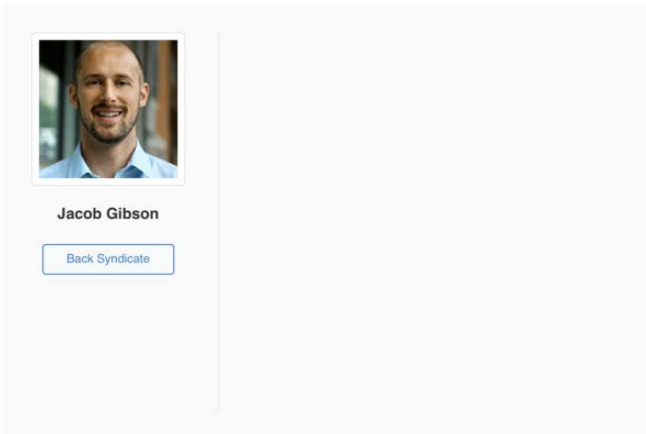
Notes: For the estimation, we use data from the months before and after the introduction of syndication (August 2013). To provide further robustness to our finding that intermediaries expand access to capital to new regions, in the table we use two matching estimators to build a control group: propensity score matching (PSM), and a matching, difference-in-difference estimator (MDID). Investors are matched using multiple variables: their location, whether they have more than 500 LinkedIn connections, have invested previously, had an investment that led to an exit, and whether they graduated from a Top 25 MBA program. In Panel A, we regress each matching variable on strata fixed effects and their interaction with a dummy for syndication. If the matching is valid, we should see non-significant coefficients for all these interactions. Our balance between the treatment (syndication) and control group (direct investment) is valid at the 1% level, and valid at the 10% level for the vast majority of the variables. We repeat the matching with five matching estimators: nearest neighbor, radius, kernel, stratification, and inverse probability weights. PSM estimation results are in columns (2)–(7) of Panel B. In Column (1) we report the OLS estimates. Column (7) controls for co-location as well as month fixed effects. We find between a 5.8% and 17.8% reduction in capital flows to startups from California, and a similar increase in other regions. We also perform a matching, difference-in-differences estimation (MDID). In Columns (8) and (9) of Panel B, there is a 7.5% to 14.2% decrease in the share of investments to startups in California, and a comparable increase in other regions. Standard errors clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-3: Validity of the Randomization: Baseline Balance

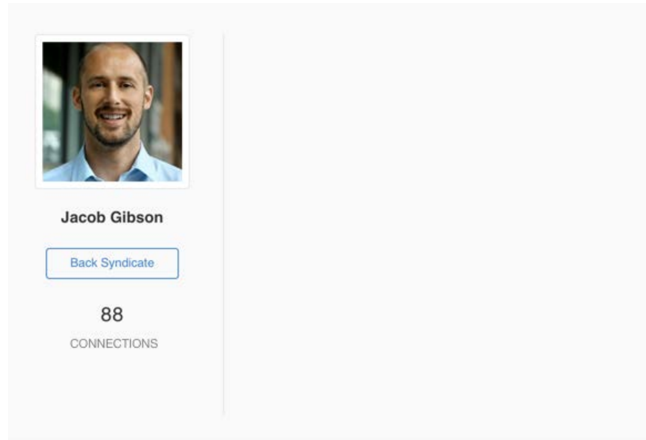
	(1) Connections	(2) Degree	(3) Portfolio	(4) Reference
1st Degree Connection–Intermediary 1	-0.003	0.001	-0.001	0.002
	(-0.970)	(0.320)	(-0.380)	(0.590)
1st Degree Connection–Intermediary 2	-0.004)	(-0.001	0.000	0.000
	(-0.750)	(-0.120)	(0.040)	(-0.090)
1st Degree Connection–Intermediary 3	-0.005	-0.001	0.000	-0.002
	(-1.250)	(-0.260)	(0.090)	(-0.550)
Num. Connections–Intermediary 1	1.491	-1.183	2.338	0.046
	(0.460)	(-0.390)	(0.770)	(0.020)
Num. Connections–Intermediary 2	1.883	-1.060	1.545	-1.288
	(0.910)	(-0.550)	(0.810)	(-0.670)
Num. Connections–Intermediary 3	3.037	-5.020	0.969	0.900
	(0.517)	(-1.140)	(0.220)	(0.210)
Quality of Intermediary 1	-0.022	0.012	-0.020	0.030
	(-0.670)	(0.380)	(-0.660)	(0.990)
Quality of Intermediary 2	0.078	-0.122	-0.073	-0.006
	(0.670)	(-1.110)	(-0.670)	(-0.060)
Quality of Intermediary 3	-0.065	0.006	0.053	0.028
	(-0.730)	(0.070)	(1.060)	(0.240)
Startup Notoriety–Intermediary 1	-0.008	0.000	0.027	-0.018
	(-0.330)	(-0.010)	(1.260)	(-0.840)
Startup Notoriety–Intermediary 2	-0.003	-0.002	0.015	-0.020
	(-0.130)	(-0.120)	(0.019)	(-1.030)
Startup Notoriety–Intermediary 3	-0.019	-0.003	0.004	0.003
	(-0.890)	(-0.160)	(0.230)	(0.140)
Reference Notoriety–Intermediary 1	0.019	-0.017	0.001	0.010
	(0.730)	(-0.730)	(0.050)	(0.420)
Reference Notoriety–Intermediary 2	0.011	-0.008	0.011	0.005
	(1.120)	(-0.870)	(1.230)	(0.560)
Reference Notoriety–Intermediary 3	0.034	-0.006	-0.016	-0.024
	(1.590)	(-0.300)	(-0.800)	(-1.190)

Notes: The table reports differences for the following variables between each of the four treatment groups and the control group: first-degree connection between the intermediary and the potential investor; number of connections; intermediary's observable quality; the notoriety of the startups featured in the intermediary's portfolio; the notoriety of the individuals writing the endorsement for the intermediary. We do not find statistically significant differences for any of the intermediary triplets at the 10% level (t-statistics are in parentheses). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

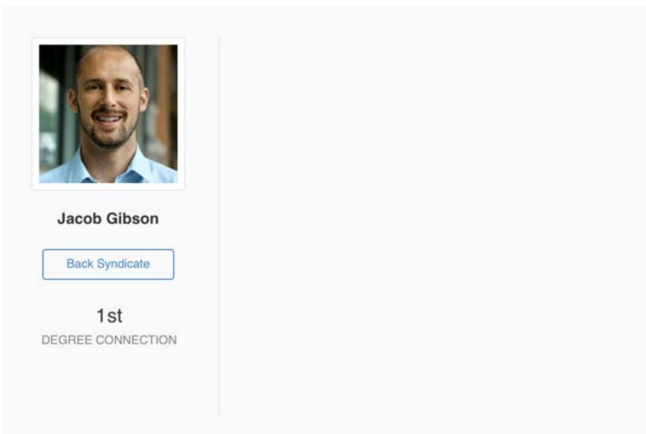
Figure A-1: Experimental Treatments



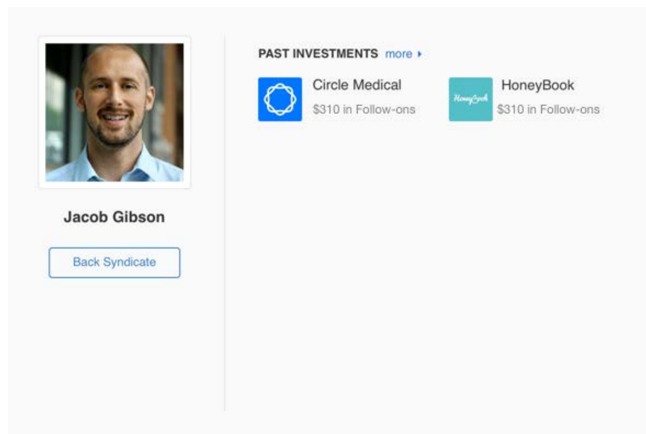
(a) Control Group



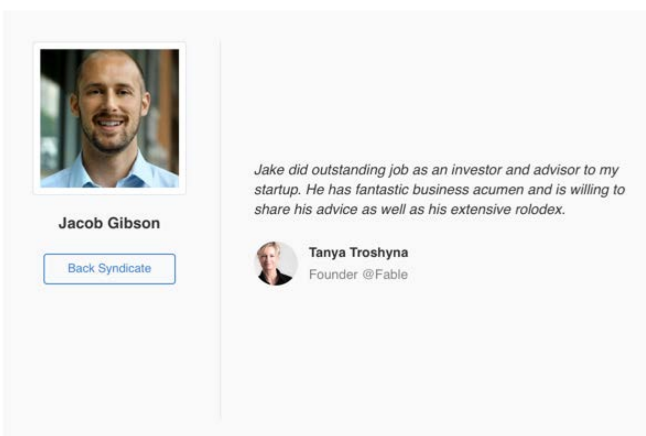
(b) 'Connections' Treatment



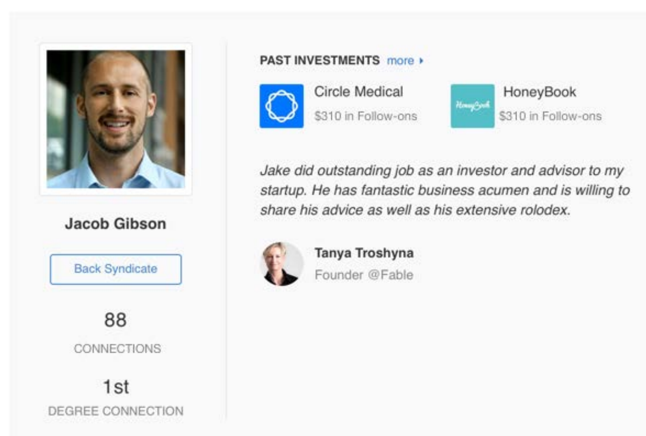
(c) 'Degree' Treatment



(d) 'Portfolio' Treatment



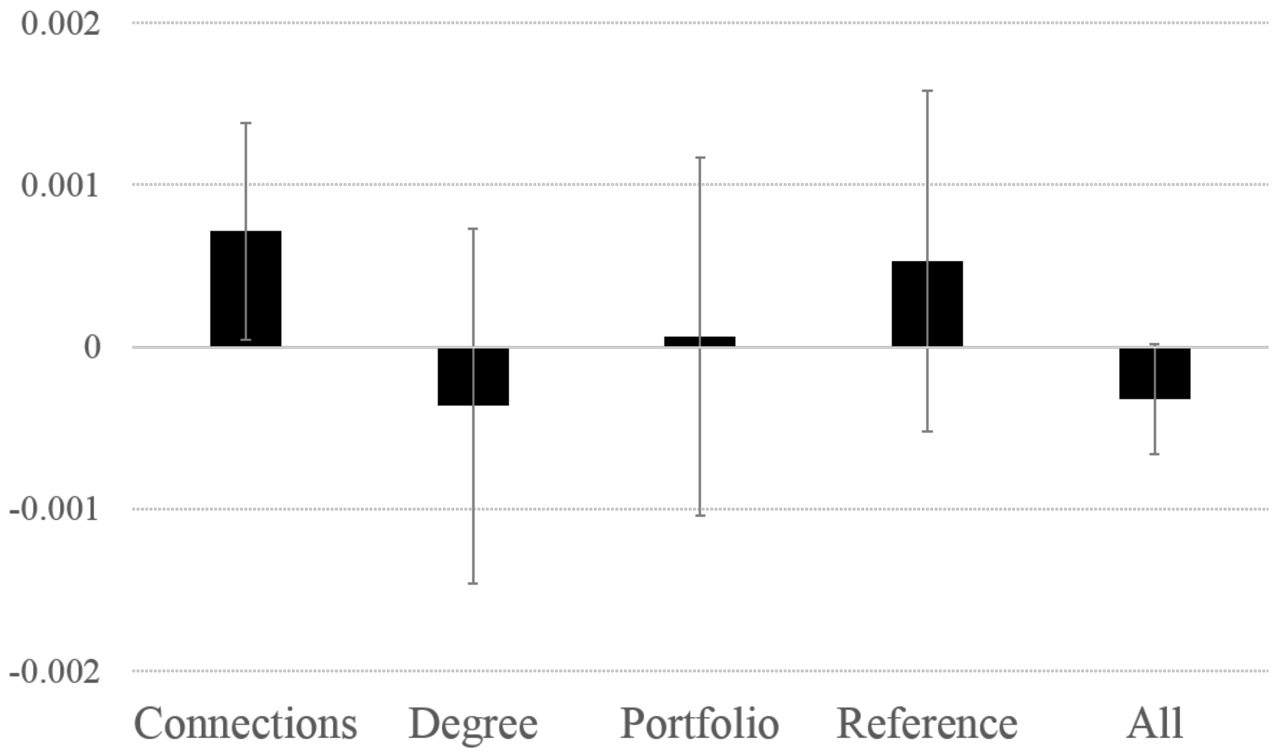
(e) 'Reference' Treatment



(f) All Treatments

The figure displays the different emails investors received under the control group (A-1a), which only included the intermediary's name, picture and profile link; the connections treatment (A-1b), which added information about the number of professional connections the intermediary has on the platform; the degree treatment (A-1c), which showed if the focal investor had a 1st, 2nd or 3rd degree or higher connection with the intermediary; the portfolio treatment A-1d, which featured past investments of the intermediary with their outcomes; the reference treatment (A-1e), which included an endorsement of the intermediary by an industry participant; and the 'all treatments' condition (A-1f).

Figure A-2: Investors Ignore Information About Intermediaries that Are Not Close to Them on the Professional Network



Notes: The outcome variable is a dummy equal to one if the investor visited the intermediary's profile page. This figure focuses on investors that were not close to the intermediaries on the professional network graph (3rd degree connection or higher, results for the other part of the sample are in Figure 1). Similar to our previous findings, the only significantly positive treatment is 'Connections', which increases the chance of clicking on the featured intermediary's profile to learn more about them by 0.07%. This effect is statistically smaller than the 0.12% for the case where there is a 1st or 2nd degree connection between investors and intermediaries (p-value=0.07). Additionally, this effect is not statistically different from the other treatments in the figure at the 10% level. In the *'connections'* treatment, we showed information about the size of an intermediary's professional network. In the *'degree'* treatment, we highlighted if the intermediary belonged to the investor's 1st degree, 2nd-degree or 3rd-degree-or-above network (similar to how LinkedIn displays social proximity on its pages). In the *'reference'* treatment, we displayed a short reference about the intermediary written by someone in the field (e.g. an entrepreneur that received investment from them). In the *'portfolio'* treatment, we listed two of the intermediary's past investments, and included a short headline about their performance (e.g. in terms of follow-on funding, acquisition etc). Last, we included a treatment which covered all the information presented in the other four treatments ('All'). For an overview of the different treatments, see Figure A-1.

Table A-4: Robustness: Reservations Vs. Finalized Investments

Panel A. Dependent Variable: Startup from California (1/0)

	(1)	(2)	(3)
Post	-0.345***	-0.492***	-0.441***
	(0.132)	(0.146)	(0.153)
Post× Investment	0.048	0.117	0.129
	(0.033)	(0.080)	(0.080)
Intermediary in Same State		0.220***	0.236***
		(0.076)	(0.077)
Intermediary in Same State×Investment		0.001	-0.001
		(0.060)	(0.058)
Majority Network in Same State			0.332*
			(0.172)
Majority Network in Same State×Investment			0.067
			(0.084)
Investment		-0.075	-0.145
		(0.071)	(0.110)
User FE	✓	✓	✓
Month FE	✓	✓	✓
Adj R-squared	0.068	0.100	0.132
N	3959	3959	3959

Panel B. Dependent Variable: Startup from Other Region (1/0)

	(1)	(2)	(3)
Post	0.254**	0.394***	0.384***
	(0.117)	(0.130)	(0.129)
Post × Investment	0.004	-0.071	-0.072
	(0.025)	(0.071)	(0.071)
Intermediary in Same State		0.199***	0.192***
		(0.067)	(0.066)
Intermediary in Same State×Investment		-0.026	-0.027
		(0.047)	(0.047)
Majority Network in Same State			0.872***
			(0.083)
Majority Network in Same State×Investment			-0.010
			(0.034)
Investment		0.091	0.092
		(0.075)	(0.075)
User FE	✓	✓	✓
Month FE	✓	✓	✓
Adj R-squared	0.061	0.092	0.138
N	3959	3959	3959

Notes: For the estimation, we use data from ten months before and after the introduction of syndication (August 2013) using a variation of specification (2). 'Investment' is a dummy variable for whether a reservation converts into a finalized investment. In the paper, we focus on reservation data since they reflect an investor's commitment to invest in a startup. If a deal is oversubscribed, some investors may be excluded from it and a reservation may not convert into an investment. Here, we show that using investment data does not change our results. In particular, we interact all explanatory variables from our main table with 'Investment'. The interaction terms are not statistically significant. Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-5: Robustness: Splitting Results for Other Regions by the Observable Quality of Startups

Dependent Variable: Startup in Other Region (1/0)

Panel A. Above-Median Startup Quality

	(1)	(2)	(3)	(4)	(5)
Post	0.288***	0.278***	0.278***	0.231**	0.230**
	(0.108)	(0.104)	(0.104)	(0.096)	(0.096)
Investor in Same State		0.578***	0.561***	0.224***	0.202***
		(0.121)	(0.117)	(0.083)	(0.075)
Investor Majority Network In Same State			0.211**	0.134	0.055
			(0.088)	(0.087)	(0.079)
Intermediary in Same State				0.680***	0.699***
				(0.103)	(0.099)
Intermediary Majority Network In Same State					0.709***
					(0.125)
Month FE	✓	✓	✓		
Adj R-squared	0.148	0.172	0.172	0.260	0.358
N	8882	8882	8882	8882	8882

Panel B. Below-Median Startup Quality

	(1)	(2)	(3)	(4)	(5)
Post	0.532***	0.522***	0.520***	0.480***	0.476***
	(0.159)	(0.159)	(0.158)	(0.179)	(0.180)
Investor in Same State		0.606***	0.592***	0.269***	0.265**
		(0.075)	(0.074)	(0.103)	(0.102)
Investor Majority Network In Same State			0.451***	0.440***	0.448***
			(0.080)	(0.105)	(0.105)
Intermediary in Same State				0.753***	0.755***
				(0.081)	(0.081)
Intermediary Majority Network In Same State					0.619***
					(0.078)
Month FE	✓	✓	✓		
Adj R-squared	0.113	0.139	0.141	0.184	0.191
N	10216	10216	10216	10216	10216

Notes: We use data from ten months before and after the introduction of intermediaries (August 2013) for the estimation. Quality signals are based on the profile integrity of the startups in ten dimensions, and range from 0 to 10 (see Section 4 for more details). The increase in capital flows is larger for below-median startups relative to above-median ones. This makes sense because below-median startups have weaker signals of quality, and therefore could not get as much investment before the introduction of intermediaries because of information asymmetry. Comparing columns 4 and 5 across panels, we see that for above-median startups, once we control for the location and network of the intermediary, investors location and network do not matter in terms of where they invest, while they still matter for below-median startups. Last, almost all the estimated local premium and network premium for both investors and intermediaries are larger for below-median startup. Standard errors are clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-6: Robustness: Investors Selecting Intermediaries by Intermediary Observable Quality

Panel A. Dependent Variable: Intermediary from California (1/0)

	Intermediary Quality Quartile 1	Intermediary Quality Quartile 2	Intermediary Quality Quartile 3	Intermediary Quality Quartile 4
	(1)	(2)	(3)	(4)
Investor in Same State	0.151***	0.066***	0.095***	0.008***
	(0.007)	(0.015)	(0.023)	(0.003)
Investor Majority Network	0.136***	0.051**	-0.031	-0.000
in Same State	(0.010)	(0.023)	(0.037)	(0.004)
Adj R-squared	0.056	0.017	0.008	0.001
N	45280	5225	3867	58157

Panel B. Dependent Variable: Intermediary from Other Regions (1/0)

	Intermediary Quality Quartile 1	Intermediary Quality Quartile 2	Intermediary Quality Quartile 3	Intermediary Quality Quartile 4
	(1)	(2)	(3)	(4)
Investor in Same State	0.191***	0.082***	0.076***	0.005***
	(0.012)	(0.018)	(0.027)	(0.001)
Investor Majority Network	0.387***	0.799***	0.332**	0.274**
in Same State	(0.030)	(0.039)	(0.154)	(0.118)
Adj R-squared	0.058	0.138	0.024	0.012
N	45280	5225	3867	58157

Notes: In this table we repeat the analyses from Table 5 by adding in measures for the investors' network. Similar to the change in investors' local premium in selecting intermediaries (LP^2), we find that the change in investors' reliance on their network decreases for intermediaries in California more rapidly than in other regions (where it is significant in all quartiles). Quartiles are based on intermediaries' observable quality, which ranges from 0 to 10. Standard errors clustered at the reservation level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

8.1 Formal Proofs and Model Extension

8.1.1 Formal Proofs of Hub Premium and Local Premium in Proposition 1

We first show the existence of a hub premium (HP) for investors located both in hub and non-hub regions under some parameter restrictions. For investors in hub regions, the probability of investing in their home region is:

$$
\begin{aligned}
Pr(H|L_i = H) &= Pr\left(\gamma(n^H) > \gamma(n^{NH}) - \kappa_1 | L_i = H\right) \\
&\quad + Pr\left(\gamma(n^{NH}) - \kappa_1 < \rho \bigcap \gamma(n^H) < \gamma(n^{NH}) - \kappa_1 | L_i = H\right) \\
&> Pr\left(\gamma(n^H) > \gamma(n^{NH}) | L_i = H\right) \\
&\quad + Pr\left(\gamma(n^{NH}) < \rho \bigcap \gamma(n^H) < \gamma(n^{NH}) | L_i = H\right) \\
&= Pr(n^H > n^{NH} | L_i = H) \\
&\quad + Pr(\gamma(n^{NH}) < \rho | L_i = H, n^H < n^{NH}) Pr(n^H < n^{NH} | L_i = H) \\
&= \frac{1}{2}(1 + \Delta_1) + \gamma^{-1}(\rho)\left(1 - \frac{1}{2}(1 + \Delta_1)\right) \\
&= \frac{1}{2} + \frac{1}{2}\Delta_1 + \frac{1}{2}\gamma^{-1}(\rho) - \frac{1}{2}\Delta_1\gamma^{-1}(\rho)
\end{aligned}
$$

Note that the second last relationship comes from the following two derivations:

$$
\begin{aligned}
Pr(n^H > n^{NH} | L_i = H) &= Pr(n^H > n^{NH} | L_i = H, n^{NH} > \Delta_1) Pr(n^{NH} > \Delta_1 | L_i = H) \\
&\quad + Pr(n^H > n^{NH} | L_i = H, n^{NH} \leq \Delta_1) Pr(n^{NH} \leq \Delta_1 | L_i = H) \\
&= \frac{1}{2}(1 - \Delta_1) + \Delta_1 \\
&= \frac{1}{2}(1 + \Delta_1).
\end{aligned}
$$

and

$$
\begin{aligned}
Pr(\gamma(n^{NH}) < \rho | L_i = H, n^H < n^{NH}) &= Pr(n^{NH} < \gamma^{-1}(\rho) | L_i = H, n^H < n^{NH}) \\
&\quad * Pr(n^H < n^{NH} | L_i = H) \\
&= \gamma^{-1}(\rho)\left(1 - \frac{1}{2}(1 + \Delta_1)\right).
\end{aligned}
$$

On the other hand, the probability of investing in a non-hub region for an investor in a hub region is given by:

$$
\begin{aligned}
Pr(NH|L_i = H) &= 1 - Pr(H|L_i = H) \\
&= \frac{1}{2} - \frac{1}{2}\Delta_1 - \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\Delta_1\gamma^{-1}(\rho)
\end{aligned}
$$

To show the existence of HP for investors in hub regions, we need to show that $Pr(H|L_i = H) - Pr(NH|L_i = H) > 0$. It is enough to show that $Pr(H|L_i = H) > \frac{1}{2}$:

$$\begin{aligned} Pr(H|L_i = H) &> \frac{1}{2} + \frac{1}{2}\Delta_1 + \frac{1}{2}\gamma^{-1}(\rho) - \frac{1}{2}\Delta_1\gamma^{-1}(\rho) \\ &= \frac{1}{2} + \frac{1}{2}\Delta_1(1-\gamma^{-1}(\rho)) + \frac{1}{2}\gamma^{-1}(\rho) \\ &> 0 \end{aligned}$$

Having established the existence of a hub premium for investors in hub regions, we show the existence of a hub premium for investors in non-hub regions. The probability of non-hub investors investing in a hub region is:

$$\begin{aligned} Pr(H|L_i = NH) &= Pr\left(\gamma(n^H) - \kappa_1 > \gamma(n^{NH})|L_i = NH\right) \\ &\quad + Pr\left(\gamma(n^{NH}) < \rho \bigcap \gamma(n^H) - \kappa_1 < \gamma(n^{NH})|L_i = NH\right) \\ &= Pr(n^H > n^{NH} + \eta(\kappa_1)|L_i = NH) \\ &\quad + Pr(\gamma(n^{NH}) < \rho|L_i = NH, n^H < n^{NH})Pr(n^H < n^{NH}|L_i = NH) \\ &= 1 - \left[(\frac{1}{2} + \eta(\kappa_1))(1-\Delta_2) + \Delta_2\right] + \gamma^{-1}(\rho)\frac{1}{2}(1+\Delta_2) \\ &= \frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_1) + \Delta_2\eta(\kappa_1) + \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\gamma^{-1}(\rho)\Delta_2 \end{aligned}$$

Note that the last relationship comes from the following two derivations:

$$\begin{aligned} Pr(n^{NH} + \eta(\kappa_1) > n^H|L_i = NH) &= Pr(n^{NH} + \eta(\kappa_1) > n^H|L_i = NH, n^H > \Delta_2) \\ &\quad * Pr(n^H > \Delta_2|L_i = NH) \\ &\quad + Pr(n^{NH} + \eta(\kappa_1)| > n^H|L_i = NH, n^H \leq \Delta_2) \\ &\quad * Pr(n^H \leq \Delta_2|L_i = NH) \\ &= (\frac{1}{2} + \eta(\kappa_1))(1-\Delta_2) + \Delta_2 \end{aligned}$$

and

$$\begin{aligned} Pr(\gamma(n^{NH}) < \rho|L_i = NH, n^H < n^{NH}) &= Pr(n^{NH} < \gamma^{-1}(\rho)|L_i = NH, n^H < n^{NH}) \\ &\quad * Pr(n^H < n^{NH}|L_i = NH) \\ &= \gamma^{-1}(\rho)\frac{1}{2}(1+\Delta_2). \end{aligned}$$

On the other hand, the probability of investing in a non-hub region for a non-hub investor is:

$$\begin{aligned} Pr(NH|L_i = NH) &= 1 - Pr(H|L_i = NH) \\ &= \frac{1}{2} + \frac{1}{2}\Delta_2 + \eta(\kappa_1) - \Delta_2\eta(\kappa_1) - \frac{1}{2}\gamma^{-1}(\rho) - \frac{1}{2}\gamma^{-1}(\rho)\Delta_2 \end{aligned}$$

To show the existence of HP for investors in non-hub regions, we need to show that $Pr(H|L_i =$

$NH) - Pr(NH|L_i = NH) > 0$. It is enough to show that $Pr(H|L_i = NH) > \frac{1}{2}$ if the average quality difference between hub and non-hub deals is not too small:

$$\begin{aligned} Pr(H|L_i = NH) &= \frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_1) + \Delta_2\eta(\kappa_1) + \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\gamma^{-1}(\rho)\Delta_2 \\ &> \frac{1}{2} \\ &if \quad \rho > \gamma\left(\frac{\Delta_2 + 2\eta(\kappa_1)(1-\Delta_2)}{1+\Delta_2}\right) \end{aligned}$$

Having established a hub premium for investors from both regions, we now show the existence of a local premium (LP^1) for deals in both regions under some parameter restrictions. For hub deals, the difference between them attracting local versus distant investors is:

$$\begin{aligned} Pr(H|L_i = H) - Pr(H|L_i = NH) &= \frac{1}{2} + \frac{1}{2}\Delta_1 + \frac{1}{2}\gamma^{-1}(\rho) - \frac{1}{2}\Delta_1\gamma^{-1}(\rho) \\ &\quad - \left[\frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_1) + \Delta_2\eta(\kappa_1) + \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\gamma^{-1}(\rho)\Delta_2\right] \\ &= \frac{1}{2}\Delta_1(1-\gamma^{-1}(\rho)) + \frac{1}{2}\Delta_2(1-\gamma^{-1}(\rho)) + (1-\Delta_2)\eta(\kappa_1) \\ &> 0 \end{aligned}$$

For deals in non-hubs, the difference between them attracting local versus distant investors is:

$$\begin{aligned} Pr(NH|L_i = NH) - Pr(NH|L_i = H) &= \frac{1}{2} + \frac{1}{2}\Delta_2 + \eta(\kappa_1) - \Delta_2\eta(\kappa_1) - \frac{1}{2}\gamma^{-1}(\rho) - \frac{1}{2}\gamma^{-1}(\rho)\Delta_2 \\ &\quad - \left[\frac{1}{2} - \frac{1}{2}\Delta_1 - \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\Delta_1\gamma^{-1}(\rho)\right] \\ &= \frac{1}{2}\Delta_2(1-\gamma^{-1}(\rho)) + (1-\Delta_2)\eta(\kappa_1) + \frac{1}{2}\Delta_1(1+\gamma^{-1}(\rho)) \\ &> 0 \end{aligned}$$

8.1.2 Formal Proofs of Hub Premium and Local Premium in Proposition 2

First, we show the existence of a local premium (LP^2) when investors select intermediaries in both regions under some parameter restrictions. For intermediaries in a hub region, the difference between them being chosen by local versus distant investors is:

$$\begin{aligned} Pr(H|L_i = H) - Pr(H|L_i = NH) &= \frac{1}{2} + \frac{1}{2}\Delta_1 + \frac{1}{2}\tilde{\gamma}^{-1}(\rho) - \frac{1}{2}\Delta_1\tilde{\gamma}^{-1}(\rho) \\ &\quad - \left[\frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_2) + \Delta_2\eta(\kappa_2) + \frac{1}{2}\tilde{\gamma}^{-1}(\rho) + \frac{1}{2}\tilde{\gamma}^{-1}(\rho)\Delta_2\right] \\ &= \frac{1}{2}\Delta_1(1-\tilde{\gamma}^{-1}(\rho)) + \frac{1}{2}\Delta_2(1-\tilde{\gamma}^{-1}(\rho)) + (1-\Delta_2)\eta(\kappa_2) \\ &> 0 \end{aligned}$$

For intermediaries in non-hub regions, the difference between the probability of them being selected by a local versus distant investors is:

$$
\begin{aligned}
Pr(NH|L_i = NH) - Pr(NH|L_i = H) &= \frac{1}{2} + \frac{1}{2}\Delta_2 + (\kappa_2) - \Delta_2\eta(\kappa_2) - \frac{1}{2}\tilde{\gamma}^{-1}(\rho) - \frac{1}{2}\tilde{\gamma}^{-1}(\rho)\Delta_2 \\
&\quad - \left[\frac{1}{2} - \frac{1}{2}\Delta_1 - \frac{1}{2}\tilde{\gamma}^{-1}(\rho) + \frac{1}{2}\Delta_1\tilde{\gamma}^{-1}(\rho)\right] \\
&= \frac{1}{2}\Delta_2(1 - \tilde{\gamma}^{-1}(\rho)) + (1 - \Delta_2)\eta(\kappa_2) + \frac{1}{2}\Delta_1(1 + \tilde{\gamma}^{-1}(\rho)) \\
&> 0
\end{aligned}
$$

Having showed the existence of a local premium (LP^2) when investors choose intermediaries in both regions, we now show that intermediaries exhibit a local premium (LP^3) in their selection of startups. The proof follows the same structure used in deriving LP^1. Last, we show that the hub premium is smaller for investments syndicated by an intermediary relative to direct investments. For investors in hubs, the probability of investing in a hub region under direct investment is:

$$
\begin{aligned}
Pr^{Dir}(H|L_i = H) &= Pr\left(\gamma(n^H) > \gamma(n^{NH}) - \kappa_1 | L_i = H\right) \\
&\quad + Pr\left((n^{NH}) - \kappa_1 < \rho \bigcap \gamma(n^H) < \gamma(n^{NH}) - \kappa_1 | L_i = H\right) \\
&= (\frac{1}{2} + \eta(\kappa_1))(1 - \Delta_1) + \Delta_1 + \gamma^{-1}(\rho + \kappa_1)\frac{1}{2}(1 + \Delta_1)
\end{aligned}
$$

Note that the last relationship comes from the following two derivations:

$$
\begin{aligned}
Pr^{Dir}\left(\gamma(n^H) > \gamma(n^{NH}) - \kappa_1 | L_i = H\right) &= Pr(n^H > n^{NH} - \eta(\kappa_1)|L_i = H) \\
&= Pr(n^H > n^{NH} - \eta(\kappa_1)|L_i = H, n^{NH} > \Delta_1) \\
&\quad * Pr(n^{NH} > \Delta_1|L_i = H) \\
&\quad + Pr(n^H > n^{NH} - \eta(\kappa_1)|L_i = H, n^{NH} \leq \Delta_1) \\
&\quad * Pr(n^{NH} \leq \Delta_1|L_i = H) \\
&= (\frac{1}{2} + \eta(\kappa_1))(1 - \Delta_1) + \Delta_1
\end{aligned}
$$

and

$$
\begin{aligned}
Pr^{Dir}\left((n^{NH}) - \kappa_1 < \rho \bigcap \gamma(n^H) < \gamma(n^{NH}) - \kappa_1 | L_i = H\right) &= Pr(n^{NH} < \gamma^{-1}(\rho + \kappa_1)|L_i = H, n^H < n^{NH}) \\
&\quad * Pr(n^H < n^{NH}|L_i = H) \\
&= \gamma^{-1}(\rho + \kappa_1)\frac{1}{2}(1 + \Delta_1).
\end{aligned}
$$

Now, for investors in a hub region, the probability of investing in a hub region with an intermediary is:

$$\begin{aligned} Pr^{Synd}(H|L_i = H) &= Pr\left(\tilde{\gamma}(n^H) > \tilde{\gamma}(n^{NH}) - \kappa_2 | L_i = H\right) \\ &\quad + Pr\left(\tilde{\gamma}(n^{NH}) - \kappa_2 < \rho \bigcap \tilde{\gamma}(n^H) < \tilde{\gamma}(n^{NH}) - \kappa_2 | L_i = H\right) \\ &= (\frac{1}{2} + \eta(\kappa_2))(1 - \Delta_1) + \Delta_1 + \tilde{\gamma}^{-1}(\rho + \kappa_2)\frac{1}{2}(1 + \Delta_1) \end{aligned}$$

Note that the last relationship comes from:

$$\begin{aligned} Pr^{Synd}\left(\tilde{\gamma}(n^H) > \tilde{\gamma}(n^{NH}) - \kappa_2 | L_i = H\right) &= Pr(n^H > n^{NH} - \eta(\kappa_2)|L_i = H) \\ &= Pr(n^H > n^{NH} - \eta(\kappa_2)|L_i = H, n^{NH} > \Delta_1) \\ &\quad * Pr(n^{NH} > \Delta_1 | L_i = H) \\ &\quad + Pr(n^H > n^{NH} - \eta(\kappa_2)|L_i = H, n^{NH} \leq \Delta_1) \\ &\quad * Pr(n^{NH} \leq \Delta_1 | L_i = H) \\ &= (\frac{1}{2} + \eta(\kappa_1))(1 - \Delta_1) + \Delta_1 \end{aligned}$$

and

$$\begin{aligned} Pr^{Synd}\left((n^{NH}) - \kappa_2 < \rho \bigcap \tilde{\gamma}(n^H) < \tilde{\gamma}(n^{NH}) - \kappa_2 | L_i = H\right) &= Pr(n^{NH} < \tilde{\gamma}^{-1}(\rho + \kappa_2)|L_i = H, n^H < n^{NH}) \\ &\quad * Pr(n^H < n^{NH}|L_i = H) \\ &= \tilde{\gamma}^{-1}(\rho + \kappa_2)\frac{1}{2}(1 + \Delta_1). \end{aligned}$$

To see that the hub premium is smaller for investments syndicated by an intermediary relative to direct investment for investors in a hub region, note that:

$$\begin{aligned} Pr^{Synd}(H|L_i = H) - Pr^{Dir}(H|L_i = H) &= (\frac{1}{2} + \eta(\kappa_2))(1 - \Delta_1) + \Delta_1 + \tilde{\gamma}^{-1}(\rho + \kappa_2)\frac{1}{2}(1 + \Delta_1) \\ &\quad - \left[(\frac{1}{2} + \eta(\kappa_1))(1 - \Delta_1) + \Delta_1 + \gamma^{-1}(\rho + \kappa_1)\frac{1}{2}(1 + \Delta_1)\right] \\ &> 0 \\ if &\quad \kappa_2 < \tilde{\gamma}(\gamma^{-1}(\rho + \kappa_1)) - \rho \end{aligned}$$

For investors in a non-hub region, we have established that the probability of investing in a hub region under direct investment is:

$$\begin{aligned} Pr^{Dir}(H|L_i = NH) &= Pr\left(\gamma(n^H) - \kappa_1 > \gamma(n^{NH})|L_i = NH\right) \\ &\quad + Pr\left(\gamma(n^{NH}) < \rho \bigcap \gamma(n^H) - \kappa_1 < \gamma(n^{NH})|L_i = NH\right) \\ &= \frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_1) + \Delta_2\eta(\kappa_1) + \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\gamma^{-1}(\rho)\Delta_2) \end{aligned}$$

For investors in a non-hub region, the probability of investing in a hub region with an intermediary is similarly given as:

$$
\begin{aligned}
Pr^{Synd}(H|L_i = NH) &= Pr\left(\tilde{\gamma}(n^H) - \kappa_2 > \tilde{\gamma}(n^{NH})|L_i = NH\right) \\
&\quad + Pr\left(\tilde{\gamma}(n^{NH}) < \rho \bigcap \tilde{\gamma}(n^H) - \kappa_2 < \tilde{\gamma}(n^{NH})|L_i = NH\right) \\
&= \frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_2) + \Delta_2\eta(\kappa_2) + \frac{1}{2}\gamma^{\tilde{-}1}((\rho) + \frac{1}{2}\tilde{\gamma}^{-1}(\rho)\Delta_2)
\end{aligned}
$$

To see that the amount of hub premium is smaller for intermediated investments relative to direct investment for investors in a non-hub region, note that:

$$
\begin{aligned}
Pr^{Synd}(H|L_i = NH) - Pr^{Dir}(H|L_i = NH) &= \frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_2) + \Delta_2\eta(\kappa_2) + \frac{1}{2}\gamma^{\tilde{-}1}((\rho) + \frac{1}{2}\tilde{\gamma}^{-1}(\rho)\Delta_2) \\
&\quad - \left[\frac{1}{2} - \frac{1}{2}\Delta_2 - \eta(\kappa_1) + \Delta_2\eta(\kappa_1) + \frac{1}{2}\gamma^{-1}(\rho) + \frac{1}{2}\gamma^{-1}(\rho)\Delta_2)\right] \\
&> 0 \\
if &\quad \kappa_2 < \tilde{\gamma}(\gamma^{-1}(\rho + \kappa_1)) - \rho
\end{aligned}
$$

8.1.3 Extension of the Theoretical Framework

The objective of this section is to provide an extension of our theoretical framework (Section 3) by explicitly incorporating the process of face-to-face due diligence, and observable startup and intermediary quality into the model. We start by writing out the return function for investors under direct investment:

$$
\Pi_i^D = max\{f_1(n_i^{Hub}), f_2(n_i^{NHub}), \phi^H + \rho\} - \kappa_1 d_{id}.
$$

The return function of investing directly using their professional network is

$$
\begin{aligned}
f_1(n_i^{Hub}) &= \gamma_1(n_i^{Hub})max\{\phi^H + \rho, (\phi^L + \rho + \epsilon_d)\mathbb{1}_{d_{id}=0}\} \\
f_2(n_i^{NHub}) &= \gamma_2(n_i^{NHub})max\{\phi^H, (\phi^L + \epsilon_d)\mathbb{1}_{d_{id}=0}\},
\end{aligned}
$$

Compared to the return function from Section 3, we introduce a few changes. First, the return functions $f_1(\cdot)$ and $f_2(\cdot)$ are now the product of two components, the old $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$, and a new component which captures investors' choices for startups with a high quality signal (ϕ^H) or with a low quality signal (ϕ^L). Note that the network terms enters multiplicatively as a scaling factor – an "iceberg" search friction – and the investment return is scaled down by this factor. Startup quality is the sum of its observable signal, plus a term that captures its unobservable quality. Without loss of generality, we define ϵ_d to be the unobservable quality term of low-signal local startups relative to that of high-signal local startups, and $\epsilon_d \in \{-\alpha, 0, \alpha\}$. This term explicitly captures the possible existence of local information on the quality of a startup. Local investors observe a noisy signal of a startup's unobservable quality $\widehat{\epsilon_i}$ with precision p.

$$
\widehat{\epsilon_i} = \begin{cases} \epsilon_d & \text{with probability p,} \qquad \epsilon_d \in \{-\alpha, 0, \alpha\} \\ 0 & \text{with probability 1-p} \end{cases}
$$

This leads to a new proposition:

Proposition 3 *(Geography of Direct Investments by Startup Quality) In direct investments, the hub premium (HP^1) is higher for startups with a low quality signal.*

Proof. *(Informal proof)* A high quality signal is observed and desirable to both local and non-local investors. Local investors, however, have a chance to learn the unobservable quality of local startups from local incidental information. This increases the value of investing in low-signal startups for local investors relative to distant investors. Therefore, the local premium is higher for startups with a low quality signal. An increased local premium in peripheral regions reduces the probability that investors randomly pick a startup in the hub region, therefore decreasing the hub premium. ■

We then turn to intermediated investments. These investments involve two steps, namely investors choosing intermediaries and then intermediaries choosing startups. We start by explicitly modelling the second part:

$$\mathcal{P}_s = max\{g_1(n_s^{Hub}), g_2(n_s^{NHub}), \phi^H + \rho\} - \kappa_2 d_{sd}.$$

The return function for an intermediary s, $\mathcal{P}_s$, is similar to that for investors in that the intermediary can choose to invest using her professional network, or through a random draw from startups with a high signal in a hub region. A key difference is that intermediaries have potentially different return functions from their network, g_1 and g_2, which are given as

$$\begin{aligned} g_1(n_i^{Hub}) &= \delta_1(n_i^{Hub}) max\{\phi^H + \rho, (\phi^L + \rho + \epsilon_d)\mathbb{1}_{d_{id}=0}\} \\ g_2(n_i^{NHub}) &= \delta_2(n_i^{NHub}) max\{\phi^H, (\phi^L + \epsilon_d)\mathbb{1}_{d_{id}=0}\}, \end{aligned}$$

where the search friction depends on their network, $\delta_1(\cdot)$ and $\delta_2(\cdot)$. Intermediaries differ on three dimensions: their location, their search cost based on an unobservable number of relevant professional connections, and their unobservable ability to perform due diligence. Therefore, for a given location, there are four types of intermediaries based on combinations of their observable signals and unobservable characteristics.

For intermediaries with high ability for conducting due diligence, the signal for unobservable startup quality $\widetilde{\epsilon_s}$ is more precise. The signals for the two types are given as:

$$\widetilde{\epsilon_s^L} = \begin{cases} \epsilon_d & \text{with probability p,} \qquad \epsilon_d \in \{-\alpha, 0, \alpha\} \\ 0 & \text{with probability 1-p} \end{cases}$$

$$\widetilde{\epsilon_s^H} = \begin{cases} \epsilon_d & \text{with probability p+}\Delta\text{p,} \qquad \epsilon_d \in \{-\alpha, 0, \alpha\} \\ 0 & \text{with probability 1-p-}\Delta\text{p} \end{cases}$$

This leads to a new proposition linking the geography of intermediated investments to the strength of startup quality signals:

Proposition 4 *(Geography of Intermediaries' Choices of Startups by Startup Quality Signal) When intermediaries choose startups, the local premium (LP^3) is higher for startups with low quality signals.*

Proof. *(Informal proof)* High quality signals are observed and desirable to both local and non-local intermediaries. However, local intermediaries have a chance to learn the unobservable quality of local startups from local incidental information. This increases the value of choosing low-signal startups for local intermediaries relative to non-local ones. Therefore, the local premium (LP^3) is higher for startups with low signals. ■

Having discussed how intermediaries choose startups, we now turn to the first step: decisions made by investors selecting intermediaries. Conditional on using intermediaries, investors can choose intermediaries from hubs or non-hub regions, and with high or low signals. The investor's return function for investments syndicated by an intermediary Π_i^S is given as:

$$\begin{aligned}
&\Pi_i^S = (1-\tau)max\{\widetilde{f_1}(n_i^{Hub}), \widetilde{f_2}(n_i^{NHub})\}, where \\
&\widetilde{f_1}(n_i^{Hub}) = \widetilde{\gamma_1}(n_i^{Hub})max\{n_s^H, (n_s^L + \widetilde{\epsilon_i})\mathbb{1}_{d_{is}=0}\} \\
&\widetilde{f_2}(n_i^{NHub}) = \widetilde{\gamma_2}(n_i^{NHub})max\{n_s^H, (n_s^L + \widetilde{\epsilon_i})\mathbb{1}_{d_{is}=0}\},
\end{aligned}$$

Intermediaries have high or low observable signals based on their number of professional connections, n_s^H and n_s^L. They also differ in unobservable quality $\widetilde{\epsilon_s^H}$ and $\widetilde{\epsilon_s^L}$. This generates a few new propositions related to geography, startup quality signals, and intermediaries' quality signals.

Proposition 5 *(Geography of Investors' Choice of Intermediaries) The amount of local premium observed when investors select intermediaries (HP^2) is higher for intermediaries with low signals.*

Proof. *(Informal proof)* Intermediaries with a stronger professional network have stronger signal and are appealing to both local and non-local investors. However, local investors may also be able to get a signal of the intermediaries' unobservable quality (i.e. their ability to perform due diligence and source high quality deals). This increases the value and therefore the likelihood they will select local intermediaries with low signals in cases where these intermediaries have high unobservable quality, but are only detected as high types by local investors. ■

Proposition 6 *(Geography of Intermediaries' Choices of Startups and Investors' Choice of Startups). The local premium is stronger when intermediaries allocate capital to startups (LP^3) relative to when the online crowd does so directly (LP^1).*

Proof. *(Informal proof)* Intermediaries with high unobservable quality are better at due diligence, i.e. they have higher precision in their ability to observe the true quality of startups relative to investors. When local startups have higher quality (observable signal plus unobservable quality), then intermediaries are more likely to invest in them. Otherwise, intermediaries are less likely to invest locally. ■

Proposition 7 *(Difference between Intermediaries' and Investors' Local Premiums) The difference in local premiums when intermediaries allocate funds and when investors allocate funds ($LP^3 - LP^1$) is larger for startups with low quality signals.*

Proof. *(Informal proof)* Intermediaries with high unobservable quality have higher precision in detecting high quality local statups relative to investors. Therefore intermediaries are able to identify high-quality local startups with low signal when these are indeed high quality. ■