The Review of Financial Studies

Social Proximity to Capital: Implications for Investors and Firms

Theresa Kuchler
Stern School of Business, New York University, NBER, and CEPR

Yan Li
School of Accounting, Southwestern University of Finance and Economics

Lin Peng
Zicklin School of Business, Baruch College

Johannes Stroebel
Stern School of Business, New York University, NBER, and CEPR

Dexin Zhou
Zicklin School of Business, Baruch College

We show that institutional investors are more likely to invest in firms from regions to which they have stronger social ties but find no evidence that these investments earn a differential return. Firms in regions with stronger social ties to locations with many institutional investors have higher valuations and liquidity. These effects are largest for small firms with little analyst coverage, suggesting that the investors' behavior is explained by their increased

We thank Elsa Allman, Yakov Amihud, Nate Baum-Snow (discussant), Zhi Da, Phil Dybvig, Patrice Fontaine, Carolina Fugazza (discussant), Edward Glaeser, Paul Gompers (discussant), Itay Goldstein, Abhinav Gupta, Arpit Gupta (discussant), Zhiguo He, Rawley Heimer, Xing Huang (discussant), Chih-Ching Hung, Narasimhan Jegadeesh, Marcin Kacperczyk, Yelena Larkin, Gabriele Lattanzio (discussant), Charles Lee, Ramana Nanda, Marina Niessner (discussant), Cameron Peng, Veronika Pool (discussant), Wenlan Qian, Oliver Rehbein, Simon Rother, Noah Stoffman, Wei Xiong, Hongjun Yan, and Jun Yang, as well as seminar participants at Baruch College, Cambridge University, Imperial College, NYU Stern, New York Fed, UC Irvine, Bonn University, Chicago Fed, Temple University, University of Miami, University of Rochester, University of Indiana, Renmin University, University of Virginia, HKUST, University of Connecticut, Dartmouth College, CUHK, DIW Berlin, INSEAD, PUC Chile, and Florida State University and conference participants at AFA, AREUEA National, EFA, SFS Cavalcade, NBER Urban Economics SI, Annual Conference in Financial Economics Research, Paris December Finance Meeting, FMA, UEA, Econometric Society Winter Meeting, and the International Macroeconomics and Finance Conference for helpful comments. Lin Peng gratefully acknowledges financial support from the Krell research grant and the Wasserman research grant. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Dexin Zhou at dexin.zhou@baruch.cuny.edu.

The Review of Financial Studies 35 (2022) 2743–2789

doi:10.1093/rfs/hhab111 Advance Access publication September 20, 2021

awareness of firms in socially proximate locations. Our results highlight that the social structure of regions affects firms' access to capital and contributes to geographic differences in economic outcomes. (*JEL* G2, G3, G4)

Received May 26, 2020; editorial decision July 14, 2021 by Editor Itay Goldstein. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

There are large regional differences in economic outcomes across the United States. For example, firms located in metropolitan and coastal regions are often more productive, more innovative, and more valuable than firms in other parts of the country (see Baum-Snow and Pavan 2011), and Moretti (2012). A growing literature explores a variety of explanations for these disparities. For example, Dougal, Parsons, and Titman (2018) argue that coastal cities with better amenities are able to attract more high-skilled workers, allowing firms from these areas to capture some of the resulting increases in productivity. In this paper, we propose a new and complementary explanation. In particular, we argue that the geographic structure of a region's social network influences the allocation of capital to local firms, and thereby contributes to the observed differences in firm outcomes.

We first show that—conditional on physical distance and other controls—institutional investors are more likely to invest in firms located in regions to which they have stronger social ties, with larger effects for investments in smaller firms with lower analyst coverage. We further document that firms in regions with stronger social links to the headquarters' locations of institutional investors have meaningfully higher liquidity and higher valuations, and argue that this relationship results from the increased social proximity to capital instead of omitted variables. Despite the strong observed effects of social ties on investor behavior and firm outcomes, investors do not generate differential returns from their investments in socially proximate regions. Based on these and other findings, we conclude that social connectedness raises investors' awareness of lesser-known firms—and thus increases their likelihood of investing in those firms—without providing the investors with an informational advantage.

To measure the social connectedness between firm and investor locations, we use the Social Connectedness Index (SCI), introduced by Bailey et al. (2018b). This measure is based on friendship links on Facebook, the world's largest online social networking service, with 239 million active users in the United States and Canada as of the end of 2017. Given Facebook's scale, the relative representativeness of its user body, and the fact that Facebook is primarily used to connect real-world friends and acquaintances, the SCI provides a comprehensive measure of the geographic structure of U.S. social networks. We quantify the social connectedness between a firm and an institutional investor as the relative probability of two Facebook users located in the headquarters' counties to be connected via a friendship link on Facebook.[1]

Our focus on the role of social connections in shaping public equity investments by institutional investors is driven by three considerations. First, institutional investors play a key role in providing capital, market liquidity, and corporate governance for U.S. firms (see Gompers and Metrick 2001; Blume and Keim 2012; Aghion, Van Reenen, and Zingales 2013). Access to institutional capital can therefore be crucial for firms to finance their operations and grow. Second, data on both the headquarters' location and the investments of institutional investors are publicly available. Third, there is substantial geographic variation in the location of institutional investors across the United States. For example, the Tri-State area (i.e., New York, New Jersey, and Connecticut), with the largest concentration of institutional assets under management (AUM), accounts for only about one-third of total U.S. institutional AUM. This variation in the location of institutional investors, combined with substantial differences in the geographic structure of social networks across U.S. counties, creates sizable variation in firms' social proximity to institutional capital.

We first analyze institutional portfolio holdings as of June 2016 and document that institutional investors tend to invest more in firms located in regions that the investors are socially connected to. In our baseline specification, a 10% increase in the social connectedness between firm and investor locations is associated with a 1.9% increase in the weight of the firm in the investor's portfolio. Our extensive set of fixed effects and controls alleviates concerns that our results might be picking up possible confounding factors, such as investors locating in regions that are socially connected to industry clusters important to the investors' mandates. We also show that our results are not driven by similarity between institution and firm headquarters counties or by institutions being more aware of a firm due to the firm's economic presence in an institution's headquarters' state.

An influential literature has documented that investors have a preference for geographically proximate firms, a feature often referred to as "home bias" or "local bias" (e.g., Coval and Moskowitz 1999; Bernile, Kumar, and Sulaeman 2015).[2] Consistent with this literature, we find that investments are decreasing in the geographic distance between firms and investors. However, when we

[1] Present-day friendship links between regions are determined by many factors, including historical migration patterns. For example, the Great Migration of African Americans from southern to northern industrial cities in the 1940s-1960s shows up as stronger present-day friendship links between Chicago and Mississippi. As a result, we would argue that the investor Northern Trust, based in Chicago, is disproportionately connected to the firm Trustmark Corporation, based in Mississippi. See Bailey et al. (2018b, 2019b) for detailed discussions of the determinants of social connectedness between locations.

[2] See also Huberman (2001), Ivković and Weisbenner (2005), Massa and Simonov (2006), Baik, Kang, and Kim (2010), Seasholes and Zhu (2010), and Becker, Cronqvist, and Fahlenbrach (2011). In addition to domestic home bias or local bias in investing, a lack of international diversification is documented and discussed in French and Poterba (1991), Cooper and Kaplanis (1994), Kang, Stulz et al. (1997), and Coeurdacier and Rey (2013).

control for the social connectedness between firm and investor locations, the negative effect of physical distance on investments disappears. Moreover, the inclusion of controls for physical distance does not affect the estimated effect of social connectedness on investments. These findings suggest both that in prior studies of local bias, physical proximity largely served as a proxy for social proximity and that investors are more likely to invest in geographically close firms because they are more likely to hear about these firms through their social networks.

Next, we explore heterogeneity in the relationship between social connectedness and investments along characteristics of both investors and firms. Social connections have the largest effect on investments in firms that are small or have low analyst coverage. While investors are likely to be less familiar with these firms on average, our findings suggest that social connections allow investors to become aware of firms' existence or to have better true or perceived information about them. In addition, social connectedness has larger effects on the investments of institutions that rely more on nonfinancial and intangible factors rather than quantitative measures. Smaller institutional investors with fewer resources are also more likely to overweight firms from locations they are socially connected to.[3]

To further minimize concerns about omitted variables at the investor-firm-pair level, we next explore panel data of institutional investment holdings between June 2007 and December 2016. This analysis allows us to directly include institution $\times$ firm fixed effects to capture any time-invariant determinants of an institution's preference for holding a particular stock. We find that time-series variation in social connectedness within an institution-firm pair—variation that is driven by changes in the headquarters' locations of firms—continues to explain investment patterns: when a firm moves its headquarters from a location that is weakly connected to an investor to a location that is more strongly connected to that investor, the investor increases its investment in that firm. This effect increases with the time since the headquarters move. Together with our heterogeneity analysis, this result dramatically reduces the scope for potential omitted variables to explain the observed relationships between social connectedness and investment behavior. Instead, our findings are more consistent with a causal effect of social connectedness on investment decisions, most likely because social connections raise an investor's awareness of small and lesser-known firms in those locations where the investor has friends.

We next ask whether the tendency of institutional investors to invest disproportionately in areas to which they are socially connected aggregates

[3] This finding is consistent with the results of Pool, Stoffman, and Yonker (2012), who document that managers from smaller fund families disproportionately overweight firms from their home states in their investment decisions. Similarly, Hirshleifer, Peng, and Wang (2020) find that earnings announcements made by firms with greater social network centrality attract more attention from both institutional and retail investors, with larger effects for smaller firms.

up to affect equilibrium capital market outcomes for firms. That is, do firms located in regions with higher social proximity to institutional capital attract more overall institutional investment? And, since institutional investors play an essential role in financial markets, does social proximity to capital affect other firm-level capital market outcomes such as valuations and secondary market liquidity?

To address these questions, we first construct a measure of each location's social proximity to institutional capital. Specifically, for each county, we calculate the weighted average institutional AUM in all other U.S. counties, where the weight is the social connectedness between the focal county and the other counties. Under this measure, counties with more friendship links to locations with high-AUM institutions are said to be more socially proximate to capital. We also construct a corresponding measure of physical proximity to institutional capital as a key control variable.

We first show that our institution-firm-pair-level results aggregate up and that, conditional on physical proximity to capital and other controls, firms located in regions that are more socially proximate to institutional capital have higher institutional ownership. Quantitatively, a 10% increase in social proximity to capital is associated with a 20.4 basis points increase in total institutional ownership.

We then examine whether social proximity to capital affects firms' valuations. There are at least two possible mechanisms for such a relationship. First, since firms in regions with more social connections to capital are more broadly held by institutional investors, these investors can better share those firms' risks and would thus demand a lower rate of return. This is similar to predictions from the equilibrium model in Merton (1987), in which more widely known firms have larger investor bases, which results in better risk sharing and higher valuations. Second, in the presence of short-sale constraints, valuations disproportionately reflect the assessments of the most optimistic investors (see Miller 1977; Scheinkman and Xiong 2003). Since firms with a broader investor base are also more likely to attract the attention of particularly optimistic investors, this provides a second channel through which valuations might increase with social proximity to capital.

Consistent with these potential mechanisms, we find that firms from regions that have stronger social links to the locations of large investors have higher valuations. Quantitatively, a 10% increase in the social proximity to capital is associated with a 1.1% increase in a firm's market-to-book ratio. These results are robust to including a large number of time-varying firm and county controls, as well as controlling for the physical proximity to capital and state $\times$ industry fixed effects. The results are also robust to including firm fixed effects in a specification that only exploits variation in a firm's social proximity to capital coming from firms moving headquarters and investors changing their AUM over time. We also find that the effect of social proximity to capital on firm valuation is generally higher for smaller firms with lower analyst coverage, precisely those firms for which we previously found the largest effects of social connectedness on institutional investment flows.

Given that institutional investors play an important role in liquidity provision (Blume and Keim 2012; Rubin 2007), we also analyze the effect of social proximity to capital on firms' secondary market liquidity. We find that a 10% increase in social proximity to capital is associated with a 0.94% reduction in effective spreads. As before, these results survive the addition of firm fixed effects and are generally larger for smaller firms with lower analyst coverage.

Our empirical approach does not exploit quasi-random variation in social proximity to capital to estimate its effects on secondary market liquidity and valuations. In fact, it is unlikely that any such variation exists, since social ties are rarely, if ever, randomly assigned. Instead, we control for many firm and county characteristics to absorb other factors that could affect our outcomes of interest. Nevertheless, one might worry that there is something about firms in counties with a high social proximity to capital that leads them to have higher liquidity and valuations independent of the effects of social connectedness on access to capital. For example, it could be that firms in places with high social proximity to capital just happen to be less risky for reasons that are not accounted for by the comprehensive set of control variables. One might also speculate that firms in places with high social proximity to capital are more well known in general, conditional on industry and other controls, and thus will attract higher liquidity provision from all investors. Such a concern may even extend to our specifications with firm fixed effects, where one might attempt to argue—even if not very persuasively—that firms with independently rising valuations would end up moving their headquarters to counties with higher social proximity to capital. We find such reasoning to be unconvincing. It would not only fail to explain our earlier result that it is only the connected institutions that invest more in these firms, but could also not explain why small firms benefit disproportionately from high social proximity to capital.

Nevertheless, to alleviate such concerns, we exploit an exogenous shock to a subset of capital providers that resulted in differential cross-sectional liquidity impact due to firms' heterogeneous social links to areas affected by the shock. Specifically, we study the effects of Hurricane Sandy, the second-costliest hurricane in U.S. history. Sandy's landfall in the mid-Atlantic region on October 22, 2012, resulted in severe disruptions to the Tri-State area's commuting networks.

We first provide evidence that, during this period, institutional investors in the affected mid-Atlantic region substantially reduced their liquidity provision. We then focus on the liquidity dynamics of firms located in areas that were not directly affected by Sandy. Consistent with our overall narrative, we find that during Hurricane Sandy, firms with high social proximity to institutional capital in the mid-Atlantic states experienced a relative reduction in their secondary market liquidity compared to firms with the same overall social proximity to capital but lower social exposure to the mid-Atlantic states. This finding provides further support that our results are not driven by omitted firm-level characteristics that affect the liquidity provision by all investors. Instead, the variation provided by Hurricane Sandy reinforces the notion that what matters for explaining the higher liquidity of firms with greater social proximity to capital is the liquidity provision from investors in those parts of the country to which a firm has social connections.

Last, we explore the implications of our results for the investment performance of institutional investors. Do investors achieve higher returns by allocating their investments to regions they are socially connected to? If institutional investors obtain an information advantage through their social connections, we should observe higher risk-adjusted returns for institutions that invest more in areas they are socially connected to (see Cohen, Frazzini, and Malloy 2008; Hong, Kubik, and Stein 2005; Hong and Xu 2019; Pool, Stoffman, and Yonker 2015). On the other hand, it is possible that investments in socially connected firms are not driven by superior information but are instead explained by investor awareness of such firms, as in Merton (1987). In this case, we would not expect investors to outperform when investing in socially proximate locations, and suboptimally diversified portfolios could even lead them to underperform (e.g., Huberman 2001; Massa and Simonov 2006; Seasholes and Zhu 2010; Pool, Stoffman, and Yonker 2015).

We examine institutional investors' performance along three dimensions: (i) across-institution comparisons of returns for investors with a differential propensity to invest in connected firms, (ii) within-institution comparisons between an institution's high-connectedness holdings and its low-connectedness holdings, and (iii) within-institution comparisons that evaluate an institution's high-connectedness holdings to high-connectedness stocks that the institution chooses not to hold. Across these tests, we find no evidence that institutional investors obtain valuable information through social connections as measured by Facebook friendship links.[4] Instead, our results are consistent with an interpretation in which institutional investments in socially proximate firms are driven by an increased awareness of these firms.

Our paper contributes to a literature on how social interactions affect economic decisions, including investment decisions.[5] While a number of papers in this literature have documented that various social interactions can affect

[4] The diverging conclusions in the literature on whether interactions through social networks can help improve investment performance may be because different types of social networks vary in their ability to convey useful information. For example, our results show that friendship networks as measured by Facebook do not appear to convey useful information, even though they affect investment patterns. This is not inconsistent with other evidence that suggests that professional networks are more able to convey useful information (see, e.g., the evidence on this presented in Cohen, Frazzini, and Malloy 2008).

[5] See, e.g., Kelly and O'Grada (2000), Duflo and Saez (2002), Duflo and Saez (2003), Hong, Kubik, and Stein (2004); Cohen, Frazzini, and Malloy (2008, 2010), Cohen, Gurun, and Malloy (2017), Hochberg, Ljungqvist, and Lu (2007), Brown et al. (2008), Shive (2010), Chen et al. (2010), Kaustia and Knüpfer (2012), Shue (2013), Banerjee et al. (2013), Pool, Stoffman, and Yonker (2015), Heimer (2016), Ahern (2017), Crawford, Gray, and Kern (2017), Maturana and Nickerson (2018), Bailey et al. (2018a), Hirshleifer, Peng, and Wang (2020), Ouimet and Tate (2020). See Hirshleifer (2020) and Kuchler and Stroebel (2020) for reviews.

investment choices, our novel measure of social connectedness allows us to provide the first evidence that these effects can aggregate up to influence equilibrium capital market outcomes for firms. By doing so, our work suggests one important mechanism through which the geographic structure of social networks can shape regional variation in economic outcomes.

1. Social Connectedness and Institutional Investments

We first document that institutional investors are more likely to invest in firms located in counties to which they have stronger social ties. We begin by describing our measures of social connectedness between U.S. counties, as well as the construction of our investment variables.

1.1 Data and measurement

1.1.1 Measuring social connectedness. To measure social connectedness between U.S. counties, we use the Social Connectedness Index first introduced by Bailey et al. (2018b).[6] This measure was created using anonymized information on the universe of friendship links between U.S.-based Facebook users as of April 2016. Facebook is the world's largest online social networking service: by the end of 2017, it had more than 2.1 billion monthly active users globally and 239 million active users in the United States and Canada. A survey of Facebook users from 2015 found that more than 68% of the U.S. adult population and 79% of online adults in the United States used Facebook (Duggan, Greenwood, and Perrin 2016). That same survey showed that Facebook usage rates among U.S.-based online adults were relatively constant across income groups, education levels, and race, as well as among urban, rural, and suburban residents; usage rates were slightly declining in age. In the United States, Facebook mainly serves as a platform for real-world friends and acquaintances to interact online, and people usually add connections on Facebook only to individuals they know in the real world. As a result, networks formed on Facebook more closely resemble real-world friendship networks than those on other online platforms, such as Twitter, where unidirectional links to nonacquaintances are common. Consistent with this, Bailey et al. (2018b,a, 2019a,b,c, 2021, 2020a,b), Chetty et al. (2021), Kuchler, Russel, and Stroebel (2021), and Rehbein and Rother (2020) provide evidence that friendships observed on Facebook are a good proxy for real-world U.S. social connections.

To construct the Social Connectedness Index, Bailey et al. (2018b) map Facebook users to their respective county locations using information such as the users' regular IP addresses. They then construct a measure of the relative number of friendship links between each county pair, $Friendships_{i,j}$. Our

[6] The Social Connectedness Index data for the United States and many other countries are available freely and without usage restrictions at https://data.humdata.org/dataset/social-connectedness-index.

measure of social connectedness between two counties corresponds to the (relative) probability that a Facebook user in county i is friends with a Facebook user in county j:

$$Social\ Connectedness_{i,j} = \frac{Friendships_{i,j}}{Population_i \times Population_j}, \tag{1}$$

where $Population_i$ corresponds to the total population in county i.[7] Figure 1 shows heat maps of our measure of $Social\ Connectedness_{i,j}$ for San Francisco County, California, in panel A and for Cook County, Illinois, in panel B. Darker colors correspond to stronger social connections to the focal counties. Both San Francisco County and Cook County are home to a substantial number of institutional investors, so these maps show differences in the relative connectedness to institutional capital in those locations. San Francisco County is strongly connected to nearby counties in coastal California. However, social connectedness is not determined solely by physical proximity. For example, San Francisco County also has strong connections to other urban areas, such as New York and Chicago, as well as to college towns across the United States. This is likely driven by connections of college graduates moving from college towns to work in San Francisco. Cook County, which includes the city of Chicago, is strongly connected to counties in the southern states along the Mississippi River. This pattern is likely the result of the large-scale migration of African Americans from southern states to northern industrial cities during the Great Migration (1916–70). More generally, these plots show that two adjacent counties can have very different social connectedness to institutional investors in San Francisco and Chicago. Such variation will help us distinguish between the effects of physical proximity and social proximity to capital.

1.1.2 Institutional holdings data. We obtain information on institutional investors' holdings from the Thomson Reuters Institutional (13F) Holdings data set. Information is reported at the level of the fund family, not the level of the individual fund. In our baseline analysis, we analyze institutional investments from June 2016, which most closely corresponds to the time when we observe social connectedness. We also expand our analysis to panel regressions with holdings data from 2007 to 2016. We combine institutional investors' holdings data with information on stock prices from CRSP to construct measures of the total investment by each fund in each firm.[8] In particular, for each institution-firm pair, we construct a measure of institutional holding, $\%PF_{i,j}$,

7 Subsequent research has confirmed that in the United States, this Social Connectedness Index is nearly perfectly correlated with a version of the index that has the product of the number of Facebook users in the denominator (Bailey et al. 2021).

8 We limit our analysis to stocks listed on NYSE, NASDAQ, and NYSE American that have a price greater than \$5. We also only consider fund families that hold at least five stocks. We only analyze firms and funds located in the 48 contiguous U.S. states.

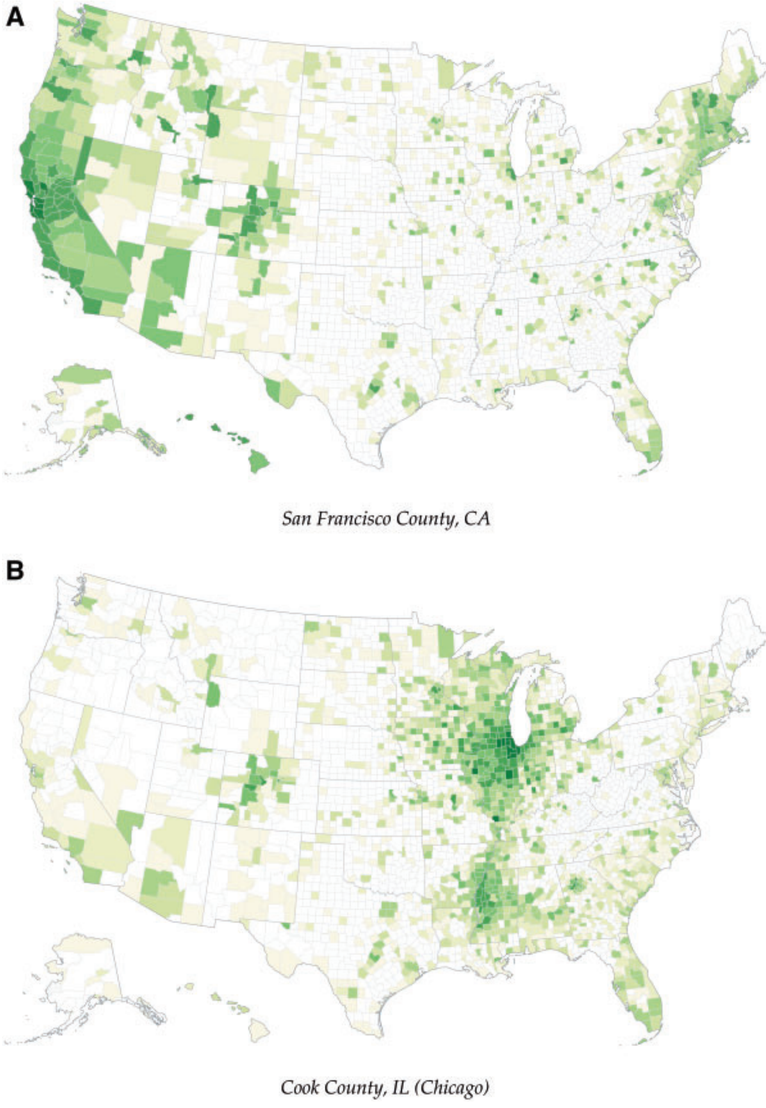

Figure 1
Examples of social connectedness
This figure shows county-level heat maps of the social connectedness to San Francisco County, CA, in panel A, and Cook County, IL, in panel B. Darker colors indicate higher social connectedness to the focal county.

which corresponds to the share of firm i in investor j's portfolio, where each institution's assets under management—its AUM—is the sum of the equity values held by that investor:

$$\%PF_{i,j} = \frac{\text{Ownership (\$) of investor } i \text{ in firm } j}{\text{AUM (\$) of investor } i}. \tag{2}$$

We obtain institutional investors' headquarters' locations from Bernile, Kumar, and Sulaeman (2015) and Bernile et al. (2019), who collect this information from Nelson's Directory of Investment Managers and by searching SEC filings.[9] We obtain firms' historical headquarters' locations from Compustat.

Overall, we have information on 3,083 firms and 2,820 fund families. Panel A of Table 1 presents summary statistics on institution-firm pairs (see also Internet Appendix Table IA.1). The mean of *%PF* indicates that for the average firm-institution pair, the firm constitutes 0.04% of the institution's public U.S. equity portfolio. Most firm-institution pairs (up to the 90th percentile) have zero investments by the fund in the firm, consistent with the fact that many institutions do not hold highly diversified portfolios.

1.2 Empirical analysis: Baseline specification

We use the following regression specification to investigate how the social connectedness between the location of firm i's headquarters and the location of institutional investor j's headquarters affects investor j's decision to invest in firm i:

$$\%PF_{i,j} = \exp[\beta \, \mathit{Log\ Social\ Connectedness}_{i,j} + \gamma X_{i,j} + \psi_i + \xi_{j \times ind(i)}] \cdot \epsilon_{i,j}. \tag{3}$$

This functional form is motivated by the binscatter plots in Figure 2, which suggest a linear relationship between *Log %PF* and $\mathit{Log\ Social\ Connectedness}_{i,j}$, both with and without controlling for the geographic distance between firms and investors. The vector $X_{i,j}$ includes controls for various measures of the geographic distance between firm i and investor j, as well as indicator variables for whether the firm and investor are located in the same state or county. Our baseline specification also includes firm and institution $\times$ firm Fama-French 48 industry fixed effects. These fixed effects ensure that our findings are not explained either by institutions choosing to locate in areas that are socially connected to industry clusters relevant to the institutions' mandate or by characteristics that might make firms located in socially connected counties more prone to attracting institutional investments on average. Regression (3) is estimated using Poisson Pseudo Maximum Likelihood (PPML) to account for the censoring of investments at zero.[10] We cluster standard errors by firm and institution.

---

[9] We are grateful to Gennaro Bernile, Alok Kumar, and Johan Sulaeman for sharing these data sets. We extend the data set by collecting additional institutional location data from SEC filings.

[10] This estimation procedure is widely used in the trade literature, which also faces a left-censoring of trade flows between countries (see the discussion in Silva and Tenreyro 2006; Bailey et al. 2021). We use the estimation procedure implemented in Correia, Guimarães, and Zylkin (2019).

Table 1
Summary statistics

A. Institution-firm observations (as of June 2016)

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
%PF	0.04	0.50	0	0	0	0	0.01
Log Social Connectedness	6.06	1.29	4.45	4.71	5.83	7.53	8.40
Log Distance	6.52	1.37	3.96	5.08	6.82	7.79	7.84

B. Firm-level variables (from 2007 to 2016)

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
%TIO	58.63	27.62	4.09	14.10	65.18	90.09	95.54
Log Effective Spread	1.02	1.19	−0.67	−0.39	0.88	2.75	3.24
Log M/B	0.70	0.89	−0.49	−0.25	0.57	1.79	2.25
Log Social Proximity to Capital	23.08	1.13	21.49	21.82	22.94	24.51	25.65
Log Physical Proximity to Capital	10.93	1.47	9.08	9.35	10.65	13.01	14.37

C. Institution characteristics, by institution type

		AUM (Million USD) as of June 2016						
Institutional type	N	MEAN	ST.DEV	P5	P10	MEDIAN	P90	P95
Dedicated	75	4,216	16,124	53	85	565	6,982	8,346
Quasi-Indexer	1,741	5,625	44,622	29	58	292	4,626	14,066
Transient	724	4,182	47,348	15	32	345	4,308	9,306
Not identified	543	273	1,108	7	15	88	373	722

This table reports summary statistics for our key variables. Statistics at the firm-institution level as of June 2016 are presented in panel A. *Social Connectedness* is defined as the number of Facebook links between a firm's headquarters' county and an institution's headquarters' county, scaled by the product of the populations in these two counties (multiplied by 10^{12}). *Distance* is the distance in miles between a firm's headquarters' county and an institution's headquarters' county. *Log Distance* is defined as *log(1+Distance)*. *%PF* is the percentage of AUM of an investor allocated to a given stock, where AUM is measured by the value of the institution's equity holdings. If an institution does not report holding in a given firm, *%PF* is equal to 0. Summary statistics for the firm-level variables are presented in panel B. Our quarterly firm-level panel spans from June 2007 to December 2016. *%TIO* measures the percentage ownership by all institutions, defined as the number of shares owned by all institutions, divided by the firm's shares outstanding. *Log Social Proximity to Capital* is $log(\sum[County\ AUM \times Social\ Connectedness])$, where County AUM caputres the AUM of all institutions headquartered in a given county and is measured in millions of USD and the summation is taken across all U.S. counties. *Log Physical Proximity to Capital* is $log(\sum[County\ AUM \div (1+Distance)])$, where the summation is taken across all U.S. counties. *Effective Spread* is the dollar-weighted percentage effective spread (multiplied by 10^3). *M/B* is defined as the ratio of market value of equity and book value of equity. Summary statistics for institutional investors based on their institution type are presented in panel C. Institution type is based on Bushee (2001). We report the number of institutions and the distribution of AUM for each institution type as of June 2016. The Appendix presents detailed variable definitions. Our sample includes institutions and firms located in the contiguous states of the United States. We require an institution to hold at least five different stocks. We study common stocks listed on NYSE, NASDAQ, and NYSE American (formerly AMEX). We exclude penny stocks (price $<$ \$5).

We report results from Regression (3) in Table 2. The first column shows our baseline estimates. The coefficient on $Log\ Social\ Connectedness_{i,j}$ is positive and statistically significant, consistent with institutional investors investing more in firms that are headquartered in counties to which the investors are socially connected. The coefficient estimate implies an elasticity of 0.189, suggesting that a 10% increase in social connectedness is associated with a 1.89% increase in $\%PF$.

In column 2 of Table 2, we examine the role of physical distance on investment choices. Consistent with the local bias literature, we find that

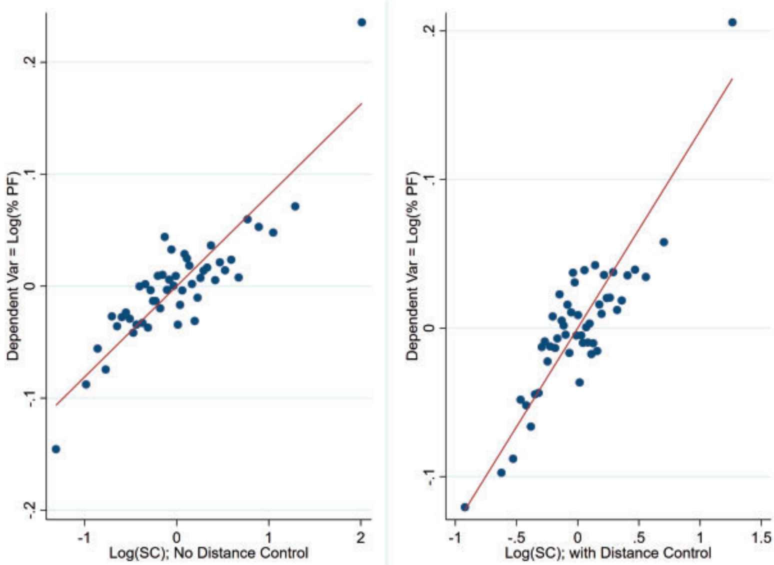

Figure 2
Binscatter plot
The vertical axis plots the log of the portfolio share of firm i in investor j's portfolio, $\%PF_{i,j}$, for nonzero investments. The horizontal axis plots the log of the social connectedness between the headquarters' locations of firm i and investor j. To produce these binscatter plots, *Log Social Connectedness* as of June 2016 is sorted into 50 bins. For each bin, the conditional mean of *Log Social Connectedness* and the conditional mean of *Log %PF*, is plotted as a scatter point. Each panel also includes the line of best fit from an OLS regression. In the left panel, we condition on firm fixed effects, institution × industry fixed effects, same state fixed effects, and same county fixed effects. In the right panel, we further condition on percentile distance dummies between the headquarters' locations of firm i and investor j.

institutional investors tend to overweight firms that are headquartered in close geographic proximity (see Coval and Moskowitz 1999, 2001; Baik, Kang, and Kim 2010; Bernile, Kumar, and Sulaeman 2015). Column 3 shows that, when we include both *Log Social Connectedness* and *Log Distance* as explanatory variables, *Log Social Connectedness* remains positively related to a firm's weight in institutional portfolios, with a coefficient estimate that is, if anything, slightly higher than that from column 2. The effect of geographic distance becomes smaller and even changes sign, suggesting that a large part of the mechanism driving local bias is that people are more likely to be socially connected to others who are geographically proximate.

In column 4 of Table 2, we control for geographic distance in a more flexible way to alleviate concerns that our measure of social connectedness might be picking up a potentially nonlinear relationship between physical distance and institutional investments. Specifically, we control for the distance between county pairs using dummies for percentiles of the distance distribution. The

Table 2
Social connectedness and institutional investment

	(1)	(2)	(3)	(4)	(5)	(6)
Log Social Connectedness	0.189***		0.253***	0.299***	0.294***	
	(14.97)		(10.65)	(11.86)	(11.08)	
Log Distance		−0.107***	0.054***			
		(−9.93)	(2.94)			
Log Social Connectedness × I(≤100 miles)						0.392***
						(4.06)
Log Social Connectedness × I(>100 miles)						0.102***
						(3.85)
Firm FE	YES	YES	YES	YES	YES	YES × SPLIT
Institution × Industry FE	YES	YES	YES	YES	YES	YES × SPLIT
Distance Percentile FE	NO	NO	NO	YES	YES	YES × SPLIT
Same State FE	NO	NO	NO	NO	YES	YES × SPLIT
Same County FE	NO	NO	NO	NO	YES	YES × SPLIT
N	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060
Pseudo R^2	0.506	0.504	0.506	0.507	0.507	0.549

This table explores the relationship between social connectedness and institutional investors' portfolio decisions. Estimates are obtained using PPML regressions. Our sample includes all firm-institution pairs in June 2016. The dependent variable is *%PF*, defined as the percentage of investor AUM allocated to a stock. If an institution does not report holdings in a given firm, *%PF* is assigned zero. *Log Social Connectedness* is defined as *log(Social Connectedness)*, where *Social Connectedness* is the number of Facebook links between a firm's headquarters' county and an institution's headquarters' county, scaled by the product of the populations in these counties. *Log Distance* is *log(1+Distance)*, where *Distance* measures the distance in miles between a firm's headquarters' county and an institution's headquarters' county. *I(≤100 miles)* (*I(> 100 miles)*) is an indicator variable that equals one when the institution-firm pair is physically less than (over) 100 miles apart. We consider Firm, Institution×Industry, Distance Percentile, Same County, and Same State fixed effects. *Same County* (*Same State*) is a dummy variable equal to one if the institution and the firm are located in the same county (state) and zero otherwise. *Distance Percentile* indicators are 100 dummy variables indicating the percentile of the distance between the firm and the institution based on all firm-institution pairs as of June 2016. The Appendix includes detailed variable definitions. Industry classification is based on Fama-French 48 industries. Standard errors are double clustered by institution and firm, and *t*-statistics are reported in parentheses. * $p < .1$; ** $p < .05$; *** $p < .01$.

coefficient of *Social Connectedness* remains highly significant and similar to that obtained in column 3.[11]

In the absence of reliable data on social networks, prior research on the effects of social interactions on investment decisions has explored co-location in the same city as a proxy for social connectedness.[12] Motivated by these studies, we next examine if our estimates are largely driven by whether the firm-institution pair is located in the same county or the same state (given that within-county and within-state connections are generally stronger than connections to individuals farther away). Column 5 includes same county or same state indicators and shows that the coefficient of *Social Connectedness* is almost identical to that in column 4, suggesting that our social connectedness measure is able to capture social interactions beyond these effects of shared locations.[13]

We further analyze how the effect of social connectedness on investments depends on the physical distance between the investor and the firm. Specifically, we define indicator variables for whether the investor-firm pair is located more or less than 100 miles apart. We then interact *Log Social Connectedness* with these indicator variables and present the results in column 6. Social connectedness has a larger effect on the probability of investing in firms that are physically closer to the institution, consistent with the ease of in-person interactions at short distances amplifying the effect of social connections (as argued, for example, by Hong, Kubik, and Stein 2005 and Pool, Stoffman, and Yonker 2015). However, even for institution-firm pairs that are more than 100 miles apart—distances at which in-person interactions are more costly—the effect of social connectedness on investments remains both economically meaningful and statistically significant. For our subsequent analysis, we present the average effect over firm-institution pairs at all distances, allowing us to capture the effects of social interactions on investment probabilities through both local and nonlocal social networks.

[11] Our results are not sensitive to the choice of number of distance indicators. In Internet Appendix Table IA.4, we consider alternative controls for physical distance using 5-, 10-, 50-, and 500-tile indicators as distance controls and obtain coefficient estimates for *Social Connectedness* ranging from 0.235 to 0.308.

[12] For example, Hong, Kubik, and Stein (2005) show that fund managers who work in the same city tend to have correlated holdings and trades, consistent with the social transmission of investment ideas by these managers. Using a more refined distance measure, Pool, Stoffman, and Yonker (2015) show that the portfolio overlaps of fund managers are more pronounced for those who reside in the same neighborhood, further suggesting that the correlated holdings are likely to be attributable to social interactions and less likely to be confounding local factors such as managers' exposure to common local media outlets, being visited by the same corporate executives during investor-relations roadshows, or by herding induced by job market segmentation.

[13] In the Appendix, we present additional findings to ensure the robustness of our baseline results to alternative specifications. Column 1 in Internet Appendix Table IA.2 reports the result when we drop firms and institutions located in the Tri-State area and California, two regions with a heavy presence of both investors and firms. Column 2 reports the result when we winsorize our dependent variable at the 99th percentile to reduce the effects of outliers. In column 3, we consider a county pair-level regression where the dependent variable is the fraction of the institution county's total AUM allocated to firm counties. Columns 4 and 5 report the result using OLS regressions instead of PPML regressions: in column 4, the dependent variable is $\%PF$, and in column 5, it is a dummy variable indicating whether there is nonzero institutional holding as the dependent variable.

1.2.1 Ruling out alternative interpretations. Our preferred interpretation of the previous findings is that investors are more likely to invest in firms headquartered in socially connected locations because they are more likely to become aware of these firms through their social networks. Before providing additional evidence in favor of this mechanism, we summarize evidence presented in Internet Appendix IA.1 that rules out two competing interpretations.

First, our results could be driven by a similarity in preferences between the residents of socially connected counties—after all, a large literature has shown that similar individuals are more likely to be friends. To explore whether such homophily can explain our results, we add control variables that capture similarities between institution and firm locations along the dimensions of population, age, employment, income, education, immigration history, marital status, political inclination, industry composition, and whether they both belong to metropolitan areas. As shown in columns 1 and 2 of Internet Appendix Table IA.3, the relationship between social connectedness and portfolio holdings is not affected by these additional controls.

Second, we consider the possibility that our results are driven by firms' economic links to institutions' locations. Such economic links could have an independent effect on investments if, for example, firms have major operations in socially connected locations, and institutions headquartered in those locations learn about these firms not through their social networks but by observing their activity in their own county. To explore whether this mechanism explains our findings, we explicitly control for two measures of the economic linkages between firm and investor counties: mentions of the investors' headquarters' states in the firms' 10-K filings (see Bernile, Kumar, and Sulaeman 2015) and the locations of firms' major customers. We find that our results are not affected by including these additional controls.

Thus, we conclude that the effect of social connectedness on investments does not capture direct economic connections or similar preferences. More generally, the stability of the estimates of β to the incremental addition of controls and fixed effects in Table 2 and Internet Appendix Table IA.3 reduces the likelihood that omitted unobserved characteristics confound our analysis (see Altonji, Elder, and Taber 2005; Oster 2019).

1.3 Heterogeneity analysis

A key takeaway from the previous analysis is that institutional investors tend to invest disproportionately in firms from areas that the investors are socially connected to. We argue that this is driven by an increase in investors' awareness of firms located in socially connected places. We next explore heterogeneity in this effect across characteristics of both firms and investors to provide further evidence for this mechanism behind the observed relationship between social connections and investments.

1.3.1 Heterogeneity by investor characteristics. We first split institutional investors based on their Bushee classification. Bushee (2001) groups fund families into three categories. “Transient” fund families are short-term-focused investors with high levels of portfolio turnover and diversification. “Quasi-indexer” families are characterized by low portfolio turnover and high diversification. “Dedicated” fund families tend to take large stakes in firms and have low portfolio turnover. They rely less on quantitative accounting measures and are more likely to use nonfinancial and intangible factors to make investment decisions. Based on these descriptions of investment strategies, our proposed channel should make the investments of “dedicated” institutions most sensitive to social ties, since the investment process for those fund families contains the largest scope for investment manager discretion.

To explore this prediction, we interact the Bushee-type indicators with $\mathit{Log\ Social\ Connectedness}_{i,j}$. The results are reported in columns 1 and 2 of Table 3. Column 1 shows the baseline result, where we control for firm $\times$ institution Bushee-type, institution $\times$ industry, and institution Bushee-type $\times$ distance percentile fixed effects. We find that, consistent with our hypothesis, “dedicated” institutions have the highest propensity to invest in firms headquartered in counties the investors are socially connected to. The investments of transient investors are least affected by social connectedness, and differences between investor types are large and statistically significant. Column 2 shows that this result remains robust after further controlling for same county and same state indicators, each interacted with institution Bushee-type.

In columns 3 and 4 of Table 3, we split investors into terciles based on their total AUM. We find that the investments of small institutions respond more than twice as much to social connectedness as the investments of large institutions do. Panel C of Table 1 shows that this is not just the result of “dedicated” investors having lower AUM. The split by institution size therefore provides additional information about the heterogeneity in institutions’ tendencies to invest based on social connectedness. As before, the heterogeneous response is robust to a variety of controls for the geographic distance between funds and firms. This result is consistent with the interpretation that small institutions have fewer resources to conduct large-scale research. Thus, their managers are more prone to invest in the limited number of stocks that they are aware of (e.g., as suggested by Pool, Stoffman, and Yonker 2012).

1.3.2 Heterogeneity by firm characteristics. Next, we explore the heterogeneity of investment sensitivity to social connectedness across various firm characteristics. Following Hong, Lim, and Stein (2000), Parsons, Sabbatucci, and Titman (2020), and others, we split firms based on their size and analyst coverage. There are at least two reasons why investments in small and informationally opaque firms may be disproportionately affected by social connectedness. First, to the extent that social connectedness increases investor awareness of firms—our preferred interpretation—these effects are likely to be less important for large and well known firms. Second, investors may perceive to have an information advantage for opaque firms they are socially connected

Table 3
Social connectedness and investment: by institution characteristics

		(1)	(2)		(3)	(4)
Log Social Connectedness						
	× Transient	0.149***	0.149***	× Small AUM	0.338***	0.329***
		(4.41)	(4.12)		(8.86)	(8.02)
	× Quasi-Index	0.318***	0.315***	× Mid AUM	0.305***	0.294***
		(11.63)	(10.95)		(8.31)	(7.70)
	× Dedicated	0.401**	0.375**	× Large AUM	0.235***	0.238***
		(2.56)	(2.27)		(6.56)	(6.35)
Institution Split × Firm FE		YES	YES		YES	YES
Institution × Industry FE		YES	YES		YES	YES
Institution Split × Distance Percentile FE		YES	YES		YES	YES
Institution Split × Same County FE		NO	YES		NO	YES
Institution Split × Same State FE		NO	YES		NO	YES
F test	No heterogeneity	17.67***	15.16***	Small = Large	4.74**	3.31*
N		7,162,800	7,162,800		8,694,060	8,694,060
Pseudo R^2		0.539	0.539		0.523	0.523

This table shows the results on how the effect of social connectedness varies by institutional investors' characteristics. The dependent variable is *%PF*, defined as the percentage of institutional AUM allocated to a stock. We report results on the heterogeneity across Bushee (2001) institution type. We interact *Log Social Connectedness* with Bushee-type dummies (Dedicated, Quasi-Indexer, and Transient) in columns 1 and 2. Institutions without a Bushee type are dropped from this sample. We report results on the heterogeneity across institution size terciles in columns 3 and 4. We interact *Log Social Connectedness* with dummy variables based on institution AUM terciles. *Log Social Connectedness* is defined as *log(Social Connectedness)*, where *Social Connectedness* is the number of Facebook links between a firm's headquarters' county and an institution's headquarters' county, scaled by the product of the populations in these two counties. We consider Firm, Institution×Industry, Distance Percentile, Same County, and Same State fixed effects, interacted with the respective institutional characteristics dummies. *Same County* (*Same State*) is a dummy variable equal to one if the institution and the firm are located in the same county (same state) and zero otherwise. *Distance Percentile* indicators indicate the percentile of the distance between the firm and the institution based on all firm-institution pairs as of June 2016. Refer to the Appendix for detailed variable definitions. Industry classification is based on Fama-French 48 industries. Standard errors are double clustered by institution and firm, and t-statistics are reported in parentheses. * $p < .1$; ** $p < .05$; *** $p < .01$.

to (though our results in later sections suggest that they do not, in fact, appear to have such an information advantage).

Columns 1 and 2 of Table 4 show the heterogeneity of the effect of social connectedness on investments when splitting firms into size terciles based on their market equity (Small Cap, Mid Cap, and Large Cap). The effect of social connectedness on investments is indeed declining in firm size. We also split firms based on analyst coverage, as analysts are an important information intermediary in financial markets, and analyst coverage can raise awareness of firms among investors. Columns 3 and 4 of Table 4 show that investments in low-analyst-coverage firms respond twice as much to social connectedness as investments in high-analyst-coverage firms do. Taken together, our firm heterogeneity tests suggest that investments in small firms with little analyst coverage are most strongly affected by social connectedness to potential investors, consistent with our interpretation that social interactions help raise investors' awareness of lesser-known firms.[14]

1.4 Within-firm identification using panel data

Our previous results show that institutional investors tend to invest disproportionately in firms located in counties that the investors are socially connected to. Despite the many controls in our cross-sectional analysis, one might worry about remaining omitted variables at the firm-investor-pair level that could correlate with the social connectedness between firm and investor locations, but that could also independently affect investors' propensities to hold stocks of a particular firm. We next address such concerns by exploiting within-firm-investor-pair variation in our measure of social connectedness generated by firms moving headquarters across locations.[15]

For this analysis, we explore quarterly panel data of institutional investment holdings between June 2007 and December 2016. The first three columns of Table 5 reproduce the specifications in columns 1, 4, and 5 of Table 2, respectively (since we introduce the time dimension, we now control for institution $\times$ quarter and firm $\times$ quarter fixed effects). These specifications confirm that the baseline results from the June 2016 cross-section are replicated in the panel, both qualitatively and quantitatively.

In the following specifications, we also include firm $\times$ institution fixed effects, which control for any time-invariant firm-institution-pair-specific

---

[14] Since firm size and analyst coverage are positively correlated, we further explore whether analyst coverage explains the relationship between social connectedness and investments beyond its correlation with firm size. We first perform an independent two-way sort that classifies firms into three terciles by firm size and analyst coverage, respectively. Then we map the 3×3 size–analyst coverage matrix into nine indicator variables and interact them with *Log Social Connectedness*. Results from this regression are reported in Internet Appendix Table IA.6. We find that even after controlling for firm size tercile, investment sensitivity to social connectedness is highest for firms with low analyst coverage.

[15] There is no time-series variation in the Social Connectedness Index, though prior work has shown that social connectedness between counties is highly stable over time (Bailey et al. 2021).

Table 4
Social connectedness and investment: by firm characteristics

		(1)	(2)		(3)	(4)
Log Social Connectedness						
	× Small Cap	0.565***	0.599***	× Low Coverage	0.568***	0.580***
		(7.62)	(7.72)		(8.27)	(7.99)
	× Mid Cap	0.448***	0.450***	× Mid Coverage	0.518***	0.526***
		(8.73)	(8.42)		(12.02)	(11.80)
	× Large Cap	0.259***	0.255***	× High Coverage	0.231***	0.222***
		(9.57)	(9.02)		(8.69)	(8.13)
Firm FE		YES	YES		YES	YES
Firm Split × Institution × Industry FE		YES	YES		YES	YES
Firm Split × Distance Percentile FE		YES	YES		YES	YES
Firm Split × Same County FE		NO	YES		NO	YES
Firm Split × Same State FE		NO	YES		NO	YES
F test	Small = Large	15.09***	17.48***	Low = High	21.03***	21.37***
N		8,694,060	8,694,060		8,694,060	8,694,060
Pseudo R^2		0.580	0.580		0.588	0.588

This table shows the results on how the effect of social connectedness varies with firm characteristics. The dependent variable is *%PF*, defined as the percentage of institutional AUM allocated to a stock. We report results on the heterogeneity across firm size. We interact *Log Social Connectedness* with dummy variables based on firm size terciles in columns 1 and 2. We report results on the heterogeneity across firms' analyst coverage. Analyst coverage is the number of one-year-ahead analyst estimates at the end of the quarter. We interact *Log Social Connectedness* with dummy variables based on firm analyst coverage terciles in columns 3 and 4. *Log Social Connectedness* is defined as *log(Social Connectedness)*, where *Social Connectedness* is defined as the number of Facebook links between a firm's headquarters' county and an institution's headquarters' county, scaled by the product of the populations in these two counties. We consider Firm, Institution×Industry, Distance Percentile, Same County, and Same State fixed effects, interacted with the respective firm characteristics dummies. *Same County* (*Same State*) is a dummy variable equal to one if the institution and the firm are located in the same county (same state) and zero otherwise. *Distance Percentile* indicators indicate the percentile of the distance between the firm and the institution based on all firm-institution pairs each quarter. Refer to the Appendix for detailed variable definitions. Industry classification is based on Fama-French 48 industries. Standard errors are double clustered by institution and firm, and *t*-statistics are reported in parentheses. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table 5
Identification using panel data

	(1)	(2)	(3)	(4)	(5)
Log Social Connectedness	0.180***	0.272***	0.261***	0.090***	0.045*
	(15.86)	(11.97)	(10.83)	(3.20)	(1.67)
Log Social Connectedness × Years After HQ Change					0.005***
					(3.17)
Firm × Quarter FE	YES	YES	YES	YES	YES
Institution × Quarter FE	YES	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	NO	NO
Institution × Firm FE	NO	NO	NO	YES	YES
Distance Percentile FE	NO	YES	YES	YES	YES
Same State FE	NO	NO	YES	YES	YES
Same County FE	NO	NO	YES	YES	YES
N	2.881e+08	2.881e+08	2.881e+08	2.881E+08	2.851E+08
Pseudo R^2	0.441	0.442	0.442	0.862	0.862

This table shows the results on how social connectedness affects institutional investors' portfolio decisions using a panel data. The sample represents institutional holdings from June 2007 to December 2016. We eliminate observations after a firm's second headquarters move in column 5. The dependent variable is *%PF*, which is defined as the percentage of AUM allocated to a stock. *Social Connectedness* is defined as the number of Facebook links between a firm's headquarters' county and an institution's headquarters' county, scaled by the product of the populations in these two counties. *Log Social Connectedness* is defined as *log(Social Connectedness)*. *Years After HQ Change* is the number of years after a firm's a headquarters move. Firms without a headquarters move are assigned a value of 10 years. We consider Firm × Quarter, Institution × Quarter, Institution × Industry, Firm × Institution, Distance Percentile, Same County, and Same State fixed effects. *Same County* (*Same State*) is a dummy variable equal to one if the institution and the firm are located in the same county (state) and zero otherwise. *Distance Percentile* are indicators that indicate the percentile of the distance between the firm and the institution based on all firm-institution pairs as of June 2016. Refer to the Appendix for detailed variable definitions. Industry classification is based on the Fama-French 48 industries. Standard errors are clustered by institution, firm, and quarter, and t-statistics are reported in parentheses. * p < .1; ** p < .05; *** p < .01.

unobservables that might be correlated with both social connectedness and investment probabilities. In these specifications, any variation in social connectedness within a firm-investor pair comes from headquarters moves of firms.[16] Overall, about 10% of firms in our sample changed headquarters locations during the sample period.

The within-firm-investor-pair variation in social connectedness continues to affect investment patterns: when a firm moves its headquarters from a location that is weakly connected to a particular investor to a location that is more strongly connected to that investor, the investor increases its investment in that firm. The estimated effect of social connectedness on investments in column 4 of Table 5 is statistically significant but smaller than the effect estimated in column 3. This is unsurprising, since investor awareness is unlikely to adjust immediately following a firm's headquarters move. In particular, investors with links to the original headquarters' county do not immediately "forget" about the firm the moment it moves to another county and would certainly not

16 While institutions also change their headquarters from time to time, our data for investor locations is based only on a single cross-section. As a result, we cannot use time-varying institutional investor locations to provide additional identification.

remove it from their portfolio immediately. Similarly, it would take some time before investors with links to the new headquarters' county hear about the firm through their social networks. Consistent with this interpretation, column 5 shows that changes in social connectedness between firms and investors due to firm relocations lead to greater changes in institutions' investments as the time since the relocation increases.

Overall, the findings in this section substantially reduce the scope for potential omitted variables to explain the observed relationships between social connectedness and investment behavior.

1.5 Fund manager characteristics, social connectedness, and fund investments

Our next analysis turns to a sample of actively managed U.S. domestic equity mutual funds.[17] Focusing on individual funds rather than institutions allows us to shed light on how different fund manager characteristics affect the relationship between social connectedness and managers' investment decisions.[18]

Table 6 reports a set of analyses corresponding to Regression (3). In the baseline specification in column 1, we confirm the previously established positive and significant relationship between *Log Social Connectedness* and $\%PF$, even after including fixed effects for various measures of geographic proximity and firm × fund style and fund × industry fixed effects (fund styles are based on the Lipper fund investment objective). Column 2 reports a placebo test using a sample of index funds. Since index fund managers do not actively select stocks, their social networks should not significantly affect their holdings. Consistent with this hypothesis, we find that there is no relationship between social connectedness and the holdings of index funds.

The remaining columns explore heterogeneity by fund manager characteristics. Column 3 shows that investments of funds managed by younger managers vary less with the social connectedness to firms' headquarters' locations. Column 4 analyzes whether the relationship between social connectedness and investment decisions differs between funds managed only by male managers and funds with at least one female manager. We find that both types of funds disproportionately invest in socially connected stocks, with somewhat larger

[17] The holdings data are from Thomson Reuters Mutual Fund Holdings Data. This sample differs from the institution sample, which also include entities such as banks, pension funds, hedge funds, and insurance companies.

[18] The fund managers' locations are collected from funds' N-SAR filings. We obtain data from Suzanne Chang. See Chang (2019) for a detailed data description. We further supplement her data with additional fund locations obtained from N-SAR filings. Fund manager characteristics are obtained from Morningstar and public records. See Chung (2018) for a detailed data description. We are grateful to Kiseo Chung for providing this data set. We require having a mutual fund manager's characteristics for a fund to be considered in our analysis. Our sample consists of 778 unique actively managed mutual funds. The median age of young management teams is 47 years, and that of old management teams is 56 years. Of these funds, 662 have only male managers while 156 have at least one female manager. Additionally, 316 funds are managed by a team with less than 50% managers holding MBA degrees, while 462 funds have more than 50% managers holding MBA degrees.

Table 6
Social connectedness and mutual fund investment, with fund manager characteristics

	(1) Active	(2) Index		(3) Active		(4) Active		(5) Active
Log Social Connectedness	0.099***	0.008						
	(4.07)	(0.13)						
Log Social Connectedness			× Young	0.075**	× Male only	0.090***	× MBA minority	0.120***
				(2.27)		(3.34)		(3.36)
			× Old	0.121***	× Female or both gender	0.111**	× MBA majority	0.086***
				(3.69)		(2.48)		(2.69)
Fund × Industry FE	YES	YES		YES		YES		YES
Firm × Style FE	YES	YES		YES × SPLIT		YES × SPLIT		YES × SPLIT
Distance Percentile FE	YES	YES		YES × SPLIT		YES × SPLIT		YES × SPLIT
Same County FE	YES	YES		YES × SPLIT		YES × SPLIT		YES × SPLIT
Same State FE	YES	YES		YES × SPLIT		YES × SPLIT		YES × SPLIT
F test (no heterogeneity)				1.05		0.17		0.52
N	2,155,060	529,070		2,155,060		2,155,060		2,155,060
Pseudo R^2	0.583	0.875		0.620		0.610		0.619

This table shows the results on how the effect of social connectedness on fund investments varies by mutual fund managers' characteristics. Our sample includes fund-firm pairs as of June 2016, including 778 actively managed U.S. domestic equity mutual funds, 191 index funds, and 2,770 unique firms. The dependent variable is *%PF*, defined as the percentage of fund AUM allocated to a stock. *Social Connectedness* is defined as the number of Facebook links between a firm's headquarters' county and a fund's headquarters' county, scaled by the product of the populations in these two counties. *Log Social Connectedness* is defined as *log(Social Connectedness)*. Column 1 shows the result from the baseline regression. Column 2 reports a placebo test based on the index fund sample. Column 3 reports the result on the heterogeneity across fund managers' age, where we split the funds into two groups by median age of the management team. In column 4, we split the funds into two groups by whether the fund is (co-)managed by at least one female manager. In column 5, we split the funds into two groups by whether more than 50% of the managers team has an MBA degree. We consider Firm, Fund×Industry, Firm×Style, Distance Percentile, Same County, and Same State fixed effects. *Same County* (*Same State*) is a dummy variable equal to one if the fund and the firm are located in the same county (same state) and zero otherwise. Distance Percentile fixed effects indicate the percentile of the distance between the firm and the fund based on all fund-firm pairs as of June 2016. Refer to the Appendix for detailed variable definitions. Industry classification is based on Fama-French 48 industries. Fund style is based on Lipper fund classification. In columns 3 to 5, we also interact fixed effects and controls with the indicators for the split (denoted as "SPLIT"). Standard errors are double clustered by fund and firm, and *t*-statistics are reported in parentheses. * p < .1; ** p < .05; *** p < .01.

effects for funds with female managers (though these differences are not statistically significant). Column 5 explores how fund managers' education affects their tendency to hold stocks headquartered in counties they are socially connected to. Specifically, we split funds based on whether the investment team consists primarily of managers with an MBA degree. We find that both types of funds are similarly inclined to hold socially connected stocks, suggesting that formal business education does not affect fund managers' reliance on social connectedness in their investment decisions.

2. Capital Market Implication for Firms

In the previous section, we established that institutional investors are more likely to invest in firms located in counties that the investors are socially connected to. We next show that this effect is large enough to generate better capital market outcomes for firms located in counties that are socially connected to regions with many large institutional investors—firms that we refer to as having high "social proximity to capital." We analyze three sets of capital market outcomes with quarterly panel data for the 2007–16 period. We first show that firms with higher social proximity to capital have more total institutional ownership. We then document positive effects of higher social proximity to capital on firm valuations and secondary market liquidity. We also show that the positive capital market effects of social proximity to capital are larger for smaller and more informationally opaque firms, precisely those firms for which we previously found the largest effects of social connectedness on investments.

2.1 Data and measurement

Our main explanatory variable in this section, the *Social Proximity to Capital* (SPC) of firms in county i at time t, is constructed as:

$$\textit{Social Proximity to Capital}_{i,t} = \sum_j AUM_{j,t} \times \textit{Social Connectedness}_{i,j}, \quad (4)$$

where $AUM_{j,t}$ is the total assets under management by institutions headquartered in county j in quarter t, and $\textit{Social Connectedness}_{i,j}$ is the social connectedness between counties i and j as defined in Equation (1). Increases in this measure mean that county i (and therefore any firm headquartered in that county) has closer social connections to counties that are home to institutional investors with high AUM.

Figure 3 shows a heat map of *Social Proximity to Capital* across U.S. counties; a data set with each county's social proximity to capital can be found on the authors' websites. Perhaps unsurprisingly, counties located on the East Coast, especially those in the Northeast, have the highest levels of *Social Proximity to Capital*, while counties in the middle of the country tend to have lower *Social Proximity to Capital.* However, consistent with our prior

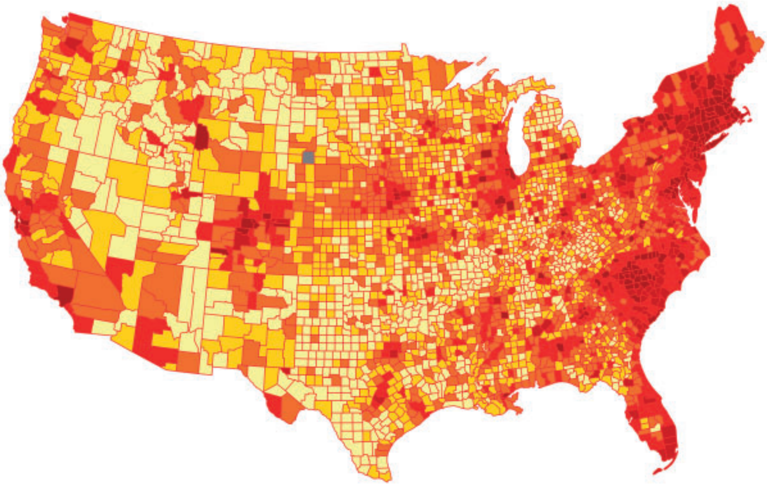

Figure 3
Heat map of social proximity to capital
This figure plots a heat map of *Social Proximity to Capital* across U.S. counties as of June 2016. *Social Proximity to Capital* of county j is defined as $\sum_i$ *County* $AUM_i \times$ *Social Connectedness*$_{i,j}$. Regions in darker colors represent higher levels of *Social Proximity to Capital*, and regions in lighter colors indicate lower levels of *Social Proximity to Capital*.

evidence that neighboring counties can have very different structures of social networks, we find that *Social Proximity to Capital* can also vary substantially between counties that are geographically close. This fact allows us to include state fixed effects in our regressions below and thereby only exploit within-state variation in *Social Proximity to Capital*.

Analogously, we construct a measure of a county's physical proximity to capital:

$$Physical\ Proximity\ to\ Capital_{i,t} = \sum_j AUM_{j,t}/(1+Distance_{i,j}), \tag{5}$$

where $Distance_{i,j}$ is the physical distance between counties i and j measured in miles. Consistent with the strong relationship between social connectedness and physical distance discussed above, we find that *Social Proximity to Capital* and *Physical Proximity to Capital* have a correlation of 0.86.

2.2 Social proximity to capital and firms' institutional ownership

Our first test explores whether institutions' overweighting of firms they are socially connected to has an aggregate effect on firms' total institutional ownership. To do this, we estimate the following regression:

$$\%TIO_{i,t} = \beta\,Log\ Social\ Proximity\ to\ Capital_{i,t-1} + \gamma X_{i,t-1} + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \tag{6}$$

where $\%TIO_{i,t}$ represents the total institutional ownership share of firm i in quarter t, and $X_{i,t-1}$ includes firm-level control variables that have been shown to affect a firm's institutional ownership share (see Baik, Kang, and Kim 2010; Green and Jame 2013) as well as numerous controls for county characteristics. As in the rest of this section, these controls are listed in the table notes and defined formally in the Appendix.

Our baseline specification, reported in column 1 of Table 7, also includes quarter, state, and industry fixed effects. We find that *Social Proximity to Capital* is significantly related to firms' institutional ownership shares: a 10% increase in *Social Proximity to Capital* is associated with a 20.4 basis points (bps) increase in the overall institutional ownership share, relative to a baseline mean of 60%. In column 2, we control for quarter $\times$ industry fixed effects, which ensures that our results are not driven by time-varying industry dynamics in institutional ownership. The point estimate of β is essentially unchanged. To further improve identification of the prior results, we include firm fixed effects in column 3. In this specification, within-firm variation in social proximity to capital is driven both by firms changing headquarters and by changes in the AUM of a given investor (that is the *Social Proximity to Capital* of a county with strong links to San Francisco increases when the AUM held by San Francisco–based investors increase). We find similar results, though the statistical significance declines somewhat.

We next investigate whether the relationship between *Social Proximity to Capital* and institutional investor share differs across firm characteristics. We documented above that social connectedness is particularly important for attracting institutional investments to small firms and firms with low analyst coverage. Consistent with those results, columns 4 and 5 of Table 7 show that *Social Proximity to Capital* has the largest effect on the institutional ownership shares of smaller firms. Quantitatively, a 10% increase in *Social Proximity to Capital* leads to a 21 bps increase in the institutional ownership share among small firms. This relationship is much smaller and statistically insignificant for midsize and large firms. Similarly, in columns 6 and 7, we find that the effect of *Social Proximity to Capital* on institutional investment share is most pronounced for firms with the lowest analyst coverage.

2.3 Social proximity to capital and firm valuation

We next investigate how firms' social proximity to capital affects their valuations. There are a number of channels through which social proximity to capital might raise a firm's valuation. The first channel is a direct implication from Merton (1987), who presents a model in which each investor knows only a subset of stocks. In equilibrium, those firms with limited investor recognition—and thus a smaller investor base—tend to have lower valuations. The intuition is that a wider investor base facilitates more risk sharing, which leads to higher valuations and a lower cost of capital. A similar argument is made by Hong, Kubik, and Stein (2008) in the context of investors' local bias. This theory finds

Table 7
Firms' social proximity to capital and institutional ownership

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Whole sample			Split by size		Split by analyst coverage	
Log Social Proximity to Capital	2.039**	1.996**	2.732*				
	(2.36)	(2.33)	(1.94)				
Low × Log Social Proximity to Capital				2.137**	2.074**	2.006**	1.925**
				(2.35)	(2.30)	(2.19)	(2.12)
Mid × Log Social Proximity to Capital				0.235	0.099	0.730	0.688
				(0.23)	(0.10)	(0.80)	(0.76)
High × Log Social Proximity to Capital				0.032	0.033	0.675	0.813
				(0.03)	(0.03)	(0.69)	(0.85)
Firm controls	YES	YES	YES	YES	YES	YES	YES
County controls	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES	YES × SPLIT	NO	YES × SPLIT	NO
Industry FE	YES	NO	NO	YES × SPLIT	NO	YES × SPLIT	NO
State FE	YES	YES	NO	YES × SPLIT	YES × SPLIT	YES × SPLIT	YES × SPLIT
Quarter × Industry FE	NO	YES	NO	NO	YES × SPLIT	NO	YES × SPLIT
Firm FE	NO	NO	YES	NO	NO	NO	NO
F test (low=high)				4.838**	4.605**	1.815	1.291
N	99,555	99,555	99,555	99,555	99,555	99,555	99,555
R^2	0.358	0.372	0.833	0.431	0.456	0.438	0.462

This table presents the panel regression results on how firms' social proximity to institutional capital affects their institutional ownership. Our sample consists of firms' quarterly institutional holding data from 2007 to 2016. The dependent variable is total institutional ownership (*%TIO*), which is the percentage of shares outstanding held by institutional investors. The main independent variable is *Log Social Proximity to Capital*, where *Social Proximity to Capital* is defined as $\sum$ *County AUM* × *Social Connectedness*. *County AUM* is measured in millions of U.S. dollars, and the summation is taken across all U.S. counties. Our firm controls include *Log Assets*, *Log M/B*, *Volatility*, *Momentum*, *Turnover*, *Price*, *R&D*, *Yield*, *S&P*, *Age*, *Advertising*, and *Exchange*. Our county controls include *Log Physical Proximity to Capital*, *Log Population*, *High School*, *College*, *Log Agglomeration*, *Income*, and *Employment*. Refer to the Appendix for detailed variable definitions. The regressions also include Quarter, State, Industry, Quarter × Industry, and Firm fixed effects. Industry fixed effects are based on the Fama-French 48 industry classification. Columns 4 and 5 exhibit the results on the heterogeneity across firm size terciles. Firm size terciles are based on firms' market capitalization at the end of the prior quarter. Columns 6 and 7 exhibit the results on the heterogeneity across analyst coverage terciles, where we rank firms into terciles based on the number of analysts covering a firm at the end of the prior quarter end. Standard errors are double clustered by quarter and firm, and t-statistics are reported in parentheses below each estimate. * $p < .1$; ** $p < .05$; *** $p < .01$.

strong support in subsequent empirical work (e.g., Lehavy and Sloan 2008; Fang and Peress 2009). We therefore hypothesize that as a result of higher investor recognition, firms with larger social proximity to capital—firms that we just showed to have a larger institutional investor base—might have higher valuations.

A second mechanism through which social proximity to capital might raise valuations is through investor disagreement. In this story, investors are more likely to consider investing in firms they are aware of. However, investors disagree about the firm's valuation. In the presence of short-sale constraints, prices disproportionately reflect the views of the most optimistic investors (Miller 1977), and when the investors' beliefs oscillate with the arrival of new information, the resulting overvaluation can persist (Scheinkman and Xiong 2003). Hence, a larger investor base has the potential to increase firm valuations by creating more scope for belief heterogeneity.

To explore how social proximity to capital affects firm valuations, we estimate:

$$Log\ Valuation_{i,t} = \beta Log\ Social\ Proximity\ to\ Capital_{i,t-1} + \gamma X_{i,t-1} + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \tag{7}$$

where $Valuation_{i,t}$ represents the market valuation of firm i in quarter t. We consider the market-to-book ratio as our primary proxy for firm valuation (in Internet Appendix Table IA.7, we show that all results are robust to using Tobin's Q as the measure of firm valuation). The dependent variable is in log form following Green and Jame (2013). $X_{i,t-1}$ includes control variables that have been shown to affect firm valuation, as well as county-level demographic and economic information.

The results from Regression (7) are reported in Table 8. In columns 1 and 2, we find a strong positive relation between a firm's social proximity to capital and its market-to-book ratio: a 10% increase in *Social Proximity to Capital* is associated with a statistically significant 1.1% increase in the market-to-book ratio. As reported in column 3, including firm fixed effects does not change this point estimate in a significant way. The rest of Table 8 highlights that the effects of *Social Proximity to Capital* on market-to-book ratios are strongest among small and midsize firms and among firms with limited analyst coverage.

2.4 Social proximity to capital and secondary market liquidity

We next examine the impact of social proximity to capital on firms' secondary market liquidity. Since institutional investors are important providers of liquidity (e.g., Rubin 2007; Blume and Keim 2012), we postulate that firms with high social proximity to institutional capital will have higher liquidity. This prediction builds on prior work that shows that stocks of firms with higher investor recognition (e.g., due to more fluent names) and firms with more competition among liquidity providers are more liquid

Table 8
Social proximity to capital and firm value

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Whole sample			Split by size		Split by analyst coverage	
Log Social Proximity to Capital	0.110***	0.110***	0.110***				
	(4.15)	(4.12)	(2.80)				
Low × Log Social Proximity to Capital				0.076***	0.079***	0.094***	0.094***
				(3.16)	(3.28)	(3.52)	(3.58)
Mid × Log Social Proximity to Capital				0.048*	0.050**	0.097***	0.097***
				(2.02)	(2.15)	(3.58)	(3.63)
High × Log Social Proximity to Capital				0.057*	0.055*	0.050*	0.043
				(2.00)	(1.96)	(1.70)	(1.48)
Firm controls	YES	YES	YES	YES	YES	YES	YES
County controls	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES	YES × SPLIT	NO	YES × SPLIT	NO
Industry FE	YES	NO	NO	YES × SPLIT	NO	YES × SPLIT	NO
State FE	YES	YES	NO	YES × SPLIT	YES × SPLIT	YES × SPLIT	YES × SPLIT
Quarter × Industry FE	NO	YES	NO	NO	YES × SPLIT	NO	YES × SPLIT
Firm FE	NO	NO	YES	NO	NO	NO	NO
F test (low=high)				0.428	0.682	2.077	2.883*
N	96,762	96,762	96,762	96,762	96,762	96,762	96,762
R^2	0.294	0.315	0.800	0.496	0.526	0.394	0.431

This table presents the panel regression results on how firms' social proximity to institutional capital affects their valuation. Our sample includes quarterly observations of firm valuation from 2007 to 2016. The dependent variable is *Log M/B*. The main independent variable is *Log Social Proximity to Capital*, where *Social Proximity to Capital* is defined as $\sum$ *County AUM* × *Social Connectedness*, where *County AUM* is measured in millions of U.S. dollars, and the summation is taken across all U.S. counties. Additional firm controls include *Log Assets*, *Profitability*, *Sale Growth*, *Asset Turnover*, *R&D*, *Advertising*, *Leverage*, *Payout*, *S&P*, *Age*, and *Exchange* dummies. Additional county controls include *Log Physical Proximity to Capital*, *High School*, *College*, *Log Agglomeration*, *Log Population*, *Income*, and *Employment*. Refer to the Appendix for detailed variable definitions. The regressions also include Quarter, State, Industry, Quarter × Industry, and Firm fixed effects. Industry fixed effects are based on the Fama-French 48 industry classification. Columns 4 and 5 exhibit the results on heterogeneity across firm size terciles. Firm size terciles are based on firms' market capitalization at the end of the prior quarter. Columns 6 and 7 exhibit the results on heterogeneity across analyst coverage terciles, where we rank firms into terciles based on the number of analysts covering a firm at the end of the prior quarter. Standard errors are double clustered by quarter and firm, and t-statistics are reported in parentheses below each estimate. * $p < .1$; ** $p < .05$; *** $p < .01$.

(see Green and Jame 2013; Liu and Wang 2016). We conduct the following regression analysis:

$$Log\ Liquidity_{i,t} = \beta Log\ Social\ Proximity\ to\ Capital_{i,t-1} + \gamma X_{i,t-1} + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \tag{8}$$

where $Liquidity_{i,t}$ is the effective percentage spread; in Internet Appendix Table IA.8, we confirm that all results are robust to analyzing the effects on the Amihud (2002) illiquidity measure. The dependent variable is in log form following Green and Jame (2013). As before, we include control variables that have been shown to affect firm liquidity, in addition to the same fixed effects as in the previous specifications.

The regression results are reported in Table 9. Column 1 shows that a 10% increase in *Social Proximity to Capital*—equivalent to a 0.08 standard deviation increase in that number—is associated with a 0.94% reduction in the effective spread. To put these magnitudes in perspective, we compare them to the effect of profitability on liquidity. Consistent with prior literature, we find that a one-standard-deviation increase in profitability is associated with a 5.8% decrease in the effective spread, while a one-standard-deviation increase in *Log Social Proximity to Capital* is associated with a 10.6% decrease in the effective spread. The coefficient remains similar when we include quarter × industry fixed effects in column 2 and firm fixed effects in column 3. Columns 4 and 5 of Table 9 show that the relationship between *Social Proximity to Capital* and effective spreads is concentrated among small firms. Columns 6 and 7 highlight that the effective spread of firms with high analyst coverage is not affected by social proximity to capital, while these effects are highly significant for firms with low and intermediate levels of analyst coverage.

In summary, we find that institutional attention to firms with high social proximity to capital may lead to higher liquidity, with the strongest effects for lesser-known firms. While our study cannot exploit quasi-random variation in *Social Proximity to Capital* across counties to obtain causal estimates, our analyses are able to control for a large number of observables at the firm level that have been shown to influence our liquidity measures. We also include county-level controls that might be correlated with both *Social Proximity to Capital* and the characteristics of local firms. More generally, we are not aware of any omitted variables that could jointly explain our findings, and in particular the heterogeneity of the effects across firm characteristics. For example, if firms in counties that were socially more proximate to capital were of higher quality on average, this would not explain the observed disproportionate investment in those firms by institutional investors in socially close counties relative to investments by institutional investors in socially distant counties. It is also unclear why only small firms in counties with high *Social Proximity to Capital* would have a higher fundamental quality or why the fundamental quality of the same firm would change when it moves headquarters from a county

Table 9
Social proximity to capital and stock liquidity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Whole sample		Split by size		Split by analyst coverage	
Log Social Proximity to Capital	−0.094***	−0.092***	−0.068**				
	(−4.50)	(−4.44)	(−2.49)				
Low × Log Social Proximity to Capital				−0.081***	−0.080***	−0.093***	−0.089***
				(−3.69)	(−3.66)	(−4.24)	(−4.04)
Mid × Log Social Proximity to Capital				−0.056***	−0.055***	−0.075***	−0.073***
				(−3.14)	(−3.08)	(−4.11)	(−4.03)
High × Log Social Proximity to Capital				0.013	0.012	−0.007	−0.007
				(0.73)	(0.66)	(−0.41)	(−0.39)
Firm controls	YES	YES	YES	YES	YES	YES	YES
County controls	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES	YES × SPLIT	NO	YES × SPLIT	NO
Industry FE	YES	NO	NO	YES × SPLIT	NO	YES × SPLIT	NO
State FE	YES	YES	NO	YES × SPLIT	YES × SPLIT	YES × SPLIT	YES × SPLIT
Quarter × Industry FE	NO	YES	NO	NO	YES × SPLIT	NO	YES × SPLIT
Firm FE	NO	NO	YES	NO	NO	NO	NO
F test (low=high)				18.075***	17.487***	13.029***	11.901***
N	100,502	100,502	100,502	100,502	100,502	100,502	100,502
R^2	0.772	0.781	0.915	0.832	0.843	0.825	0.837

We study how firms' social proximity to institutional capital affects their stock liquidity using panel regressions. Our sample covers quarterly stock liquidity variables from 2007 to 2016. The dependent variable is *Log Effective Spread*, where *Effective Spread* is the quarterly average of daily percentage effective spread calculated by dividing effective bid-ask spreads by midpoint prices. The main independent variable is *Log Social Proximity to Capital*, where *Social Proximity to Capital* is defined as $\sum$ *County AUM* × *Social Connectedness*. *County AUM* is measured in millions of U.S. dollars, and the summation is taken across all U.S. counties. Additional firm controls include *Log Assets*, *Profitability*, *Log M/B*, *Log Volatility*, *Momentum*, *Price*, *S&P*, *Age*, *Advertising*, and *Exchange* dummies. Additional county controls include *Log Physical Proximity to Capital*, *High School*, *College*, *Log Agglomeration*, *Log Population*, *Income*, and *Employment*. Refer to the Appendix for detailed variable definitions. The regressions also include Quarter, State, Industry, Quarter × Industry, and Firm fixed effects. Industry fixed effects are based on the Fama-French 48 industry classification. Columns 4 and 5 exhibit the results on the heterogeneity across firm size terciles. Firm size terciles are based on firms' market capitalization at the end of the previous quarter. Columns 6 and 7 exhibit the results on the heterogeneity across analyst coverage terciles, where we rank firms into terciles based on the number of analysts covering the firm at the end of the prior quarter. Standard errors are double clustered by quarter and firm, and *t*-statistics are reported in parentheses below each estimate. * $p < .1$; ** $p < .05$; *** $p < .01$.

with low social proximity to capital to one with high social proximity to capital. Nevertheless, we next provide additional evidence for our proposed explanation.

2.4.1 Hurricane Sandy and market liquidity. One concern with the prior specifications is that, despite our rich set of controls, there might be omitted variables that correlate with firms' social proximity to capital as well as their liquidity (and that do so more for lesser-known firms). For example, one might argue that places with high social proximity to capital have more well known firms in general, and thus will have higher liquidity provision from all institutional investors, independent of where those investors are located. To provide further evidence against such alternative interpretations, we analyze the response of firms' secondary market liquidity to a temporary shock to investors in socially connected counties.

Specifically, we explore the temporary shock to East Coast–based investors during Hurricane Sandy in late October 2012, which caused damages of nearly US$ 70 billion. Hurricane Sandy presents a unique opportunity to explore the causal effects of social proximity to capital, due to the concentration of capital in the affected areas and the fact that those investors' ability to participate in financial markets was substantially reduced in the aftermath of the hurricane. In particular, in the weeks after Sandy, many employees were unable to come physically into their offices due to the disruption in roads and public transportation. As a result, the liquidity provision by the institutions in the affected area was substantially impaired. Consistent with this interpretation, the *Wall Street Journal* quoted a trader saying, "'The market isn't officially closed, but many of the venues that supply liquidity are closed.... if people thought volume was thin recently, Monday could be the Wild West for low liquidity'" (Russolillo 2012). If our interpretation of the correlation between social proximity to capital and liquidity were correct, we would thus expect the liquidity of firms with high connectedness to East Coast–based investors to fall disproportionately during Hurricane Sandy, a prediction that we next test empirically.

In our baseline analysis, we define the area affected by Sandy to be the mid-Atlantic states (New York, New Jersey, Connecticut, District of Columbia, Pennsylvania, Delaware, Maryland, Virginia, and West Virginia), though our results are robust to broader or narrower definitions. In our regressions, we exclude those firms that are geographically close to the affected area (all firms located in the eastern United States) to avoid any spurious results on liquidity driven by uncertainty related to firms' fundamentals (e.g., Rehse et al. 2019). Our empirical specification is as follows:

$$Log\ Spread_{i,t} = \beta Affected\ Ratio_i \times I(Sandy)_t + \gamma X_{i,t} + \psi_{t,ind_i} + \xi_i + \epsilon_{i,t}. \quad (9)$$

$I(Sandy)_t$ is an indicator variable that equals to one during the Sandy period, defined here as October 22–November 7, 2012, the period when travel in the Tri-State region was substantially affected by Sandy. The key explanatory variable, *Affected Ratio*$_i$, is defined as the ratio of county i's socially proximate capital in the affected area (i.e., the mid-Atlantic states) to county i's overall social proximity to capital, measured in the quarter before Sandy:

$$\mathit{Affected\ Ratio}_i = \frac{\sum_{k \in Mid-Atlantic} AUM_k \times \mathit{Social\ Connectedness}_{i,k}}{\sum_j AUM_j \times \mathit{Social\ Connectedness}_{i,j}}. \quad (10)$$

In other words, *Affected Ratio* captures the cross-sectional exposure of firms to institutional capital in the areas affected by Hurricane Sandy. In our baseline specification, we also include firm fixed effects, which absorb any effect of *Affected Ratio* on firms' average liquidity, and day $\times$ industry fixed effects, which absorb any common time variation in liquidity among firms in the same industry. $X_{i,t}$ includes additional firm-level and county-level control variables from Table 9, each interacted with $I(Sandy)_t$.

We report these results in Table 10. The sample period covers January 2012 to July 2013. The coefficient of interest is β, which captures whether the cross-sectional variation in exposure to institutional capital affected by Sandy correlates with differential changes in the effective spread during the Sandy period. The β-coefficient in column 1 is positive and significant, highlighting that the spreads of firms with higher *Affected Ratio* widened more during the Sandy period compared to the spreads of other firms. The economic magnitude of the effect is significant: firms with a one-standard-deviation higher *Affected Ratio* experienced a 2.53% (0.253$\times$0.10) additional increase in their effective spreads.

In column 2, we consider the possibility that our results may be driven by firms' economic connections to mid-Atlantic states (remember that we already drop all firms located in the eastern United States). To explore whether such an exposure can explain our results, we exclude any firm that lists a mid-Atlantic state in the top three states in the firm's 10-K filing (see Section 1.2). As shown in column 2, we still find a positive and significant coefficient on the interaction of the Sandy indicator and *Affected Ratio*, despite losing almost half of our observations.

The effects on liquidity of social proximity to capital in the areas affected during Hurricane Sandy also vary with firm characteristics. In particular, in Section 2.4, we documented that the average effect of social proximity to capital on firm liquidity is largest for small firms and those with low analyst coverage. In columns 3 and 4 of Table 10, we show that it is generally the liquidity of those same small firms with social proximity to capital in the affected areas that falls the most during Hurricane Sandy.

In summary, shocks to the liquidity provision by institutional investors lead to the largest increases in the effective spreads of firms that are socially connected to the region affected by the shock. This effect is largest for lesser-known firms. These findings show that our cross-sectional liquidity results are not driven by firm or county characteristics that affect liquidity provision by all investors.

Table 10
The effect of social proximity to capital during Hurricane Sandy

	(1)	(2)	(3)	(4)
	Whole sample		Split by size	Split by analyst coverage
I(Sandy) × Affected Capital Ratio	0.253***	0.380***		
	(3.42)	(3.72)		
Low × I(Sandy) × Affected Capital Ratio			0.428**	0.550***
			(2.32)	(3.17)
Med × I(Sandy) × Affected Capital Ratio			0.118	0.028
			(0.77)	(0.25)
High × I(Sandy) × Affected Capital Ratio			0.152	0.249***
			(1.35)	(2.96)
Control × Sandy FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
State FE	YES	YES	YES × SPLIT	YES × SPLIT
Day × Industry FE	YES	YES	YES × SPLIT	YES × SPLIT
Sample Exclusion	EAST	ESAT & EXPOSED	EAST	EAST
F test (low=high)			1.090	1.970
N	545,295	295,384	545,295	545,295
R^2	0.766	0.782	0.785	0.784

We analyze cross-sectional differences in the impact of Hurricane Sandy on the liquidity of firms with various levels of social proximity to institutional capital in the mid-Atlantic area. The sample ranges from January 2012 to July 2013. The dependent variable is daily *Log Effective Spread*. The key variable of interest is the interaction between *I(Sandy)* and *Affected Capital Ratio*. *I(Sandy)* is an indicator variable equal to one during the affected period defined as October 22, 2012, to November 7, 2012. *Affected Capital Ratio* is defined as the *Social Proximity to the Mid-Atlantic Capital*, divided by *Social Proximity to Capital* as of the third quarter of 2012. We include the control variables from Table 9 and the *Affected Ownership Ratio*, each interacted with *I(Sandy)*. Refer to the Appendix for detailed variable definitions. We also control for Firm, State, and Day × Industry fixed effects. We exclude firms in eastern states from our sample. Additionally, firms with a mid-Atlantic state as one of their top three mentioned states in 10-K are excluded in column 2. Column 3 shows heterogeneity of the result across firm size terciles, and column 4 shows it across analyst coverage terciles. We report t-statistics based on robust standard errors clustered by firm and week in the parentheses below. * $p < .1$; ** $p < .05$; ***$p < .01$.

3. Implications for Institutional Investors

We previously showed that institutional investors tend to overweight firms located in counties to which the investors are socially connected. We now examine the implications of this behavior for investor performance. This analysis will shed light on the mechanisms behind the observed investor behavior. Specifically, if the overweighting of connected firms is driven by an informational advantage that investors obtain through their social networks, one would expect that, all else equal, investors that hold more socially connected stocks should outperform other investors that hold fewer such stocks. Similarly, one might also expect that the same investor would be able to obtain higher returns on stocks they hold from socially connected counties than on stocks they hold from counties they are not connected to. If overweighting is instead driven by investor familiarity with firms, investors with larger holdings of socially connected stocks should not expect to outperform and may even underperform.[19]

[19] Underperformance may be due to: (i) institutional investors substituting a possible value-creating stock-picking ability with relying on the non-value-creating activity of buying socially connected stocks, and (ii)

Table 11
Portfolio social connectedness and performance

	Return			σ(Return)			Sharpe ratio		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Excess	CAPM	FF5	Excess	CAPM	FF5	Excess	CAPM	FF5
Low	0.044	0.006	0.007**	1.330***	0.645***	0.533***	0.061***	0.011*	0.026***
	(1.56)	(1.39)	(2.47)	(13.91)	(18.73)	(19.36)	(4.23)	(1.81)	(7.33)
2	0.040	0.003	0.004*	1.224***	0.451***	0.367***	0.064***	0.010*	0.021***
	(1.48)	(1.20)	(1.83)	(12.48)	(16.79)	(17.64)	(4.20)	(1.82)	(3.96)
3	0.040	0.004*	0.005**	1.197***	0.393***	0.315***	0.065***	0.014***	0.025***
	(1.51)	(1.78)	(2.54)	(12.10)	(15.77)	(16.05)	(4.20)	(2.92)	(4.44)
4	0.039	0.003	0.003*	1.180***	0.336***	0.267***	0.066***	0.013**	0.028***
	(1.46)	(1.19)	(1.70)	(11.48)	(14.50)	(15.22)	(4.15)	(2.34)	(4.14)
5	0.039	0.003	0.003*	1.178***	0.331***	0.263***	0.066***	0.016***	0.025***
	(1.48)	(1.62)	(1.80)	(11.50)	(14.95)	(15.40)	(4.20)	(2.91)	(3.70)
6	0.038	0.003	0.002	1.134***	0.318***	0.258***	0.067***	0.015**	0.022***
	(1.47)	(1.52)	(1.17)	(11.80)	(15.68)	(16.07)	(4.29)	(2.53)	(3.38)
7	0.038	0.003∗	0.003	1.145***	0.342***	0.274***	0.066***	0.014**	0.017**
	(1.49)	(1.88)	(1.57)	(11.52)	(14.74)	(15.15)	(4.27)	(2.19)	(2.54)
8	0.039	0.003	0.003*	1.153***	0.380***	0.315***	0.066***	0.013**	0.016***
	(1.48)	(1.65)	(1.65)	(12.06)	(16.71)	(17.42)	(4.33)	(2.31)	(2.82)
9	0.040	0.004**	0.004**	1.171***	0.416***	0.346***	0.066***	0.011**	0.015***
	(1.52)	(2.24)	(2.46)	(12.20)	(15.76)	(16.29)	(4.35)	(2.03)	(2.77)
High	0.042	0.005*	0.006***	1.287***	0.593***	0.507***	0.061***	0.011**	0.017***
	(1.54)	(1.75)	(2.69)	(13.16)	(15.99)	(16.30)	(4.43)	(2.07)	(3.76)
High-Low	−0.003	−0.001	−0.001	−0.042	−0.052	−0.026	−0.001	0.000	−0.009
	(−0.75)	(−0.40)	(−0.45)	(−1.34)	(−1.45)	(−0.82)	(−0.22)	(0.03)	(−1.67)

This table reports daily portfolio returns, volatilities, and Sharpe ratios of institutional investors with different propensities to hold socially connected stocks. The propensity to hold socially connected stocks for each institutional investor is estimated using the following equation:

$$\%PF_{i,j} = \exp[\beta_j \, Log\ Social\ Connectedness_{i,j} + \beta_2 Same\ County_{i,j} + \beta_3 Same\ State_{i,j} +$$
$$Firm\ FE + Institution \times Industry\ FE + Distance\ Percentile\ FE] \cdot \epsilon_{i,j}.$$

We sort institutional investors into deciles based on their propensity to hold socially connected stocks (β_j). Portfolios are rebalanced at the end of each quarter using institutions' quarter-end holdings. In the first three columns, we report average daily portfolio returns (in %) of the institutions in each decile, where the returns are excess returns over the risk-free rate, CAPM, and Fama-French five-factor adjusted returns. Columns 4–6 report the portfolio returns' standard deviation of excess returns or residual returns. To calculate the standard deviation, we first calculate return volatility for each institution in each quarter using daily returns. We then report the average standard deviation in a given decile. Columns 7–9 report portfolio Sharpe ratios. For each institution-quarter, we compute a Sharpe ratio, defined as the average portfolio return divided by return standard deviation. We report the average Sharpe ratio for each decile. For average portfolio returns, we report t-statistics based on Newey and West (1994) standard errors. For standard deviations and Sharpe ratios, we compute t-statistics based on quarter and institution clustered standard errors. t-statistics are reported in parentheses. * $p < .1$; ** $p < .05$; *** $p < .01$.

We examine measures of investor performance both across and within institutions. In the across-institution tests, we sort investors by their propensity to hold socially connected stocks and compare the overall performance across investors with different propensities. In the within-institution tests, we compare the performance of the connected and nonconnected holdings of the same institution.

3.1 Across-institution performance comparison

For the across-institution tests, we first estimate each institution's propensity to hold connected stocks in June 2016 using the following cross-sectional regression:

$$\%PF_{i,j} = \exp[\beta_j \, Log\ Social\ Connectedness_{i,j} + \gamma X_{i,j} + \psi_i + \xi_{j \times ind(i)}] \cdot \epsilon_{i,j}. \quad (11)$$

This specification corresponds to that of column 5 in Table 2, except that we allow the propensity to hold socially connected stocks to vary across institutions, giving us investor-specific β_j coefficients. We next sort institutions into deciles of β_j. For each β_j-decile portfolio and each quarter, we then compute the equally weighted averages of a number of performance measures across all institutions in the decile.

The results are presented in Table 11. Columns 1–3 show the results comparing average excess returns, CAPM-adjusted returns, and Fama and French (2015) five-factor-model-adjusted returns, respectively.[20] We do not find evidence that investors that exhibit a greater propensity to overweight stocks from areas they are socially connected to outperform investors with a lower propensity.

We also investigate whether institutional investors that invest more in socially connected counties bear more risk due to underdiversification. Columns 4–6 report the average volatility of daily returns for institutions in each decile portfolio. We do not find that the portfolios of institutions with a higher propensity to hold socially connected stocks have higher volatility. Finally, columns 7–9 report institutions' Sharpe ratios. As before, we do not find any significant differences between institutions in the top decile and institutions in the bottom decile of β_j.[21] Overall, there is no evidence that the propensity to hold socially connected stocks leads to a differential performance among institutional investors.

---

the overweighting of connected stocks resulting in portfolio underdiversification and therefore a deviation from the efficient portfolio.

[20] We compute institution i's daily holding period return as $Return_{i,t} = \sum_j w_{i,j,t-1} Return_{j,t}$, where $w_{i,j,t-1}$ is i's holding of stock j at end of the prior quarter and $Return_{j,t}$ is stock j's daily return. We then equal-weight all institutions' returns within a β_j-decile to obtain returns for that portfolio.

[21] Internet Appendix Table IA.10 shows that while there is some heterogeneity in the performance spread (that is, between portfolios of different propensity to hold socially connected stocks) across Bushee (2001) types, these differences are not economically meaningful.

3.2 Within-institution performance comparison

Although we do not find any significant performance difference between institutions that have a high propensity to invest in connected firms and those with a low propensity, it is possible that these institutions are also very different on other dimensions, which might confound our across-institution analysis. Therefore, we conduct a second test to examine whether, for the same institution, holdings from locations with a high connectedness to the investor perform better than holdings from locations with a lower connectedness. This test is motivated by Coval and Moskowitz (2001), who show that investors are better at picking stocks among firms that are geographically close by.

More specifically, for each county i, we sort all other counties j that have a firm headquarters into terciles by their connectedness to i. Next, we focus on institutions in county i and classify their stock holdings into "low," "mid," and "high" connectedness portfolios based on the headquarters' location of each stock (for example, we group all stocks headquartered in counties in the lowest tercile of connectedness to i). For each institution, we also construct an institution-neutral portfolio by taking a long position in the "high" social connectedness portfolio and a short position in the "low" social connectedness portfolio. In Table 12, we report average daily excess returns for each institution's subportfolios, as well as the daily risk-adjusted returns using the CAPM and Fama and French (2015) five-factor models. We find no significant return difference between the "low" and "high" connectedness portfolios, suggesting that an institution's connected holdings do not significantly outperform its nonconnected holdings.

We next compare the returns of high social connectedness stocks held by institutions with the returns of stocks in high social connectedness counties that the investors do not hold. This comparison helps us identify whether institutions are successful in avoiding low-quality stocks from counties they are socially connected to. Table 12 also shows that "high" connectedness portfolios held by institutions do not outperform stocks in high connectedness counties that are not held by the institution.[22]

In summary, we find that institutions do not systematically benefit from investing in socially connected stocks. Unlike the local bias literature (e.g., Coval and Moskowitz 2001), which shows that mutual funds may have an information advantage for local stocks (a pattern we replicate in Internet Appendix Table IA.11), we find no evidence that institutional investors have

[22] We also explore whether there exist performance differences between high-connectedness and low-connectedness portfolios among those institutions that are most prone to investing in stocks they are connected to (i.e., institutions in the top β_j-decile, institutions with small AUM, and dedicated institutions). In Internet Appendix Table IA.12, we report the average performance difference between high-connectedness and low-connectedness sub-portfolios for each of these institution types, as well as between connected stocks held by institutions and connected stocks not held by institutions. Columns 1–3 show that there is no significant difference in excess returns, CAPM-adjusted returns, or Fama and French (2015) five-factor-model-adjusted returns, respectively. Columns 4–6 show that, with the exception of dedicated institutions, even among institutions heavily invested in socially connected stocks, there is no consistent evidence that they can avoid selecting poorly performing socially connected stocks.

Table 12
Performance of socially connected holdings

Social connectedness	(1) Excess	(2) CAPM	(3) FF5
	Stocks held by institutions		
Low (held)	0.040*	0.005	0.002
	(1.68)	(1.22)	(0.61)
High (held)	0.039*	0.002	0.003
	(1.69)	(0.90)	(1.64)
High (held) - Low (held)	−0.001	−0.002	0.001
	(−0.27)	(−0.57)	(0.32)
	Stocks not held by institutions		
High (not held)	0.038	−0.000	0.003
	(1.57)	(−0.00)	(1.14)
High (held) - High (not held)	0.001	0.002	0.000
	(0.33)	(1.33)	(0.33)
N	2,456	2,456	2,456

This table reports the daily returns (in %) of socially connected holdings within institutions' portfolios from 2007 to 2016. To assign holdings into different portions of social connectedness for an institution, we use all the counties that have at least one institution (or firm) located in that county and construct institution-firm county pairs. For each institution county, we first sort all firm counties into terciles based on social connectedness between counties and then assign the firm counties into low, mid, and high connectedness counties. Based on firms' headquarters' counties, institutional holdings are assigned into three connectedness groups, and we report the average daily returns of the three groups. We report the excess returns, CAPM, and Fama-French five-factor alpha of the high and low social connectedness portfolios and the return difference in these two portfolios. The portfolios are rebalanced at the end of each quarter using the value of institutional holdings. We also report the value-weighted returns for stocks with high social connectedness to institutions that are not part of the institutional holdings and the return difference between the high social connectedness stocks held and not held by institutions. Newey and West (1994) adjusted t-statistics are reported in parentheses. * p < .1; ** p < .05; *** p < .01.

an information edge through social connections as measured by Facebook friendships. As a result, our results are most consistent with a story in which institutional investors' investments in socially connected firms are primarily driven by awareness of these firms rather than by superior information.[23]

4. Conclusion

A growing literature explores a variety of explanations for geographic disparities of economic outcomes across the United States. We contribute to this literature by investigating how the geographic structure of social networks shapes the allocation of capital to firms and thereby contributes to differences in firm outcomes. We find that, all else equal, institutional investors invest more in firms located in regions to which they have stronger social ties. As a result, firms in regions that are socially proximate to institutional capital have

[23] Importantly, though, our evidence does not mean that investors are unable to obtain information advantages through social networks more generally (such an effect was documented in Cohen, Frazzini, and Malloy 2008; Hong, Kubik, and Stein 2005; Hong and Xu 2019; Pool, Stoffman, and Yonker 2015). Rather, our findings suggest that the social network as characterized by Facebook friendship links represents a broader type of network that is intrinsically different from networks based on factors such as shared neighborhood or education institutions.

higher liquidity and higher valuations. We thus conclude that differences in the social proximity to capital can be an important channel through which regional characteristics affect economic opportunities for firms. An interesting direction for future research is to explore the effects of real outcomes for firms, such as their investment and research and development activities.

In addition, it is likely that the structure of regions' social networks can have other effects on regional outcomes beyond the effects on firms' access to capital that are the focus of the present paper (see Chetty et al. 2021). For example, social connections can facilitate trade flows between regions (as suggested by Bailey et al. 2021). Quantifying the extent to which regional differences in economic and social outcomes are explained by the structure of social networks, along with providing evidence for the various mechanisms, is an exciting avenue for future research, and we hope that the public availability of the Social Connectedness Index will encourage other researchers to join us in pursuing these questions.

Appendix

Table A.1
Variable List

Variable	Definition
A. Institution (mutual fund)-firm pairwise variables	
Social Connectedness Index (SCI)	Number of Facebook friends linked between two counties.
Social Connectedness (SC)	SCI divided by the product of two counties' population. We scale up this variable by a factor of 10^{12}. Log Social Connectedness is *log(SC).*
Distance	Distance measures the physical distance between two counties in miles. Log Distance is defined as *log(1+Distance).*
%PF	(Shares held$\times$Price/Institution AUM)$\times$100.
Citation Share	Fraction of citation of an institution state in a firm's 10-K. *I(Cited)* equals 1 if Citation Share is greater than 0.
Revenue Share	Fraction of revenue derived from an institution state. *I(Revenue State)* equals 1 if Revenue Share is greater than 0.
|County Difference|	Absolute county-pair characteristics differences. We consider the differences along the following dimensions: log county population, population density, median age, income per capita, share of high school graduates, share of college graduates, employment rate, share immigrants, share unmarried, industry similarity, and Democratic vote share. Density is a 1–6 code obtained from the USDA. Voting data are obtained from townhall.com. The other county-level data are obtained from the American Community Survey.
Industry Similarity	Cosine similarity between vectors of employment shares of NAICS industry groups presented by the American Community Survey of the firm and the institution county.
I (Metro Pair)	Equals 1 if both firm and institution counties have Rural-Urban Continuum Codes less than or equal to 3.
I (Large Metro Pair)	Equals 1 if both firm and institution have Rural-Urban Continuum Codes equal to 1.
I (Top 100 Populous County Pair)	Equals 1 if both firm and institution belong to the top 100 populous counties as of 2018.
Citation Share	The fraction of institution state mentions in a firm's 10-K report. We also consider I (Cited), an indicator variable if the institution's state is cited in the firm's 10-K.
Revenue Share	The fraction of a firm's revenue derived from the institution state. We also consider I (Revenue State), which equals 1 if the firm reports nonzero revenue from the institution state. Firms' state revenues are obtained from Compustat segment data.
B. Firm-level variables	
%TIO	Common shares owned by institutional investors over total shares outstanding.
Illiquidity (ILLIQ)	The Amihud (2002) illiquidity measure, defined as quarterly average of $\|R\|_{i,t}/V_{i,t}$, where R is the daily return (in decimals) and $\$V$ is the dollar trading volume (scaled up by 10^6).

(*Continued*)

Table A.1
Continued

Variable	Definition
Effective Spread (Spread)	The quarterly average of daily dollar-weighted percentage effective spread (scaled up by 10^3). We obtain the daily dollar-weighted percentage effective spread from the WRDS Intraday Indicator database.
Tobin's Q	The ratio of market value and replacement cost of firm assets. Market value of equity is obtained from CRSP and is measured at the quarter end. Replacement cost of a firm's book equity plus total debt. Book equity is shareholders' equity, plus deferred taxes and investment tax credit, minus book value of preferred stock. The relevant variables are obtained from the Compustat Quarterly database and CRSP.
Market-to-Book (M/B)	The ratio of market value of equity and book value of equity. Book equity is defined as shareholders' equity, plus deferred taxes and investment tax credit, minus book value of preferred stock. The relevant variables are obtained from the Compustat Quarterly database. Market value of equity is obtained from CRSP and is measured at the quarter end.
Total Assets	Book value of total assets from Compustat. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag. We denote log total assets as Log Assets.
Return Volatility (Volatility)	Standard deviation of monthly returns over the past six months.
Share Turnover (Turnover)	Average trading volume scaled by total shares outstanding over the past six months.
Share Price (Price)	Historical price at the quarter end.
S&P 500 Dummy (S&P)	Dummy variable equal to one if the firm is included in the S&P 500 Index and zero otherwise.
Exchange Dummies	Dummy variables indicating the stock exchange of the firm.
Momentum	Cumulative monthly return from month t-2 to t-12 before the end of the quarter.
Firm Age (Age)	Number of months since a firm's first appearance in CRSP.
Dividend Yield (Yield)	Annual dividends distributed over the market price. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag from the fiscal year end.
R&D Expense (R&D)	Total research and development expenditures scaled by net sales. We set missing values of R&D to zero. We obtain the most recent data from the the Compustat Annual database and require at least a four-month lag from the fiscal year end.
Advertising Expense (Advertising)	Total advertising expenditures scaled by net sales. We set missing values to zero. We obtain the most recent data from Compustat Annual database and require at least a four-month lag from the fiscal year end.
Profitability	EBITDA scaled by book value of assets. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag from the fiscal year end.

(*Continued*)

Table A.1
Continued

Variable	Definition
Sales Growth	Sales growth is measured over the past three years. If less than three years of sales data are available, sales growth is estimated using all available data. Missing values are set to zero. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag from the fiscal year end.
Asset Turnover	Net sales over book value of total assets. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag from the fiscal year end.
Book Leverage (Leverage)	Book value of debt scaled by the book value of total assets. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag from the fiscal year end.
Dividend Payout (Payout)	Sum of dividends and repurchases divided by net income. We obtain the most recent data from the Compustat Annual database and require at least a four-month lag from the fiscal year end.
# Analyst	Number of one-year-ahead analyst estimates at the end of the quarter.
Size	Market capitalization (in millions) at the quarter end.
C. *Institution (mutual fund)-level variables*	
AUM	Assets under management in millions of dollars, based on the total market capitalization of their equity holdings.
Propensity Beta	Coefficient of Log Social Connectedness in our baseline specification for each institution as a proxy for the propensity to hold socially connected stocks.
D. *County-level vVariables*	
Social Proximity to Capital (SPC)	County j's social proximity to capital is defined as $\sum(\textit{County AUM}_i \times \textit{Social Connectedness}_{i,j})$, where county AUM is the sum of AUM of all the institutions located in a given county. Log Social Proximity to Capital is defined as $\log(SPC)$.
Physical Proximity to Capital (PPC)	County j's physical proximity to capital is defined as $\sum(\frac{\textit{County AUM}_i}{1+\textit{Distance}_{i,j}})$, where county AUM is the sum of AUM of all the institutions located in a given county. Log Physical Proximity to Capital is defined as $\log(PPC)$.
Population	County population. Missing values are replaced by state average. Obtained from the American Community Survey. Log Population is defined as $\log(Population)$.
High School Attainment Dummy (High School)	Equals one if the percentage of adults in a county who obtain a high school education is greater than the sample median. Missing values are replaced by zero. Obtained from the American Community Survey.
College School Attainment Dummy (College)	Equals one if the percentage of adults in a county who obtain a college education is greater than the sample average. Missing values are replaced by zero. Obtained from the American Community Survey.

(*Continued*)

Table A.1
Continued

Variable	Definition
Income Per Capita (Income)	County-level income per capita. Missing values are replaced by state average. Obtained from the American Community Survey.
Employment Rate (Employment)	County-level employment rate. Missing values are replaced by state average. Obtained from the American Community Survey.
Industry Agglomeration (Agglomeration)	Log number of firms in the same industry within a 50-mile radius. Missing values are replaced by state median. Obtained from the American Community Survey. We denote $\log(Agglomeration)$ as Log Agglomeration.
Affected Capital Ratio	Social proximity to the institutional capital located in the mid-Atlantic area, divided by overall social proximity to capital.
Affected Ownership Ratio	Ownership by institutions located in the mid-Atlantic area, divided by total institutional ownership.

References

Aghion, P., J. Van Reenen, and L. Zingales. 2013. Innovation and institutional ownership. *American Economic Review* 103:277–304.

Ahern, K. R. 2017. Information networks: Evidence from illegal insider trading tips. *Journal of Financial Economics* 125:26–47.

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005. Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy* 113:151–84.

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5:31–56.

Baik, B., J.-K. Kang, and J.-M. Kim. 2010. Local institutional investors, information asymmetries, and equity returns. *Journal of Financial Economics* 97:81–106.

Bailey, M., R. Cao, T. Kuchler, and J. Stroebel. 2018a. The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy* 126:2224–76.

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong. 2018b. Social connectedness: Measurements, determinants, and effects. *Journal of Economic Perspectives* 32:259–80.

Bailey, M., E. Dávila, T. Kuchler, and J. Stroebel. 2019a. House price beliefs and mortgage leverage choice. *Review of Economic Studies* 86:2403–52.

Bailey, M., P. Farrell, T. Kuchler, and J. Stroebel. 2019b. Social connectedness in urban areas. Working Paper w26029, NBER.

Bailey, M., A. Gupta, S. Hillenbrand, T. Kuchler, R. Richmond, and J. Stroebel. 2021. International trade and social connectedness. *Journal of International Economics* 129:103418–.

Bailey, M., D. Johnston, M. Koenen, T. Kuchler, D. Russel, and J. Stroebel. 2020a. Social distancing during a pandemic: The role of friends. Working Paper w28234, NBER.

Bailey, M., D. M. Johnston, T. Kuchler, J. Stroebel, and A. Wong. 2019c. Peer effects in product adoption. Working Paper, NBER.

Bailey, M., T. Kuchler, D. Russel, B. State, and J. Stroebel. 2020b. Social connectedness in Europe. Working Paper, NBER.

Banerjee, A., A. Chandrasekhar, E. Duflo, and M. O. Jackson. 2013. The diffusion of microfinance. *Science* 341: doi: 10.1126/science.1236498.

Baum-Snow, N., and R. Pavan. 2011. Understanding the city size wage gap. *Review of Economic Studies* 79:88–127.

Becker, B., H. Cronqvist, and R. Fahlenbrach. 2011. Estimating the effects of large shareholders using a geographic instrument. *Journal of Financial and Quantitative Analysis* 46:907–42.

Bernile, G., A. Kumar, and J. Sulaeman. 2015. Home away from home: Geography of information and local investors. *Review of Financial Studies* 28:2009–49.

Bernile, G., A. Kumar, J. Sulaeman, and Q. Wang. 2019. Has local informational advantage disappeared? *Review of Financial Economics* 37:38–60.

Blume, M. E., and D. B. Keim. 2012. Institutional investors and stock market liquidity: Trends and relationships. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Brown, J. R., Z. Ivković, P. A. Smith, and S. Weisbenner. 2008. Neighbors matter: Causal community effects and stock market participation. *Journal of Finance* 63:1509–31.

Bushee, B. J. 1998. The influence of institutional investors on myopic R&D investment behavior. *Accounting Review* 305–33.

———. 2001. Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting Research* 18:207–46.

Chang, S. 2019. Local industry bias in investor behavior: Evidence from mutual funds. Working Paper, SSRN.

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

Chetty, R., N. Hendren, M. O. Jackson, T. Kuchler, J. Stroebl, R. Fluegge, F. Gonzalez, M. Jacob, M. Koenen, E. Laguna-Muggenburg, et al. 2021. Social capital in the United States: Measurement, determinants, and associations with economic mobility. Working Paper.

Chung, K. 2018. Changing career incentives and risk-taking in the mutual fund industry. Working Paper, SSRN.

Coeurdacier, N., and H. Rey. 2013. Home bias in open economy financial macroeconomics. *Journal of Economic Literature* 51:63–115.

Cohen, L., A. Frazzini, and C. Malloy. 2008. The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy* 116:951–79.

———. 2010. Sell-side school ties. *Journal of Finance* 65:1409–37.

Cohen, L., U. G. Gurun, and C. Malloy. 2017. Resident networks and corporate connections: Evidence from World War II internment camps. *Journal of Finance* 72:207–48.

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

Correia, S., P. Guimarães, and T. Zylkin. 2019. PPMLHDFE: Fast Poisson estimation with high-dimensional fixed effects ISSN arXiv, preprint, http://arxiv.org/abs/1903.01690.

Coval, J. D., and T. J. Moskowitz. 1999. Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54:2045–73.

———. 2001. The geography of investment: Informed trading and asset prices. *Journal of Political Economy* 109:811–41.

Crawford, S. S., W. R. Gray, and A. E. Kern. 2017. Why do fund managers identify and share profitable ideas? *Journal of Financial and Quantitative Analysis* 52:1903–26.

Dougal, C., C. A. Parsons, and S. Titman. 2018. Urban vibrancy and firm value creation. Working Paper, SSRN.

Duflo, E., and E. Saez. 2002. Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of Public Economics* 85:121–48.

———. 2003. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics* 118:815–42.

Duggan, M., S. Greenwood, and A. Perrin. 2016. Social media update 2016. Pew Research Center. Https://www.pewresearch.org/internet/2016/11/11/social-media-update-2016/.

Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1–22.

Fang, L., and J. Peress. 2009. Media coverage and the cross-section of stock returns. *Journal of Finance* 64:2023–52.

French, K. R., and J. M. Poterba. 1991. Investor behavior and international diversification. *American Economic Review* 81:222–26.

Gompers, P. A., and A. Metrick. 2001. Institutional investors and equity prices. *Quarterly Journal of Economics* 116:229–59.

Green, T. C., and R. Jame. 2013. Company name fluency, investor recognition, and firm value. *Journal of Financial Economics* 109:813–34.

Heimer, R. 2016. Peer pressure: Social interaction and the disposition effect. *Review of Financial Studies* 29:3177–209.

Hirshleifer, D. 2020. Presidential address: Social transmission bias in economics and finance. *Journal of Finance* 75:1779–831.

Hirshleifer, D., L. Peng, and Q. Wang. 2020. Social networks and market reactions to earnings news. Working Paper, SSRN.

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

Hong, H., J. D. Kubik, and J. C. Stein. 2004. Social interaction and stock-market participation. *Journal of Finance* 59:137–63.

———. 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance* 60:2801–24.

———. 2008. The only game in town: Stock-price consequences of local bias. *Journal of Financial Economics* 90:20–37.

Hong, H., T. Lim, and J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55:265–95.

Hong, H., and J. Xu. 2019. Inferring latent social networks from stock holdings. *Journal of Financial Economics* 131:323–44.

Huberman, G. 2001. Familiarity breeds investment. *Review of Financial Studies* 14:659–80.

Ivković, Z., and S. Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *Journal of Finance* 60:267–306.

Kang, J.-K., R. M. Stulz, et al. 1997. Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan. *Journal of Financial Economics* 46:3–28.

Kaustia, M., and S. Knüpfer. 2012. Peer performance and stock market entry. *Journal of Financial Economics* 104:321–38.

Kelly, M., and C. O. O'Grada. 2000. Market contagion: Evidence from the panics of 1854 and 1857. *American Economic Review* 90:1110–24.

Kuchler, T., D. Russel, and J. Stroebel. 2021. Jue insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by facebook. *Journal of Urban Economics* 103314.

Kuchler, T., and J. Stroebel. 2020. Social finance. Working Paper, NBER.

Lehavy, R., and R. G. Sloan. 2008. Investor recognition and stock returns. *Review of Accounting Studies* 13:327–61.

Liu, H., and Y. Wang. 2016. Market making with asymmetric information and inventory risk. *Journal of Economic Theory* 163:73–109.

Massa, M., and A. Simonov. 2006. Hedging, familiarity and portfolio choice. *Review of Financial Studies* 19:633–85.

Maturana, G., and J. Nickerson. 2018. Teachers teaching teachers: The role of workplace peer effects in financial decisions. *Review of Financial Studies* 32:3920–57.

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42:483–510.

Miller, E. 1977. Risk, uncertainty and divergence of opinion. *Journal of Finance* 32:1151–68.

Moretti, E. 2012. *The new geography of jobs*. New York: Houghton Mifflin Harcourt.

Newey, W. K., and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61:631–53.

Oster, E. 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics* 37:187–204.

Ouimet, P., and G. Tate. 2020. Learning from coworkers: Peer effects on individual investment decisions. *Journal of Finance* 75:133–72.

Parsons, C. A., R. Sabbatucci, and S. Titman. 2020. Geographic lead-lag effects. *Review of Financial Studies* 33:4721–70.

Pool, V. K., N. Stoffman, and S. E. Yonker. 2012. No place like home: Familiarity in mutual fund manager portfolio choice. *Review of Financial Studies* 25:2563–99.

———. 2015. The people in your neighborhood: Social interactions and mutual fund portfolios. *Journal of Finance* 70:2679–732.

Rehbein, O., and S. Rother. 2020. Distance in bank lending: The role of social networks. Working Paper, University of Bonn and University of Mannheim, Germany.

Rehse, D., R. Riordan, N. Rottke, and J. Zietz. 2019. The effects of uncertainty on market liquidity: Evidence from Hurricane Sandy. *Journal of Financial Economics* 134:318–32.

Rubin, A. 2007. Ownership level, ownership concentration and liquidity. *Journal of Financial Markets* 10:219–48.

Russolillo, S. 2012. Morning marketbeat: Hurricane Sandy closes markets. *Wall Street Journal* October 29.

Scheinkman, J. A., and W. Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 111:1183–220.

Seasholes, M. S., and N. Zhu. 2010. Individual investors and local bias. *Journal of Finance* 65:1987–2010.

Shive, S. 2010. An epidemic model of investor behavior. *Journal of Financial and Quantitative Analysis* 45:169–98.

Shue, K. 2013. Executive networks and firm policies: Evidence from the random assignment of MBA peers. *Review of Financial Studies* 26:1401–42.

Silva, J. S., and S. Tenreyro. 2006. The log of gravity. *Review of Economics and Statistics* 88:641–58.