Social Proximity to Capital: Implications for Investors and Firms*

Theresa Kuchler†
Yan Li‡
Lin Peng§
Johannes Stroebel¶
Dexin Zhou||

Abstract

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 higher liquidity. These effects are largest for small firms with little analyst coverage, suggesting that the investors' behavior is explained by their increased 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 Codes: G2, G3, G4

Keywords: Social Networks, Social Connectedness Index, Institutional Investors, Familiarity Bias

*This draft: February 6, 2021. We thank Elsa Allman, Yakov Amihud, 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, Wenlan Qian, Yelena Larkin, Gabriele Lattanzio (discussant), Marcin Kacperczyk, Charles Lee, Ramana Nanda, Marina Niessner (discussant), Cameron Peng, Veronika Pool (discussant), Oliver Rehbein, Simon Rother, Noah Stoffman, Wei Xiong, Hongjun Yan, Jun Yang, as well as seminar participants at Baruch, Cambridge, Imperial, NYU Stern, the New York Fed, UC Irvine, Bonn University, Chicago Fed, Temple, University of Miami, Florida State University, and conference participants at AFA, 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 acknowledges the Krell research grant and the Wasserman research grant for financial support.

†Stern School of Business, New York University, NBER, CEPR. Email: theresa.kuchler@stern.nyu.edu.

‡Zicklin School of Business, Baruch College, One Bernard Baruch Way, Box 10-225, New York, NY 10010. Email: yan.li@baruch.cuny.edu.

§Zicklin School of Business, Baruch College, One Bernard Baruch Way, Box 10-225, New York, NY 10010. Phone: (646) 312-3491, Email: lin.peng@baruch.cuny.edu.

¶Stern School of Business, New York University, NBER, CEPR. Email: johannes.stroebel@stern.nyu.edu.

||Zicklin School of Business, Baruch College, One Bernard Baruch Way, Box 10-225, New York, NY 10010. Email: dexin.zhou@baruch.cuny.edu.

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; Moretti, 2012). A growing literature explores a variety of explanations for these disparities. For example, Dougal et al. (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 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. (2018a). This measure is based on friendship links on Facebook, the world's largest online social networking service with 239 million active users in the U.S. 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]

[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 the South to Northern industrial cities in the 1940s-1960s shows up as stronger present-day friendship links between Chicago and Mississippi. As a result, we argue that the investor Northern Trust, based in Chicago, is disproportionately connected to the firm Trustmark Corporation, based in Mississippi. See Bailey et al. (2018a, 2019b) for detailed discussions of the determinants of social connectedness between locations.

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 et al., 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 (New York, New Jersey, and Connecticut), with the largest concentration of institutional assets under management (AUM), only accounts for 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.[2] After flexibly controlling for the geographic distance between the locations of firms and investors, we find that a 10% increase in the social connectedness between firm and investor locations is associated with a 3.1% increase in the weight of the firm in the investor's portfolio. Our extensive set of fixed effects 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.

A large 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 et al., 2015).[3] Consistent with this literature, we find that investments are decreasing in the geographic distance between firms and investors. However, after controlling for the social connectedness between firm and investor locations, the effect of physical distance on investments essentially disappears. Moreover, the inclusion of controls for physical distance does not affect the estimated effect of social connectedness on investments. These findings suggest 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.

[2] As such, our results contribute to a literature on how social interaction affect economic decisions (e.g., Hong et al., 2004; Cohen et al., 2008, 2010, 2017; Hochberg et al., 2007; Chen et al., 2010; Kaustia and Knüpfer, 2012; Shue, 2013; Pool et al., 2015; Bailey et al., 2018b; Allen et al., 2020; Ouimet and Tate, 2020). See Hirshleifer (2020) and Kuchler and Stroebel (2020) for reviews.

[3] See also Huberman (2001), Massa and Simonov (2006), Baik et al. (2010), Seasholes and Zhu (2010), and Becker et al. (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 et al. (1997), and Coeurdacier and Rey (2013).

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 a firm's 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 non-financial 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.[4]

Despite the many controls and fixed effects in the previous cross-sectional analyses, one might worry about possible omitted variables at the firm-investor-pair level that can correlate with the social connectedness between firm and investor locations. Some of these omitted variables could possibly affect an investor's propensity to hold stocks of a particular firm through a channel other than social interactions. We strengthen our preferred interpretation of the observed patterns by directly addressing two alternative plausible explanations. First, since individuals in socially connected counties have similar preferences on average—a feature referred to as homophily—one might worry that investors are more likely to invest in firms headquartered in socially connected counties because these firms have certain business models or practices viewed more favorably by the investor. While this story would probably not generate the heterogeneities described above, we also address it directly by controlling for the similarity of firm and institution counties along a large number of demographic and economic characteristics, which would at least partially capture any role played by common preferences. We find that our results are not affected by the addition of these controls, suggesting that common preferences between connected counties are not an important driver of our findings. We also rule out a second alternative explanation for our findings whereby direct economic links between socially connected regions (for example, because firms have major operations in socially linked regions) provide an alternative mechanism through which investors become more aware of the existence of lesser-known firms.

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 through December 2016. This analysis allows us to directly include institution $\times$ firm fixed effects to capture any time-invariant deter-

[4] This finding is consistent with the results of Pool et al. (2012), who document that managers from smaller fund families disproportionately overweight firms from their home state in their investment decisions. Similarly, Hirshleifer et al. (2019) 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.

minants 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 location 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 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.

In the second part of the paper, we ask whether the tendency of institutional investors to disproportionately invest in areas to which they are socially connected aggregates 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 institutional 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 and a 0.59% increase in its Tobin's Q. 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 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 0.86% reduction in effective spreads and a 2.7% reduction in the Amihud (2002) illiquidity measure. As before, these results survive the addition of firm fixed effects and are 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 further address such concerns, we exploit an exogenous shock to a *subset* of capital providers that resulted in differential cross-sectional liquidity impacts 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 provide evidence suggesting that, during this period, institutional investors in the affected Mid-Atlantic region were likely to reduce 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 with otherwise similar firms that had the same overall social proximity to capital. 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.

Overall, these results provide—to the best of our knowledge—the first evidence that social proximity to capital affects aggregate firm outcomes such as liquidity and valuations. While a number of papers have documented that various social interactions can affect individual investment decisions, our novel measure of social connectedness allows us to show that these effects can have aggregate implications for firms. It also suggests one mechanism through which the geographic structure of social networks can shape regional variation in economic outcomes.

In the final part of the paper, 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 et al., 2008; Hong et al., 2005; Hong and Xu, 2019; Pool et al., 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 et al., 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. We first estimate each institution's propensity to invest in socially proximate firms and sort institutions into deciles based on that propensity. Comparing across these deciles, we find no significant variation in investor performance, measured by excess returns, CAPM, Fama and French (2015) five-factor alphas, risk, or Sharpe Ratios. When we look within institutions, we also find no differential performance between an institution's connected holdings and its non-connected holdings. Thus, our results do not support the idea that institutional investors obtain valuable information through social connections as measured by Facebook friendship links.[5] Instead, our results are more consistent with an interpretation in which institutional investments in socially proximate firms are driven by an increased awareness of these firms.

1 Social Connectedness and Institutional Investments

In the first part of the paper, we 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.

[5]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 might 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 affected investment patterns. This is not inconsistent with other evidence that suggests that professional networks are more able to convey useful information (see, for example, the evidence on this presented in Cohen et al., 2008).

Measuring Social Connectedness. To measure social connectedness between U.S. counties, we use the *Social Connectedness Index* first introduced by Bailey et al. (2018a). 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 U.S. 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 U.S. used Facebook (Duggan et al., 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 U.S., Facebook mainly serves as a platform for real-world friends and acquaintances to interact online, and people usually only add connections on Facebook to individuals whom 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 uni-directional links to non-acquaintances are common. Consistent with this, Bailey et al. (2018a,b, 2019a,b,c, 2020a,b,c), Kuchler et al. (2020), and Rehbein et al. (2020) provide evidence that friendships observed on Facebook are a good proxy for real-world U.S. social connections.

The Social Connectedness Index of Bailey et al. (2018a) maps Facebook users to their respective county locations using information such as the users' regular IP addresses.[6] They then construct a measure of the relative number of friendship links between each county pair, $Friendships_{i,j}$. Our 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:

$$\mathit{Social\ Connectedness}_{i,j} = \frac{\mathit{Friendships}_{i,j}}{\mathit{Population}_i \times \mathit{Population}_j}, \tag{1}$$

where $Population_i$ corresponds to the total population in county i. Figure 1 shows heatmaps of our measure of $\mathit{Social\ Connectedness}_{i,j}$ for San Francisco County, CA, in Panel A and for Cook County, IL, 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

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

County is strongly connected to nearby counties in coastal California. However, social connectedness is not solely determined by physical proximity. For example, San Francisco County also has strong connections to other urban areas, such as New York or 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 San Francisco for career opportunities. 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. 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.

[Insert Figure 1 near here]

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 use institutional investment data from June 2016, which most closely corresponds to the time when we observe social connectedness (there is no time-series data for social connectedness available to researchers). We then 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.[7] In particular, for each institution-firm pair, we construct a measure of institutional holding, $\%PF_{i,j}$, 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 et al. (2015, 2019), who collect this information from Nelson's Directory of Investment Managers and by searching SEC filings.[8] We obtain firms' historical headquarters locations from Compustat.

[7] We limit our analysis to stocks listed on NYSE, Nasdaq, and NYSE MKT 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.

[8] We are grateful for 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.

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 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.

[Insert Table 1 near here]

1.1 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:

$$\%PF_{i,j} = \exp\left[\beta \mathit{Log\ Social\ Connectedness}_{i,j} + \gamma X_{i,j} + \psi_i + \xi_j\right] \cdot \epsilon_{i,j}. \quad (3)$$

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.[9] 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 fixed effects. 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.

[Insert Figure 2 near here]

We report results from regression 3 in Table **2**. The first column shows our baseline estimates. The coefficient on $\mathit{Log\ Social\ Connectedness}_{i,j}$ 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 percent increase in social connectedness is associated with a 1.89 percent increase in $\%PF$. Importantly, the inclusion of

[9] In the analysis in Figure 2, we drop all observations where $\%PF$ equals zero. However, our PPML estimation procedure allows the inclusion of these observations in our regression analysis.

[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., 2020a). We use the estimation procedure implemented in Correia et al. (2019).

firm fixed effects ensures that this finding is not driven by characteristics that might make firms located in socially connected counties more prone to attracting institutional investments on average.

One possible explanation for the estimated coefficients is that investors may strategically set up headquarters in regions that are socially connected to industry clusters important to the investors' investment objectives. To show that such location decisions do not explain our findings, we next include institution $\times$ Fama-French 48 industry fixed effects to account for institutions' preferences to locate in areas that are socially connected to particular industry clusters. Column 2 shows that the inclusion of these fixed effects increases the regression R^2 from 32% to 51%, suggesting that industry preference explains a substantial portion of the cross-sectional variation in funds' holding choices. The coefficient estimate of β, which captures the effect of social connections on investment choices, remains unchanged, suggesting that our results are not explained by funds' strategic location decisions.

[Insert Table 2 near here]

In column 3 of Table Table 2, we examine the role of physical distance on investment choices. Consistent with the local bias literature, we find that institutional investors tend to overweight firms that are headquartered in close geographic proximity. In column 4, we include both *Log Social Connectedness* and *Log Distance* as explanatory variables. These two variables are highly correlated with one another ($\rho = -0.69$). *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. This finding therefore helps to advance the literature on local bias, which often interprets geographic local bias as a result of strong connections between fund managers and corporate executives located close by (e.g., Coval and Moskowitz, 1999, 2001; Baik et al., 2010; Bernile et al., 2015). Our results are consistent with this interpretation but indicate that *Social Connectedness* is a richer proxy for the underlying social networks than physical distance is.

The next specifications alleviate concerns that our measure of social connectedness might be picking up a potentially non-linear relationship between physical distance and institutional investments. In column 5, we control for the geographic distance between county pairs using 500-tile distance dummies (Appendix Figure IA.1 reports the coefficients on these dummies). In column 6, we further add indicators for whether the firm-institution pair is located in the same county or the same state, allowing us to control for the word-of-mouth interaction over short distances (see Hong et al., 2004). The effect of social connectedness on investments even increases somewhat in these specifications. Finally, column 7 explores the monotonicity of the relationship between social connectedness and investments by including quintile indicators of Social Connectedness as explanatory variables. Coefficients for quintiles 2 to 5 represent the relative increases in investments compared to firm-institution pairs in the first quintile of social connectedness. Consistent with the findings in Figure 2, there is a monotonic increase in portfolio weights across quintiles of social connectedness.[11]

Controlling for County Similarities and Economic Linkages. 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. We next rule out two possible alternative interpretations before presenting additional evidence in favor of our preferred interpretation.

First, we consider whether 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, a pattern that has been termed "homophily." For example, it is possible that when two counties have similar population compositions or similar political preferences, investors residing in one county are naturally more likely to invest in firms headquartered in the other county, perhaps because these firms engage in certain business models or business practices viewed more favorably by the investor.

To explore whether homophily can explain our results, column 1 of Table **3** adds additional control variables that capture potential similarities between institution and firm locations, along the dimensions of population, age, employment, income, education, immigration history, marital status, political inclination, and industry composition.[12] The relationship between social connectedness and portfolio

[11] In the Appendix, we present additional findings to ensure the robustness of our baseline results to alternative specifications. Column 1 of Panel A in 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 county pair-level regression where the dependent variable is the fraction of 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 non-zero institutional holding as the dependent variable. In columns 2 to 4 of Panel B, we consider alternative controls for physical distance using 10, 50, and 100-tile indicators as distance controls. In columns 4 and 5, we interact social proximity to capital with indicators for whether the firm-institution pair is located within 100 miles. Our results are robust to using these alternative specifications. Additionally, we find that social proximity to capital helps explain institutional investments in both local and non-local firms.

[12] Our control variables are defined as $|C_i - C_j|$, where C represents the following county characteristics: log county population, population density, median age, income per capita, share at least high school, share college, employment rate, share immigrants, share unmarried, and Democratic vote share. Industry composition similarity is measured as the cosine distance between vectors of employment shares of NAICS industry groups between the firm and the institution county.

holding is not affected by these additional controls, suggesting that our findings are not just the result of a similarity of preferences between investor and firm counties driving the observed investments.[13]

[Insert Table 3 near here]

Next, 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 end up learning about these firms not through their social networks but by observing their activity in their own county. To explore the extent to which this mechanism explains our findings, we explicitly control for two measures of the economic linkages between firm and investor counties.

First, following Bernile et al. (2015), we measure the economic links of a firm to an investor's location through mentions of the investors' headquarters state in the firm's 10-K filings. In column 2 of Table 3, we control for citation share, defined as the share of mentions of the institutions' headquarters states in all state mentions. In column 3, we instead include an indicator for whether the institution's headquarters state has been mentioned in the firm's 10-K statement. Consistent with Bernile et al. (2015), we find both measures to be positively correlated with portfolio holdings. However, controlling for these variables does not affect our estimates of the relationship between social connectedness and investments.

The second proxy for economic connectedness uses information about a firm's major customer locations. In column 4, we control for the percentage of sales derived from customers located in the institution's headquarters state.[14] In column 5, we instead include a dummy variable indicating whether the institution's headquarters is located in a state from which the firm derives major revenue. We find a positive relationship between firms' revenue share and institutions' holdings, but controlling for revenue sources does not affect the estimated relationship between social connectedness and investments.

None of the many controls described above significantly affect the estimated impact of social connectedness. Thus, we conclude that the effect of social connectedness on investments does not primarily 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 Tables 2 and 3 reduces the likelihood that omitted unobserved characteristics confound our analysis (see Altonji et al., 2005; Oster, 2019).

[13]In column 1 of Panel B of Appendix Table IA.2, we also examine whether our results are affected by whether both cities are in a metropolitan area, which might affect our results if the cost of flying between metro areas is lower, making it easier for investors to visit local firms and ultimately invest. We find that our results are not affected by this control.

[14]This relationship is measured using the Compustat Segment History dataset. Disclosures of major revenue locations are relatively sparsely populated, and we treat missing values as 0.

1.2 Heterogeneity Analysis

A key takeaway from the previous analysis is that institutional investors tend to disproportionately invest in firms from areas that the investors are socially connected to. We propose 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 proposed mechanism behind the observed relationship between social connections and investments.

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 non-financial 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 $\textit{Log Social Connectedness}_{i,j}$. The results are reported in Panel A of Table 4. Column 1 shows the baseline result, where we control for firm $\times$ Bushee-type and institution $\times$ industry 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. This finding is unaffected by including controls for the physical distance between firm and fund, even when those are interacted with Bushee-type indicators. Indeed, in the specification with the richest controls in column 4 of Table 4, the investments of dedicated investors react four times more strongly to variation in social connectedness as the investments of transient investors do.

[Insert Table 4 near here]

In Panel B of Table 4, we split investors into terciles based on their total AUM. We find that investments of small institutions respond more than twice as much to social connectedness than 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 et al., 2012).

Heterogeneity by Firm Characteristics. Next, we explore the heterogeneity of investment sensitivity to social connectedness across various firm characteristics. Following Hong et al. (2000), Parsons et al. (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 to (though our results in later sections suggest that they do not, in fact, appear to have such an information advantage).

Panel A of Table 5 presents results when splitting firms into size terciles based on their market equity. We refer to these size terciles as Small Cap, Mid Cap, and Large Cap. The specification in column 1 shows that the relationship between social connectedness and investments is indeed declining in firm size. The difference in the effect of social connectedness on the investments in Small Cap and Large Cap firms is sizable and highly statistically significant. In columns 2-4, we add our standard controls for geographic distance, interacted with fixed effects for firm size terciles. Our results are not affected by the addition of these control variables.

[Insert Table 5 near here]

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.[15] These results are reported in Panel B of Table 5. Column 1 shows the baseline results. We find that investments in low-analyst-coverage firms respond twice as much to social connectedness as investments in high-analyst-coverage firms do. Our results are robust to including controls for geographic distance inter-

[15] Analyst coverage measures the number of analysts issuing an earnings forecast for the current fiscal year. We obtain measures of analyst coverage for each firm from I/B/E/S summary file in the quarter prior to the institutional holding report date.

acted with analyst-coverage-tercile indicators. 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.[16]

1.3 Within-Firm Identification Using Panel Data

Our previous results show that institutional investors tend to disproportionately invest 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 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.

For this analysis, we explore quarterly panel data of institutional investment holdings between June 2007 and December 2016. The first three columns of Table 6 reproduce the specifications in columns 2, 5, and 6 in 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.[17]

[Insert Table 6 near here]

In the following specifications, we also include firm $\times$ institution fixed effects, which control for any time-invariant firm-institution-pair-specific 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.[18] Overall, about 10% of firms in our sample changed headquarters locations during the sample period.

[16] 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 Appendix Table IA.3. We find that even after controlling for firm size tercile, investment sensitivity to social connectedness is highest for firms with low analyst coverage.

[17] Column 1 in Panel C of Appendix Table IA.2 shows that the results are also robust to controlling for institution style (Bushee Type) $\times$ firm $\times$ quarter fixed effects. Columns 2 to 4 of that Table mirror results from Table 3, and show that controlling for county-level economic linkages does not affect the conclusion from the panel analysis.

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

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 6 is statistically significant but smaller than the effect estimated in column 3.[19] This is perhaps unsurprising. For firms with headquarters changes, investor awareness is unlikely to adjust immediately following the 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 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 headquarters moves lead to larger changes in investments a few years after the headquarters change.

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.

1.4 Fund Manager Characteristics, Social Connectedness, and Fund Investment

Our next analysis turns to a sample of actively managed U.S. domestic equity mutual funds.[20] 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.[21]

Table 7 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 $\times$ fund style and fund $\times$ industry fixed effects (fund styles are based on the Lipper fund classification). Column 2 reports a placebo test using a sample of index funds. Since index fund man-

[19] We also note that standard errors are higher for this regression in part due to reduction in the degrees-of-freedom following the inclusion of firm $\times$ institution fixed effects.

[20] 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.

[21] 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 dataset. 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 of old management teams is 56 years. 662 of these funds 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 with MBA degrees, while 462 funds have more than 50% managers with MBA degrees.

agers do not actively select stocks, their social networks should not significantly affect their holdings. Consistent with this hypothesis, we find that there is no statistically significant 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 if 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 larger 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. This result suggests that formal business education does not affect fund managers' reliance on social connectedness in their investment decisions.

[Insert Table 7 near here]

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 2007-2016 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:

$$Social\ Proximity\ to\ Capital_{i,t} = \sum_j AUM_{j,t} \times Social\ Connectedness_{i,j}, \tag{4}$$

where $AUM_{j,t}$ is the total assets under management by institutions headquartered in county j in quarter t, and $Social\ Connectedness_{i,j}$ 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.[22]

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 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*.

[Insert Figure 3 near here]

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:

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

[22]Since social connectedness has the largest effect on investments of "dedicated" investors, we also explored separate measures of social proximity to "dedicated" capital. However, funds from different groups are located in similar counties: counties with high AUM for "dedicated" investors usually also have high AUM for other investors. As a result, it is not possible to obtain variation in social proximity to "dedicated" capital that is independent of social proximity to overall capital.

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 et al., 2010; Green and Jame, 2013) as well as controls for county characteristics.[23]

Our baseline specification, reported in column 1 of Table **8**, 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 20.3 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 additionally 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 (i.e., 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.[24]

[Insert Table 8 near here]

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 **8** show that *Social Proximity to Capital* has the largest effect on the institutional ownership share 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 mid-size 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.

[23]Our firm controls include log total assets (Log Assets), log market-to-book ratio (Log M/B), return volatility (Volatility), 12-month momentum (Momentum), share turnover (Turnover), lag stock price (Price), R&D expenses over net sales (R&D), dividend yield (Yield), an S&P 500 index dummy (S&P), firm age (Age), advertising expenditures over net sales (Advertising), and exchange dummies (Exchange). Our county controls include log physical proximity to capital (Physical Proximity), log county population (Population), high school educational attainment (High School), college educational attainment (College), log number of nearby firms in the same industry (Agglomeration), income per capita (Income), and employment rate (Employment). Detailed descriptions of these variables can be found in Appendix A.1. We use the latest available information to calculate these variables at the end of the prior quarter.

[24]Appendix Table IA.4 shows that these results, as well as our other firm-level results, are generally stronger among firms with low physical proximity to capital, suggesting that physical and social proximity to capital are substitutes in facilitating access to capital.

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 narrower investor base facilitates less risk-sharing, which leads to lower valuations and a higher cost of capital. A similar argument was made by Hong et al. (2008) in the context of investors' local bias. This theory has found 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 but disagree on 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 run the following regression:

$$\mathit{Log\ Valuation}_{i,t} = \beta \mathit{Log\ Social\ Proximity\ to\ Capital}_{i,t-1} + \gamma X_{i,t-1} + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \quad (7)$$

where $\mathit{Valuation}_{i,t}$ represents the market valuation of firm i in quarter t. We consider two measures of valuation, Tobin's Q and the market-to-book to ratio.[25] The dependent variables are 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.[26] The fixed effects and stan-

[25]The market-to-book ratio is defined as market capitalization divided by book equity. Market capitalization is obtained from CRSP, and book equity is obtained from Compustat. Tobin's Q is defined as market value of the firm over the replacement cost of its assets, and is obtained from Compustat.

[26]We include the following firm-level control variables: Log Assets, Profitability, Sales Growth, Asset Turnover, R&D, Advertising, book leverage (Leverage), dividend payout (Payout), and S&P, firm age (Age), Exchange. We also control for county-level measures: Log Physical Proximity, Population, High School, College, Agglomeration, Income, and Employment. A detailed description of these variables can be found in Appendix A.1.

dard errors correspond to those in the specifications in Table 8.

The results from regression 7 are reported in Table 9. The dependent variable in Panel A is the log of the market-to-book ratio. 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 9 highlights that the effects of *Social Proximity to Capital* on market-to-book ratios are strongest among small and mid-size firms and among firms with limited analyst coverage.

[Insert Table 9 near here]

Our second measure of firm valuation is Tobin's Q. Columns 1 and 2 of Panel B of Table 9 report the full sample results. We find that a 10% increase in *Social Proximity to Capital* is associated with a 0.6% increase in Tobin's Q. To get a better sense of the economic magnitude implied by our estimates, we compare our coefficient to the estimated effect of another explanatory variable, firm age. In our sample, a one-standard-deviation increase in firm age is associated with a 3.1% decrease in Tobin's Q, which is qualitatively consistent with Green and Jame (2013). Correspondingly, our estimates imply that a one-standard-deviation increase in *Log Social Proximity to Capital* (which corresponds to a 1.10 log point increase) is associated with a 6.5% increase in Tobin's Q. When we include firm fixed effects in column 3, we do not observe a significant change in the coefficient estimate.

In columns 4 and 5, we report the differential effects of *Social Proximity to Capital* on Tobin's Q for firms of different sizes. We do not find a significant difference for firms in the top and bottom terciles of the size distribution. Columns 6 and 7 report the results separately for firms with differential analyst coverage. We find that the effect of social proximity to capital on Tobin's Q is generally stronger among firms with lower 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 (see Green and Jame, 2013; Liu and Wang, 2016). We conduct the following regression analysis:

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

where $Liquidity_{i,t}$ represents one of two measures of secondary market liquidity of firm i in quarter t: the effective percentage spread and the Amihud (2002) illiquidity measure.[27] The dependent variables are 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.[28]

[Insert Table 10 near here]

The regression results are reported in Panel A of Table **10**. 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 significant at the 1% level when we include Quarter $\times$ Industry fixed effect in column 2 and firm fixed effects in column 3.

Next, we estimate the effect of social proximity to capital separately for firms of different sizes and with different analyst coverage. Columns 4 and 5 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 Panel B of Table **10**, we report the same set of tests with the Amihud (2002) illiquidity measure as the dependent variable. In the full-sample analysis reported in column 1, we find that a 10% increase in *Social Proximity to Capital* is related to a 2.9% decrease in illiquidity. This effect is robust to including

[27] The effective percentage spread is obtained from the Intraday Indicator database in WRDS. It is defined as two times the dollar-trading-volume-weighted absolute difference between trading price and midpoint price (scaled up by 10^3). We aggregate this daily measure into a quarterly measure by taking the quarterly average. The Amihud (2002) measure is defined as quarterly average of $|RET_{i,t}|/Dollar\ Volume_{i,t}$ for stock i in day t. We scale this measure up by 10^6 when reporting summary statistics. The intuition behind this measure is that a liquid stock can allow a high trade volume to pass through in any given day without a significant change in price.

[28] We include the following firm-level controls: Log Assets, Profitability, Log M/B, Log Volatility, Momentum, Price, S&P, Age, Advertising, and Exchange. We also control for county-level measures of Log Physical Proximity, Population, High School, College, Agglomeration, Income, and Employment. A detailed description of these variables can be found in Appendix A.1.

quarter × industry fixed effects, as reported in column 2. The magnitude of the effect is economically meaningful: a one-standard-deviation increase in profitability is associated with a 13.8% reduction in the Amihud illiquidity measure, while a one-standard-deviation increase in *Log Social Proximity to Capital* is associated with a 33% decrease in the Amihud measure. Column 3, again, shows that our results are largely unaffected by the inclusion of firm fixed effects. As before, the rest of Table **10** shows that the effect of social proximity to capital on reducing illiquidity is most significant for small firms as well as for firms with lower analyst coverage.

In summary, we find that firms' social proximity to capital is negatively associated with both effective spread and illiquidity. Additionally, these effects are strongest for small firms and firms with low analyst coverage. These results suggest 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 can jointly explain our findings, and in particular the heterogeneity of the effect 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. Nevertheless, we next provide additional evidence for our proposed explanation.

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 affect 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 explore 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 physically come 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 likely impaired. Consistent with this intepretation, the Wall Street Journal quoted a trader saying that *"'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, 2014). If our interpretation of the baseline correlation between social proximity to capital and liquidity were correct, we would thus expect that the liquidity of firms with high connectedness to East Coast-based investors fell 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 (NY, NJ, CT, DC, PA, DE, MD, VA, and WV), 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 (i.e., 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:

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

Motivated by the theoretical argument in Liu and Wang (2016), who show that spreads tend to widen when the number of market makers for an asset decreases, we focus on firm i's percentage effective spread on day t, given by $Spread_{i,t}$, as the dependent variable. $I(Sandy)_t$ is an indicator variable that equals to one during the Sandy period, defined as October 22, 2012, to November 7, 2012.[29] 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:

$$Affected\ Ratio_i = \frac{\sum_{k \in Mid-Atlantic} AUM_k \times Social\ Connectedness_{i,k}}{\sum_j AUM_j \times Social\ Connectedness_{i,j}}. \tag{10}$$

[29]We chose this date range to capture the period when travel in the Tri-State region was substantially impacted by Sandy. See https://www.cnn.com/2013/07/13/world/americas/hurricane-sandy-fast-facts/index.html for details.

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 10, each interacted with $I(Sandy)_t$.[30]

[Insert Table 11 near here]

Our results are reported in Table 11. 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 5.9% (0.248$\times$0.237) 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 1.1). As shown in column 4, we find that, despite losing close to one-half of our observations, we still find a positive and significant coefficient on the interaction of the Sandy indicator and the *Affected Ratio*.

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 11, 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.

[30]We include Log Assets, Log M/B, Profitability, Log Volatility, Momentum, Price, S&P, Age, Advertising, Exchange, Log Physical Proximity, Population, High School, College, Agglomeration, Income, Employment, and *Affected Ownership Ratio*, the ratio of ownership by institutions located in the Mid-Atlantic Area to total institutional ownership.

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 would 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.[31]

We examine measures of investor performance both across and within institutions. In the across-institution test, 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 test, we compare the performance of the connected and non-connected holdings of the same institution.

3.1 Across-Institution Performance Comparison

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

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

This specification corresponds to that of column 6 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 then sort institutions into deciles of β_j.[32] 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 12. Columns 1-3 show the results for average excess returns, CAPM-adjusted returns, and Fama and French (2015) five-factor-model-adjusted returns, respectively.[33]

[31] 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) the overweighting of connected stocks resulting in portfolio under-diversification and therefore a deviation from the efficient portfolio.

[32] We use the cross-sectional estimates from June 2016 for all institution-quarters, as these propensities are highly persistent.

[33] 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

We do not find evidence that investors that exhibit greater propensity to overweight stocks from areas they are socially connected to outperform investors with lower propensity.

We also investigate whether institutional investors that invest more in socially connected counties are under-diversified and, therefore, bear more risk. 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 and do not find any significant difference between institutions in the top decile and institutions in the bottom decile of β_j.[34] Overall, there is no evidence that the propensity to hold socially connected stocks leads to a differential performance among institutional investors.

[Insert Table 12 near here]

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 many 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', 'medium', and 'high' connectedness portfolios based on the headquarters location of each stock (e.g., we group all stocks 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 **13**, we report average daily excess returns for each institution's sub-portfolios, as well as the daily risk-adjusted returns using the CAPM and Fama and French (2015) five-factor models.[35] Panel A shows no significant

j at end of the prior quarter and $Return_{j,t}$ is stock j's daily return.

[34] Table IA.5 shows that while there is some heterogeneity across Bushee (2001) types, these differences are not economically meaningful.

[35] Similar to Section 3.1, the sub-portfolio returns are based on the value-weighted daily returns, with weights determined by the value of a stock's holding as of the end of the prior quarter.

return difference between the 'low' and 'high' connectedness portfolios, suggesting that an institution's connected holdings do not significantly outperform its non-connected 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. Panel A of Table **13** also shows that 'high' connectedness portfolios held by institutions do not outperform stocks in high connectedness counties that are not held by the institution.

We also explore whether there exist performance differences between high-connectedness and low-connectedness portfolios among those institutions that are most susceptible to invest in stocks they are connected to (i.e., institutions in the top β_j decile, institutions with small AUM, and dedicated institutions). In Panel B of Table **13**, 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 alpha, or Fama-French 5-factor adjusted alpha, 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.

[Insert Table 13 near here]

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 Appendix Table IA.6), we find no evidence that institutional investors have 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 firms rather than by superior information.[36]

[36]Importantly, though, our evidence does not mean that investors are unable to obtain information advantages through other social networks (such an effect was documented in Cohen et al., 2008; Hong et al., 2005; Hong and Xu, 2019; Pool et al., 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 the network based on factors such as shared neighborhood or education institutions. The finding that this broader type of network has both economically large and statistically significant impacts on institutions' portfolio decisions and equilibrium asset prices complements the prior literature on the role of social networks on financial markets. Similarly, Da et al. (2020) show that air travel reduces the effects of geographical proximity as well as local investment bias.

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 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.

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. For example, social connections can facilitate trade flows between regions (as suggested by Bailey et al., 2020a). Quantifying the extent to which regional differences in economic and social outcomes are explained by the structure of social networks, and 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 pursue these and related questions.

References

Aghion, Philippe, John Van Reenen, and Luigi Zingales, 2013, Innovation and institutional ownership, *American economic review* 103, 277–304.

Allen, Linda, Lin Peng, and Yu Shan, 2020, Social networks and credit allocation on fintech lending platforms, *Available at SSRN* .

Altonji, Joseph G, Todd E Elder, and Christopher R Taber, 2005, Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, *Journal of political economy* 113, 151–184.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of financial markets* 5, 31–56.

Baik, Bok, Jun-Koo Kang, and Jin-Mo Kim, 2010, Local institutional investors, information asymmetries, and equity returns, *Journal of financial economics* 97, 81–106.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, 2018a, Social connectedness: Measurements, determinants, and effects, *Journal of Economic Perspectives* 32, 259–80.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, 2018b, The economic effects of social networks: Evidence from the housing market, *Journal of Political Economy* 126, 2224–2276.

Bailey, Michael, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel, 2019a, House price beliefs and mortgage leverage choice, *The Review of Economic Studies* 86, 2403–2452.

Bailey, Michael, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel, 2019b, Social connectedness in urban areas, *NBER Working Paper* .

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel, 2020a, International trade and social connectedness, *Working Paper* .

Bailey, Michael, Drew Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, and Johannes Stroebel, 2020b, Social distancing during a pandemic: The role of friends, *NBER Working Paper* .

Bailey, Michael, Drew M Johnston, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, 2019c, Peer effects in product adoption, Technical report, National Bureau of Economic Research.

Bailey, Michael, Theresa Kuchler, Dominic Russel, Bogdan State, and Johannes Stroebel, 2020c, Social connectedness in Europe, *Working Paper* .

Baum-Snow, Nathaniel, and Ronni Pavan, 2011, Understanding the city size wage gap, *The Review of economic studies* 79, 88–127.

Becker, Bo, Henrik Cronqvist, and Rüdiger Fahlenbrach, 2011, Estimating the effects of large shareholders using a geographic instrument, *JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS* 46, 907–942.

Bernile, Gennaro, Alok Kumar, and Johan Sulaeman, 2015, Home away from home: Geography of information and local investors, *The Review of Financial Studies* 28, 2009–2049.

Bernile, Gennaro, Alok Kumar, Johan Sulaeman, and Qin Wang, 2019, Has local informational advantage disappeared?, *Review of Financial Economics* 37, 38–60.

Blume, Marshall E, and Donald B Keim, 2012, Institutional investors and stock market liquidity: trends and relationships, *Available at SSRN 2147757* .

Bushee, Brian J, 1998, The influence of institutional investors on myopic r&d investment behavior, *Accounting review* 305–333.

Bushee, Brian J, 2001, Do institutional investors prefer near-term earnings over long-run value?, *Contemporary Accounting Research* 18, 207–246.

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

Chen, Henry, Paul Gompers, Anna Kovner, and Josh Lerner, 2010, Buy local? the geography of venture capital, *Journal of Urban Economics* 67, 90–102.

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

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

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008, The small world of investing: Board connections and mutual fund returns, *Journal of Political Economy* 116, 951–979.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2010, Sell-side school ties, *The Journal of Finance* 65, 1409–1437.

Cohen, Lauren, Umit G Gurun, and Christopher Malloy, 2017, Resident networks and corporate connections: Evidence from world war ii internment camps, *The Journal of Finance* 72, 207–248.

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

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, 2019, Ppmlhdfe: Fast poisson estimation with high-dimensional fixed effects, *arXiv preprint arXiv:1903.01690* .

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

Coval, Joshua D, and Tobias J Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of political Economy* 109, 811–841.

Da, Zhi, Umit G. Gurun, Bin Li, and Mitch Warachka, 2020, Investment in a smaller world: The implications of air travel for investors and firms, *Management Science* forthcoming.

Dougal, Casey, Christopher A Parsons, and Sheridan Titman, 2018, Urban vibrancy and firm value creation, *Working Paper* .

Duggan, Maeve, Shannon Greenwood, and Andrew Perrin, 2016, Social media update 2016. Pew Research Center.

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of financial economics* 116, 1–22.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *The Journal of Finance* 64, 2023–2052.

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

Gompers, Paul A, and Andrew Metrick, 2001, Institutional investors and equity prices, *The quarterly journal of Economics* 116, 229–259.

Green, T Clifton, and Russell Jame, 2013, Company name fluency, investor recognition, and firm value, *Journal of Financial Economics* 109, 813–834.

Hirshleifer, David, 2020, Presidential address: Social transmission bias in economics and finance, *The Journal of Finance* .

Hirshleifer, David, Lin Peng, and Wang Qiguang, 2019, Social interaction and market reaction to earnings news, Technical report, Working Paper.

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

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2004, Social interaction and stock-market participation, *The journal of finance* 59, 137–163.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2005, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, *The Journal of Finance* 60, 2801–2824.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2008, The only game in town: Stock-price consequences of local bias, *Journal of Financial Economics* 90, 20–37.

Hong, Harrison, Terence Lim, and Jeremy C Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *The Journal of Finance* 55, 265–295.

Hong, Harrison, and Jiangmin Xu, 2019, Inferring latent social networks from stock holdings, *Journal of Financial Economics* 131, 323–344.

Huberman, Gur, 2001, Familiarity breeds investment, *The Review of Financial Studies* 14, 659–680.

Kang, Jun-Koo, Rene 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, Markku, and Samuli Knüpfer, 2012, Peer performance and stock market entry, *Journal of Financial Economics* 104, 321–338.

Kuchler, Theresa, Dominic Russel, and Johannes Stroebel, 2020, The geographic spread of covid-19 correlates with structure of social networks as measured by facebook, Technical report, National Bureau of Economic Research.

Kuchler, Theresa, and Johannes Stroebel, 2020, Social finance, Technical report, National Bureau of Economic Research.

Lehavy, Reuven, and Richard G Sloan, 2008, Investor recognition and stock returns, *Review of Accounting Studies* 13, 327–361.

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

Massa, Massimo, and Andrei Simonov, 2006, Hedging, familiarity and portfolio choice, *The Review of Financial Studies* 19, 633–685.

Merton, Robert C, 1987, A simple model of capital market equilibrium with incomplete information, *The journal of finance* 42, 483–510.

Miller, Edward, 1977, Risk, uncertainty and divergence of opinion,, *The journal of finance* 32, 1151–1168.

Moretti, Enrico, 2012, *The new geography of jobs* (Houghton Mifflin Harcourt).

Newey, Whitney K, and Kenneth D West, 1994, Automatic lag selection in covariance matrix estimation, *The Review of Economic Studies* 61, 631–653.

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

Ouimet, Paige, and Geoffrey Tate, 2020, Learning from coworkers: Peer effects on individual investment decisions, *The Journal of Finance* 75, 133–172.

Parsons, Christopher A, Riccardo Sabbatucci, and Sheridan Titman, 2020, Geographic lead-lag effects, *The Review of Financial Studies* 33, 4721–4770.

Pool, Veronika K, Noah Stoffman, and Scott E Yonker, 2012, No place like home: Familiarity in mutual fund manager portfolio choice, *The Review of Financial Studies* 25, 2563–2599.

Pool, Veronika K, Noah Stoffman, and Scott E Yonker, 2015, The people in your neighborhood: Social interactions and mutual fund portfolios, *The Journal of Finance* 70, 2679–2732.

Rehbein, Oliver, Simon Rother, et al., 2020, Distance in bank lending: The role of social networks, Technical report, University of Bonn and University of Mannheim, Germany.

Rehse, Dominik, Ryan Riordan, Nico Rottke, and Joachim Zietz, 2019, The effects of uncertainty on market liquidity: Evidence from hurricane sandy, *Journal of Financial Economics* .

Rubin, Amir, 2007, Ownership level, ownership concentration and liquidity, *Journal of financial Markets* 10, 219–248.

Russolillo, Steven, 2014, Traders brace for ‘frankenstorm’, *Wall Street Journal* .

Scheinkman, Jose A, and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of political Economy* 111, 1183–1220.

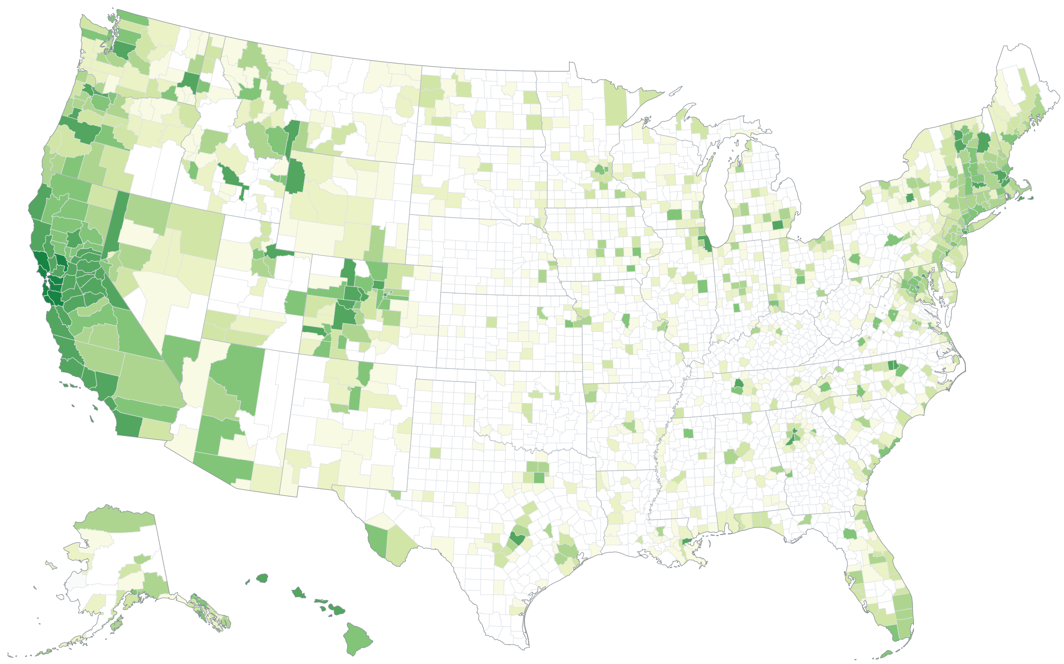
Seasholes, Mark S, and Ning Zhu, 2010, Individual investors and local bias, *The Journal of Finance* 65, 1987–2010.

Shue, Kelly, 2013, Executive networks and firm policies: Evidence from the random assignment of mba peers, *The Review of Financial Studies* 26, 1401–1442.

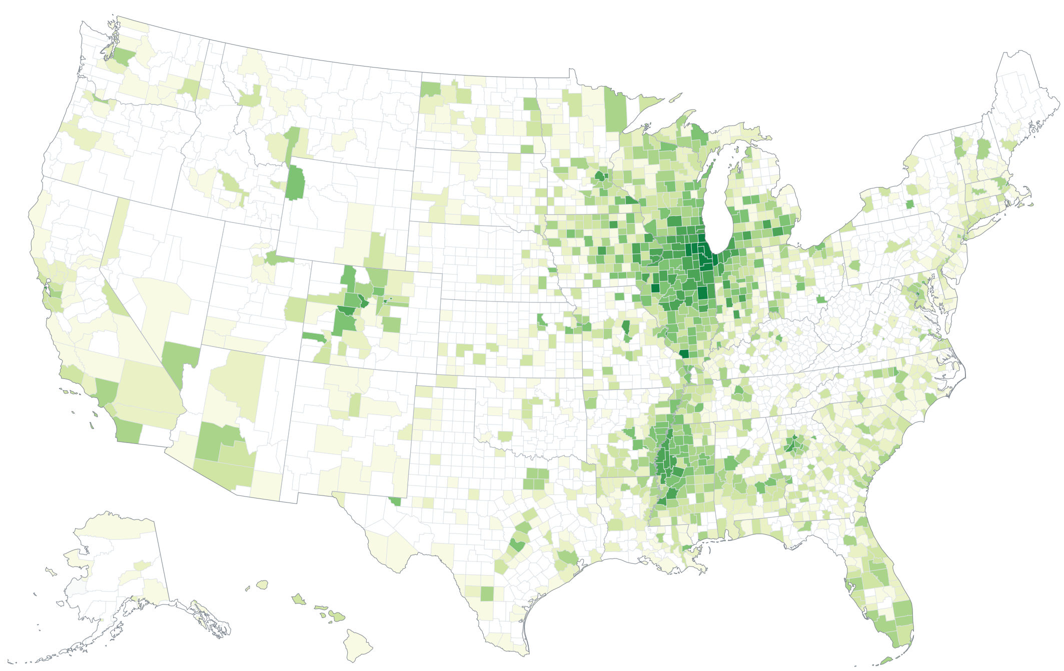
Silva, JMC Santos, and Silvana Tenreyro, 2006, The log of gravity, *The Review of Economics and statistics* 88, 641–658.

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.



Panel A: San Francisco County, CA



Panel B: Cook County, IL (Chicago)

Figure 2: **Binscatter Plot**

To produce these binned scatter plots, *Log Social Connectedness* as of June 2016 is sorted into 50 bins. For each bin, the conditional mean of *Log Social Connectedness* and conditional mean of the dependent variable, *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 include firm fixed effects, institution $\times$ industry fixed effects, same state fixed effects, and same county fixed effects. We further include 500-tile distance dummies as our distance control in the right panel.

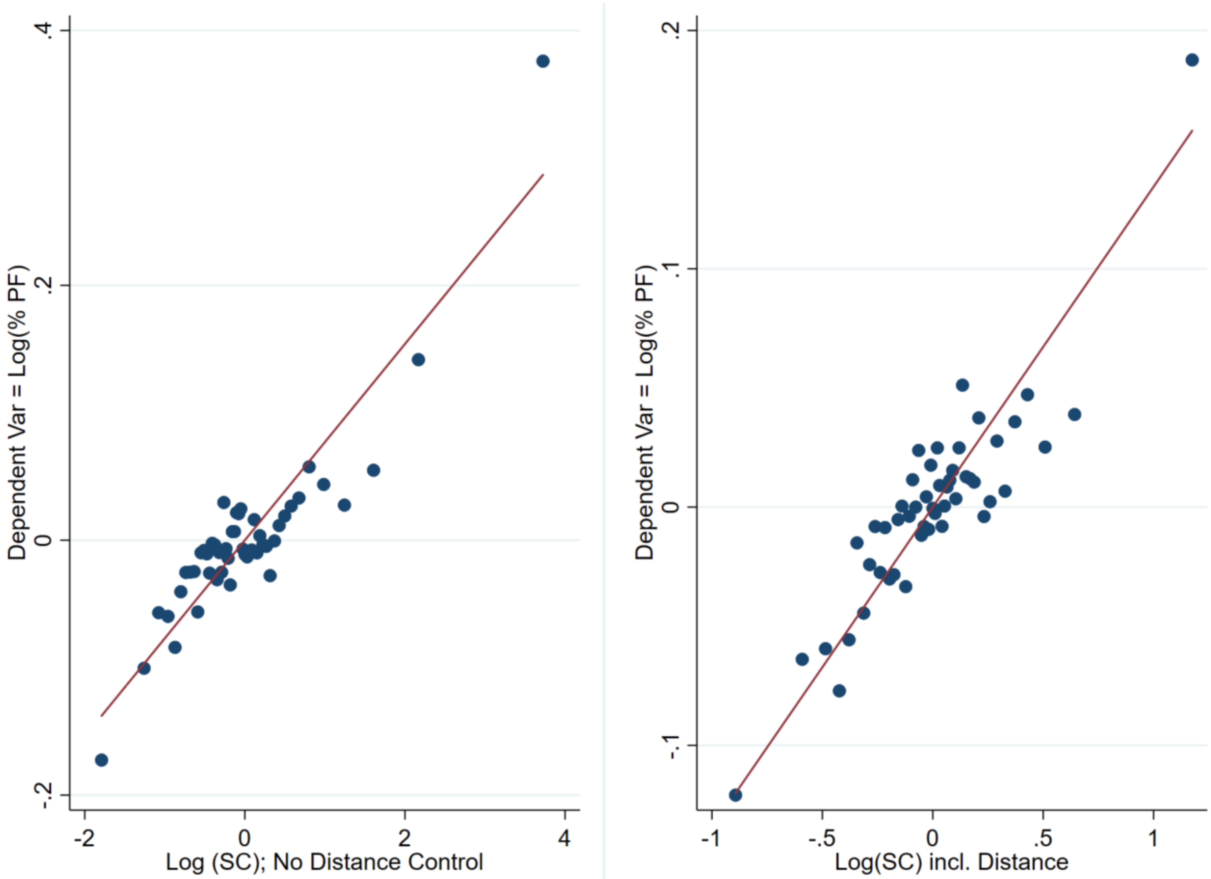


Figure 3: **Heat Map of Social Proximity to Capital**

This figure plots the 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 red have higher levels of *Social Proximity to Capital*, and regions in blue indicate lower levels of *Social Proximity to Capital*.

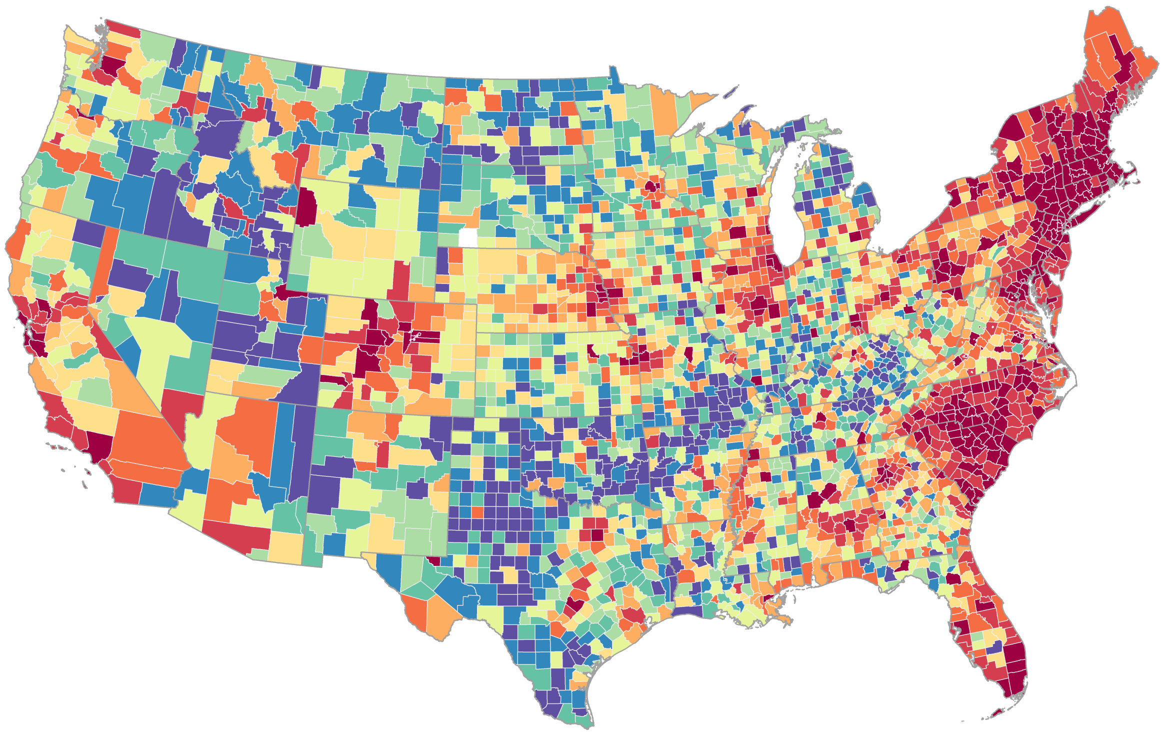


Table 1: Summary Statistics

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 allocated to a 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 assigned 0. Summary statistics for the firm-level variables are presented in Panel B. Our 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 is measured in millions of US dollars 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. *ILLIQ* is the Amihud (2002) illiquidity measure at the quarterly horizon (multiplied by 10^6). *Effective Spread* is the dollar-weighted percentage effective spread (multiplied by 10^3). *Tobin's Q* is the ratio of market value and replacement cost of assets. *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. Appendix A.1 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 MKT (formerly AMEX). We exclude penny stocks (Price$<$ \$5).

Panel A: Institution-Firm Pair Observations (as of June 2016)

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
Log Social Connectedness	6.06	1.29	4.45	4.71	5.84	7.53	8.40
Log Distance	6.52	1.37	3.96	5.09	6.82	7.80	7.84
% PF	0.04	0.50	0	0	0	0	0.01

Panel 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 ILLIQ	-5.38	3.16	-9.75	-9.06	-5.79	-1.07	1.06
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 Tobin's Q	0.44	0.56	-0.15	-0.06	0.30	1.22	1.55
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

Panel C: Institution Characteristics, by Institution Type

		AUM (Million USD) as of June 2016						
Institution Type	**N**	**MEAN**	**ST. DEV**	**P5**	**P10**	**MEDIAN**	**P90**	**P95**
Dedicated	75	4,215.57	16,124.16	52.61	85.05	565.09	6,982.40	8,345.77
Quasi-Indexer	1741	5,624.85	44,621.73	28.98	58.00	292.18	4,626.11	14,066.13
Transient	724	4,182.23	47,348.04	15.28	31.97	344.69	4,308.45	9,305.58
Not Identified	543	272.289	1,107.78	7.47	15.315	88.42	373.35	721.91

Table 2: Social Connectedness and Institutional Investment

This table explores the relationship between social connectedness and institutional investors' portfolio decisions. Estimates are obtained using Poisson Pseudo Maximum Likelihood (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 0. *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. We also consider *Social Connectedness Quintile* indicators as independent variables. *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. We consider Firm, Institution, Institution$\times$Industry, Distance 500-tile, 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 500-tile* indicators are 500 dummy variables indicating the quantile of the distance between the firm and the institution based on all firm-institution pairs as of June 2016. Appendix A.1 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. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Social Connectedness	0.189***	0.189***		0.253***	0.321***	0.314***	
	(13.21)	(14.97)		(10.65)	(12.89)	(11.78)	
Log Distance			-0.107***	0.054***			
			(-9.93)	(2.94)			
Social Connectedness Quintile = 2 (Low)							0.041*
							(1.79)
Social Connectedness Quintile = 3							0.152***
							(5.07)
Social Connectedness Quintile = 4							0.270***
							(6.46)
Social Connectedness Quintile = 5 (High)							0.466***
							(8.02)
Firm FE	YES	YES	YES	YES	YES	YES	YES
Institution FE	YES	NO	NO	NO	NO	NO	NO
Institution $\times$ Industry FE	NO	YES	YES	YES	YES	YES	YES
Distance 500-tile FE	NO	NO	NO	NO	YES	YES	YES
Same State FE	NO	NO	NO	NO	NO	YES	YES
Same County FE	NO	NO	NO	NO	NO	YES	YES
N	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060
Pseudo R^2	0.320	0.506	0.504	0.506	0.508	0.508	0.507

Table 3: County Similarity and Economic Relationship

This table shows the results on whether the relationship of social connectedness and institutional investors' holdings are explained by county similarities or economic linkages. The dependent variable is % *PF* defined as the percentage of AUM allocated to a stock. If an institution does not report holdings in a given firm, % *PF* is assigned 0. *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 include Firm × Industry, Distance 500-tile, 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 500-tile* indicators are 500 dummy variables indicating the quantile of the distance between the firm and the institution. In column 1, we control for county differences, defined as $|C_i - C_j|$, where C represents a county characteristic variable. We consider the following county characteristics: log county population, population density code, median age, income per capita, population share with at least high school education, population share with at least college education, employment, share of immigrants, share of unmarried, and Democrat voter share. We also control for industry composition similarity, measured by the cosine similarity between firm and institution counties. In columns 2 and 3, we include controls for state citations in 10-K filings, where column 2 includes *Citation Share* and column 3 includes a state citation indicator (*I(Cited)*). We include state revenue controls in columns 4 and 5, where column 4 includes a control for *Revenue Share* and column 5 includes *I(Revenue State)*, an indicator variable which equals 1 if the institution is located in the firm's revenue state. Refer to Appendix A.1 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. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)
Log Social Connectedness	0.300***	0.309***	0.306***	0.304***	0.309***
	(12.15)	(11.66)	(11.50)	(11.47)	(11.65)
Citation Share		0.787***			
		(5.08)			
I (Cited)			0.079***		
			(4.00)		
Revenue Share				0.771*	
				(1.69)	
I (Revenue State)					0.162
					(1.17)
Firm FE	YES	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES	YES
Same county FE	YES	YES	YES	YES	YES
Same state FE	YES	YES	YES	YES	YES
Distance 500-tile FE	YES	YES	YES	YES	YES
County Characteristic Difference	YES	NO	NO	NO	NO
N	8,219,537	7,997,302	7,997,302	8,311,768	8,694,060
R^2	0.509	0.521	0.521	0.510	0.509

Table 4: Social Connectedness and Investment: by Institution Characteristics

This table shows the result 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. In Panel A, 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). Institutions without a Bushee type are dropped from this sample. In Panel B, we report results on the heterogeneity across institution size terciles. 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 *Log Distance*, defined as *log(1+Distance)*, as a control. We consider Firm, Institution×Industry, Distance 500-tile, 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 500-tile indicators indicate the quantile of the distance between the firm and the institution based on all firm-institution pairs as of June 2016. Refer to Appendix A.1 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. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

Panel A: Heterogeneity across Bushee Institution Types

	(1)	**(2)**	**(3)**	**(4)**
Transient × Log Social Connectedness	0.116***	0.128***	0.177***	0.175***
	(6.78)	(4.19)	(5.11)	(4.77)
Quasi-Indexer × Log Social Connectedness	0.206***	0.286***	0.334***	0.331***
	(15.46)	(11.64)	(12.34)	(11.50)
Dedicated × Log Social Connectedness	0.332***	0.345**	0.634***	0.714***
	(4.86)	(2.45)	(4.43)	(4.82)
Institution Type × Firm FE	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES
Institution Type × Log Distance	NO	YES	NO	NO
Institution Type × Distance 500-tile FE	NO	NO	YES	YES
Institution Type × Same County FE	NO	NO	NO	YES
Institution Type × Same State FE	NO	NO	NO	YES
F Test (No Heterogeneity)	27.9***	20.23***	19.89***	21.09***
N	7,162,800	7,162,800	7,162,800	7,162,800
Pseudo R^2	0.536	0.537	0.544	0.544

Table 4: (Continued)

Panel B: Heterogeneity across Institution AUM Groups

	(1)	(2)	(3)	(4)
Small AUM × Log Social Connectedness	0.253***	0.300***	0.357***	0.344***
	(14.05)	(8.35)	(9.20)	(8.22)
Mid AUM × Log Social Connectedness	0.183***	0.256***	0.322***	0.309***
	(10.64)	(7.71)	(8.75)	(7.86)
Large AUM × Log Social Connectedness	0.120***	0.195***	0.263***	0.267***
	(7.09)	(6.34)	(7.20)	(6.93)
Institution Type × Firm FE	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES
Institution Type × Log Distance	NO	YES	NO	NO
Institution Type × Distance 500-tile FE	NO	NO	YES	YES
Institution Type × Same County FE	NO	NO	NO	YES
Institution Type × Same State FE	NO	NO	NO	YES
F Test (Small = Large)	38.18***	6.71***	3.50**	2.14
N	8,694,060	8,694,060	8,694,060	8,694,060
Pseudo R^2	0.522	0.522	0.527	0.527

Table 5: Social Connectedness and Investment: by Firm Characteristics

This table shows the result 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. In Panel A, we report results on the heterogeneity across firm size. We interact *Log Social Connectedness* with dummy variables based on firm size terciles. In Panel B, we report results on the heterogeneity across firms' analyst coverage. We interact *Log Social Connectedness* with dummy variables based on firm analyst coverage terciles. *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 *Log Distance*, defined as *log(1+Distance)*, as a control. We consider Firm, Institution×Industry, Distance 500-tile, 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 500-tile indicators indicate the quantile of the distance between the firm and the institution based on all firm-institution pairs each quarter. Refer to Appendix A.1 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. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

Panel A: Heterogeneity across Firm Size Groups

	(1)	(2)	(3)	(4)
Small Cap × Log Social Connectedness	0.351***	0.421***	0.536***	0.553***
	(7.08)	(4.34)	(6.24)	(6.13)
Mid Cap × Log Social Connectedness	0.327***	0.483***	0.539***	0.548***
	(13.11)	(10.43)	(10.45)	(10.22)
Large Cap × Log Social Connectedness	0.174***	0.224***	0.276***	0.268***
	(13.10)	(8.99)	(10.65)	(9.71)
Firm FE	YES	YES	YES	YES
Firm Type × Institution × Industry FE	YES	YES	YES	YES
Firm Type × Log Distance	NO	YES	NO	NO
Firm Type × Distance 500-tile FE	NO	NO	YES	YES
Firm Type × Same County FE	NO	NO	NO	YES
Firm Type × Same State FE	NO	NO	NO	YES
F Test (Small = Large)	37.61***	26.05***	25.48***	26.84***
N	8,694,060	8,694,060	8,694,060	8,694,030
Pseudo R^2	0.565	0.565	0.570	0.570

Table 5: (Continued)

Panel B: Heterogeneity by Analyst Coverage Groups

	(1)	(2)	(3)	(4)
Low Coverage × Log Social Connectedness	0.393***	0.532***	0.681***	0.737***
	(9.61)	(7.94)	(9.47)	(9.37)
Mid Coverage × Log Social Connectedness	0.290***	0.429***	0.447***	0.442***
	(13.84)	(11.32)	(9.79)	(9.20)
High Coverage × Log Social Connectedness	0.162***	0.214***	0.265***	0.253***
	(12.13)	(8.60)	(10.40)	(9.37)
Firm FE	YES	YES	YES	YES
Firm Size Tercile × Institution× Industry FE	YES	YES	YES	YES
Firm Type × Log Distance	NO	YES	NO	NO
Firm Type × Distance 500-tile FE	NO	NO	YES	YES
Firm Type × Same County FE	NO	NO	NO	YES
Firm Type × Same State FE	NO	NO	NO	YES
F Test (Low=High)	47.01***	34.98***	35.93***	39.97***
N	8,694,060	8,694,060	8,694,060	8,694,060
Pseudo R^2	0.577	0.577	0.583	0.583

Table 6: Identification Using Panel Data

This table shows the results on how social connectedness affects institutional investors' portfolio decisions using a panel data. The sample used in columns 1 to 4 represents institutional holdings from June 2007 to December 2016. We eliminate firm-quarter observations after a firm's first 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)*. *I(No Recent HQ Move)* is an indicator variable that equals one if the firm has not moved its headquarters in the previous 3 years. We consider Firm × Quarter, Institution × Quarter, Institution × Industry, Firm × Institution, Distance 500-tile, 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 500-tile indicators that indicate the quantile of the distance between the firm and the institution based on all firm-institution pairs as of June 2016. Refer to Appendix A.1 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. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)
Log Social Connectedness	0.180***	0.286***	0.276***	0.101***	0.073**
	(15.86)	(12.48)	(11.21)	(3.37)	(2.42)
Log Social Connectedness × I(No Recent HQ Move)					0.022**
					(2.00)
Firm FE × Quarter	YES	YES	YES	YES	YES
Institution × Quarter	YES	YES	YES	YES	YES
Institution × Industry FE	NO	YES	YES	NO	NO
Institution × Firm FE	NO	NO	NO	YES	YES
Distance 500-tile 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	0.883

Table 7: Social Connectedness and Mutual Fund Investment, with Fund Manager Characteristics

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 results 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 management team has an MBA degree. We consider Firm, Fund×Industry, Firm×Style, Distance 500-tile, 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 500-tile fixed effects indicate the quantile of the distance between the firm and the fund based on all fund-firm pairs as of June 2016. Refer to Appendix A.1 for detailed variable definitions. Industry classification is based on Fama-French 48 industries. Fund style is based on Lipper fund classification. In columns 2 to 4, 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. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

	(1)	(2)		(3)		(4)		(5)
	Active	**Index**		**Active**		**Active**		**Active**
Log Social Connectedness	0.116***	0.018						
	(4.39)	(0.28)						
			× Young	0.091**	**× Male only**	0.094***	**× MBA Minority**	0.122***
				(2.54)		(3.13)		(3.12)
Log Social Connectedness			**× Old**	0.141***	**× Female or Both Gender**	0.184***	**× MBA Majority**	0.116***
				(3.93)		(3.85)		(3.35)
Fund × Industry FE	YES	YES		YES		YES		YES
Firm FE	NO	YES		NO		NO		NO
Firm × Style FE	YES	NO		YES		YES		YES
Distance 500-tile 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-STAT				2.58		1.00		0.01
N	2,155,060	529,070		2,155,060		2,155,060		2,155,060
Pseudo R^2	0.584	0.702		0.622		0.623		0.614

Table 8: Firms' Social Proximity to Capital and Institutional Ownership

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* $\times$ *Social Connectedness*. *County AUM* is measured in millions of US dollars and the summation is taken across all U.S. counties. Additional firm controls include *Log Assets*, *Log M/B*, *Volatility*, *Momentum*, *Turnover*, *Lag Price*, *R&D*, *Yield*, *S&P*, *Age*, *Advertising*, and *Exchange* dummies. Additional county controls include *Log Physical Proximity to Capital*, *High School*, *College*, *Agglomeration*, *Population*, *Income*, and *Employment*. Refer to Appendix A.1 for detailed variable definitions. The regressions also include Quarter, State, Industry, Quarter $\times$ 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. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

	(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 $\times$ Log Social Proximity to Capital				2.137**	2.074**	2.006**	1.925**
				(2.35)	(2.30)	(2.19)	(2.12)
Mid $\times$ Log Social Proximity to Capital				0.235	0.099	0.730	0.688
				(0.23)	(0.10)	(0.80)	(0.76)
High $\times$ 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 $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
Industry FE	YES	NO	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
State FE	YES	YES	NO	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE
Quarter $\times$ Industry FE	NO	YES	NO	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE
Firm FE	NO	NO	YES	NO	NO	NO	NO
F-stat (Null: 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

Table 9: Social Proximity to Capital and Firm Value

This table presents the panel regression result 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 variables are *Log M/B* in Panel A and *Log Tobin's Q* in Panel B. The main independent variable is *Log Social Proximity to Capital*, where *Social Proximity to Capital* is defined as $\sum$ *County AUM* $\times$ *Social Connectedness*, where *County AUM* is measured in millions of US 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*, *Agglomeration*, *Population*, *Income*, and *Employment*. Refer to Appendix A.1 for detailed variable definitions. The regressions also include Quarter, State, Industry, Quarter $\times$ 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. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

Panel A: Social Proximity to Capital and Market-to-Book

	(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 $\times$ Log Social Proximity to Capital				0.076***	0.079***	0.094***	0.094***
				(3.16)	(3.28)	(3.52)	(3.58)
Mid $\times$ Log Social Proximity to Capital				0.048*	0.050**	0.097***	0.097***
				(2.02)	(2.15)	(3.58)	(3.63)
High $\times$ 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 $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
Industry FE	YES	NO	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
State FE	YES	YES	NO	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE
Quarter $\times$ Industry FE	NO	YES	NO	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE
Firm FE	NO	NO	YES	NO	NO	NO	NO
F-stat (Null: 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

Table 9: (Continued)

Panel B: Social Proximity to Capital and Tobin's Q

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Whole Sample			Split by Size		Split by Analyst Coverage	
Log Social Proximity to Capital	0.059***	0.058***	0.059**				
	(3.69)	(3.62)	(2.38)				
Low × Log Social Proximity to Capital				0.037***	0.037**	0.054***	0.053***
				(2.72)	(2.71)	(3.54)	(3.53)
Mid × Log Social Proximity to Capital				0.032**	0.032**	0.047***	0.046***
				(2.14)	(2.18)	(2.95)	(2.90)
High × Log Social Proximity to Capital				0.036*	0.032*	0.032*	0.027
				(1.99)	(1.81)	(1.75)	(1.46)
Firm Controls	YES	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES	YES × TERCILE	NO	YES × TERCILE	NO
Industry FE	YES	NO	NO	YES × TERCILE	NO	YES × TERCILE	NO
State FE	YES	YES	NO	YES × TERCILE	YES × TERCILE	YES × TERCILE	YES × TERCILE
Quarter × Industry FE	NO	YES	NO	NO	YES × TERCILE	NO	YES × TERCILE
Firm FE	NO	NO	YES	NO	NO	NO	NO
F-stat (Null: Low=High)				0.010	0.098	1.404	2.134
N	96,761	96,761	96,761	96,761	96,761	96,761	96,761
R^2	0.332	0.353	0.839	0.552	0.577	0.435	0.469

Table 10: Social Proximity to Capital and Stock Liquidity

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* in Panel A, where *Effective Spread* is the quarterly average of daily percentage effective spread. The dependent variable is *Log ILLIQ* in Panel B, where *ILLIQ* is the quarterly average of Amihud (2002) illiquidity measure defined as daily absolute return, divided by dollar volume. The main independent variable is *Log Social Proximity to Capital*, where *Social Proximity to Capital* is defined as $\sum$ *County AUM* $\times$ *Social Connectedness*. *County AUM* is measured in millions of US dollars and the summation is taken across all U.S. counties. Additional firm controls include *Log Assets*, *Profitability*, *Log M/B*, *Log Volatility*, *Momentum*, *Lag Price*, *S&P*, *Age*, *Advertising*, and *Exchange* dummies. Additional county controls include *Log Physical Proximity to Capital*, *High School*, *College*, *Agglomeration*, *Population*, *Income*, and *Employment*. Refer to Appendix A.1 for detailed variable definitions. The regressions also include Quarter, State, Industry, Quarter $\times$ 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. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

Panel A: Social Proximity to Capital and Effective Spread

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Whole Sample			Split by Size		Split by Analyst Coverage	
Social Proximity to Capital	-0.094***	-0.092***	-0.068**				
	(-4.50)	(-4.44)	(-2.49)				
Low $\times$ Social Proximity to Capital				-0.081***	-0.080***	-0.093***	-0.089***
				(-3.69)	(-3.66)	(-4.24)	(-4.04)
Mid $\times$ Social Proximity to Capital				-0.056***	-0.055***	-0.075***	-0.073***
				(-3.14)	(-3.08)	(-4.11)	(-4.03)
High $\times$ 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 $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
Industry FE	YES	NO	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
State FE	YES	YES	NO	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE
Quarter $\times$ Industry FE	NO	YES	NO	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE
Firm FE	NO	NO	YES	NO	NO	NO	NO
F-stat (Null: 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

Table 10: (Continued)

Panel B: Social Proximity to Capital and Illiquidity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Whole Sample			Split by Size		Split by Analyst Coverage	
Social Proximity to Capital	-0.292***	-0.285***	-0.210***				
	(-4.73)	(-4.66)	(-2.99)				
Low × Social Proximity to Capital				-0.253***	-0.250***	-0.271***	-0.261***
				(-3.49)	(-3.45)	(-3.88)	(-3.74)
Mid × Social Proximity to Capital				-0.115**	-0.116**	-0.179***	-0.175***
				(-2.42)	(-2.42)	(-3.83)	(-3.77)
High × Social Proximity to Capital				-0.035	-0.033	-0.096**	-0.092**
				(-0.73)	(-0.68)	(-2.13)	(-2.05)
Firm Controls	YES	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES	YES × TERCILE	NO	YES × TERCILE	NO
Industry FE	YES	NO	NO	YES × TERCILE	NO	YES × TERCILE	NO
State FE	YES	YES	NO	YES × TERCILE	YES × TERCILE	YES × TERCILE	YES × TERCILE
Quarter × Industry FE	NO	YES	NO	NO	YES × TERCILE	NO	YES × TERCILE
Firm FE	NO	NO	YES	NO	NO	NO	NO
F-stat (Null: Low=High)				11.099***	11.208***	6.515**	6.096**
N	101,008	101,008	101,008	101,008	101,008	101,008	101,008
R^2	0.795	0.801	0.945	0.850	0.856	0.846	0.853

Table 11: The Effect of Social Proximity to Capital during Hurricane Sandy

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 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 **10** and the *Affected Ownership Ratio*, each interacted with *I(Sandy)*. Refer to Appendix A.1 for detailed variable definitions. We also control for firm fixed effects, state fixed effects, 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 3 mentioned states in 10-K are excluded in column 2. Column 3 shows heterogeneity of the result across firm size terciles and column 4 across analyst coverage terciles. We report t-statistics based on robust standard errors clustered by firm and week in the parentheses below. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)
	Whole Sample		Split by Size	Split by Analyst Coverage
I(Sandy) × Affected Ratio	0.248***	0.329***		
	(3.39)	(2.92)		
Low × I(Sandy) × Affected Ratio			0.420**	0.543***
			(2.29)	(3.18)
Med × I(Sandy) × Affected Ratio			0.113	0.021
			(0.75)	(0.19)
High × I(Sandy) × Affected Ratio			0.151	0.243***
			(1.35)	(2.88)
Control × Sandy FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
State FE	YES	YES	YES × TERCILE	YES × TERCILE
Day × Industry FE	YES	YES	YES × TERCILE	YES × TERCILE
Sample Exclusion	EAST	EAST & Exposed	EAST	EAST
F-stat (Low=High)			1.060	2.010
N	545,295	295,384	545,295	545,295
R^2	0.766	0.781	0.785	0.784

Table 12: Portfolio Social Connectedness and Performance

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:

$$\%PF_{i,j} = \exp[\beta_{SC,i} \textit{Log Social Connectedness}_{i,j} + \beta_2 \textit{Same County}_{i,j} + \beta_3 \textit{Same State}_{i,j} + \textit{Firm FE} + \textit{Institution} \times \textit{Industry FE} + \textit{Distance 500-tile FE}] \cdot \epsilon_{i,j}.$$

We sort institutional investors into deciles based on their propensity to hold socially connected stocks ($\beta_{SC,i}$). Portfolios are rebalanced at the end of each quarter using institutions' previous 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 5-Factor adjusted returns. Columns 4 to 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 4 to 6 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. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

	Return			**σ(Return)**			**Sharpe Ratio**		
	(1) Excess	**(2) CAPM**	**(3) FF5**	**(4) Excess**	**(5) CAPM**	**(6) FF5**	**(7) Excess**	**(8) CAPM**	**(9) 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)

Table 13: Performance of Socially Connected Holdings

This table reports the daily returns (in %) of socially connected holdings within institutions' portfolio 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, median, 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 in panel A. We report the excess returns, and CAPM, and Fama-French 5-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 but are not part of the institutional holding and the return difference between the high social connectedness stocks held and not held by institutions. In panel B, we report the return difference between the high connectedness and low connectedness holding returns and the return difference between high connectedness holding and high connectedness stocks institutions do not hold for three institution subgroup, including those with high social connectedness beta (top β_{SC} decile), low AUM (bottom decile), and dedicated institutions. Newey and West (1994) adjusted t-statistics are reported in parentheses. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

Panel A: Social Connectedness and Holding Performance

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

Table 13: (Continued)

Panel B: Performance of Socially Connected Holdings in Subgroups

Subgroup	(1) Excess	(2) CAPM	(3) FF5	(4) Excess	(5) CAPM	(6) FF5
	High (Held) - Low (Held)			High (Held) - High (Not Held)		
High β_{SC}	0.003	-0.000	0.004	0.004	0.006	0.004
	(0.53)	(-0.04)	(0.82)	(0.99)	(1.64)	(1.34)
Low AUM	-0.002	-0.005	-0.000	-0.003	0.000	-0.000
	(-0.34)	(-0.77)	(-0.01)	(-0.53)	(0.08)	(-0.02)
Dedicated	0.015	0.009	0.013	0.015*	0.011	0.013**
	(1.61)	(1.02)	(1.38)	(1.78)	(1.46)	(2.39)
N	2,456	2,456	2,456	2,456	2,456	2,456

Appendix A.1 Variable List

Variable	Definition
Panel 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, and Democratic vote share. Density is a 1-6 code obtained from USDA. Voting data are obtained from townhall.com. The other county-level data are obtained from 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 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 non-zero revenue from the institution state. Firms' state revenues are obtained from Compustat Segement data
Panel 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)
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 WRDS Intraday Indicator database

Appendix A.1: (Continued)

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 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 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 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 6 months.
Share Turnover (Turnover)	Average trading volume scaled by total shares outstanding over the past 6 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 S&P500 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 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 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 Compustat Annual database and require at least a four-month lag from the fiscal year end.
Sales Growth	Sales growth is measured over the past three years. If less than three years of sales data is available, sales growth is estimated using all available data. Missing values are set 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.
Asset Turnover	Net sales over book value of total assets. We obtain the most recent data from 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 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 Compustat Annual database and require at least a four-month lag from the fiscal year end.
Size	Market capitalization (in millions) at the quarter end.
# Analyst	Number of one year-ahead analyst estimates at the end of the quarter.

Appendix A.1: (Continued)

Panel 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.
Panel D: County-Level Variables	
Social Proximity to Capital	County j's social proximity to capital is defined as $log(\sum(County\ AUM_i \times Social\ Connectedness_{i,j}))$, where county AUM is the sum of AUM of all the institutions located in a given county.
Physical Proximity to Capital	County j's physical proximity to capital is defined as $log(\sum(\frac{County\ AUM_i}{1+Distance_{i,j}}))$, where county AUM is the sum of AUM of all the institutions located in a given county.
Population	Log county population. Missing values are replaced by state average. Obtained from American Community Survey.
High School Attainment Dummy (High School)	Equals 1 if percent of adults in a county that obtain a high school education is greater than the sample median. Missing values are replaced by 0. Obtained from American Community Survey.
College School Attainment Dummy (College)	Equals 1 if percent of adults in a county that obtain a college education is greater than the sample average. Missing values are replaced by 0. Obtained from American Community Survey.
Income Per Capita (Income)	County-level income per capita. Missing values are replaced by state average. Obtained from American Community Survey.
Employment Rate (Employment)	County-level employment rate. Missing values are replaced by state average. Obtained from 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 American Community Survey.
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.

Social Proximity to Capital: Implications for Investors and Firms

Internet Appendix

We present the following information in this internet appendix:

- Table IA.1: Additional summary statistics
- Table IA.2: Robustness Tests for Regressions
- Table IA.3: Two-Dimensional Heterogeneity Based on Firm Size and Analyst Coverage
- Table IA.4: Social Proximity to Capital, Physical Proximity to Capital, and Firm Outcomes
- Table IA.5: Portfolio Social Connectedness and Performance by Institution Types
- Table IA.6: Performance of Socially Connected Holdings: Local vs Non-Local Firms
- Figure IA.1: Coefficients of 500-Tile Distance Indicators

Table IA.1: Additional Summary Statistics

This table reports additional summary statistics for our key variables. Summary statistics for the firm-institution level variables are presented in Panel A. Summary statistics for firm-level panels from 2007 to 2016 are presented in Panel B. Summary statistics of variables used in Hurricane Sandy analysis are presented in Panel C. Summary statistics of variables used in mutual fund analyses are presented in Panel D. Refer to Appendix A.1 for detailed variable definitions.

Panel A: Summary Statistics for Institution-Firm Pairs as of Jun 2016

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
Same State	0.02	0.13	0	0	0	0	0
Same County	0.06	0.24	0	0	0	0	1
|Diff. in Median Age|	3.22	2.77	0.10	0.40	2.60	6.90	8.60
|Diff. in Income per Capita|	62.89	51.05	2.17	6.40	48.79	148.83	159.40
|Diff. in % Bachelor|	13.53	9.72	0.70	1.80	11.90	28.50	30.30
|Diff. in % High School|	4.51	3.70	0.10	0.50	3.70	9.30	12.00
|Diff. in % Immigrants|	6.05	4.32	0.27	0.83	5.38	12.47	13.83
|Diff. in % Employment|	1.45	1.12	0.08	0.21	1.21	2.98	3.56
|Diff. in % Unmarried|	8.74	6.57	0.43	1.08	7.50	18.29	20.89
|Diff. in % Democrat|	19.44	14.18	1.19	2.88	16.91	39.13	46.48
|Diff. in Population|	1.07	0.86	0.04	0.14	0.87	2.31	2.75
|Diff. in Population Density|	0.58	1.09	0	0	0	2	3
Industry Similarity	0.94	0.05	0.84	0.87	0.95	0.98	0.99
I (Metro Pair)	0.94	0.24	0	1	1	1	1
I (Large Metro Pair)	0.63	0.48	0	0	1	1	1
I (Top 100 Populous County Pair)	0.51	0.50	0	0	1	1	1
Citation Share	0.07	0.17	0	0	0.01	0.21	0.36
I (Cited)	0.53	0.50	0	0	1	1	1
Revenue Share	0.00	0.03	0	0	0	0	0
I (Revenue State)	0.01	0.08	0	0	0	0	0

Table IA.1: (Continued)

Panel B: Summary Statistics for Firm-level Variables from 2007 to 2016

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
Log Total Assets	7.07	1.89	4.15	4.78	6.96	9.55	10.37
Return Volatility	10.56	8.42	3.03	3.90	8.76	18.82	23.52
Share Turnover	0.19	0.22	0.02	0.03	0.14	0.39	0.51
Share Price	43.61	1129.11	6.11	7.32	20.89	63.22	83.77
Sales Growth	0.14	1.82	-0.15	-0.08	0.05	0.31	0.45
Asset Turnover	0.86	0.94	0.05	0.06	0.69	1.82	2.35
Book Leverage	0.57	0.29	0.14	0.20	0.56	0.91	0.93
Payout	0.73	24.28	-0.23	0.00	0.28	1.63	2.59
S&P 500 Dummy	0.157	0.364	0	0	0	1	1
Momentum	14.20	64.87	-46.81	-34.21	7.14	61.02	91.37
Firm Age	240.33	212.38	13	27	188	511	633
Dividend Yield	0.02	0.05	0	0	0	0.04	0.05
R&D Expense	1.26	80.20	0	0	0	0.15	0.25
Advertising Expense	0.01	0.08	0	0	0	0.03	0.05
Profitability	0.09	0.17	-0.09	0.01	0.10	0.22	0.28
College	0.48	0.50	0	0	0	1	1
High School	0.48	0.50	0	0	0	1	1
Income per Capita	152.07	44.05	101.44	109.27	139.52	209.97	259.37
Employment	92.39	2.54	87.77	88.83	92.83	95.31	95.87
Agglomeration	8.28	1.45	5.29	6.44	8.49	9.91	10.27
Population	13.72	1.03	11.86	12.28	13.74	15.16	15.47

Table IA.1: (Continued)

Panel C: Daily Panel around Hurricane Sandy from Jan 2012 to Jul 2013

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
Affected Capital Ratio	0.33	0.10	0.14	0.18	0.36	0.42	0.45
Affected Ownership Ratio	0.45	0.15	0.19	0.28	0.44	0.63	0.71

Panel D: Summary Statistics for Active Mutual Fund-Firm Pair Variables as of Jun 2016

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
% PF	0.04	0.30	0	0	0	0	0
Log Social Connectedness	6.23	1.30	4.60	4.87	6.00	7.83	8.86
Same County	0.02	0.14	0	0	0	0	0
Same State	0.06	0.24	0	0	0	0	1
Med. Age	50.37	8.28	37.96	40.33	49.46	60.75	66.25
% Female	0.20	0.40	0	0	0	1	1
% MBA	0.51	0.39	0	0	0.50	1	1

Panel E: Summary Statistics for Active Index Fund-Firm Pair Variables as of Jun 2016

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
% PF	0.04	0.43	0	0	0	0	0
Log Social Connectedness	6.20	1.31	4.62	4.87	5.97	7.97	8.86
Same County	0.01	0.12	0	0	0	0	0
Same State	0.08	0.28	0	0	0	0	1

Table IA.2: Robustness Tests for Pairwise Regressions

This table exhibits the robustness tests for pairwise regressions in Table 2. In Panel A, we present robustness tests for cross-sectional pairwise tests, based on institutional holdings as of June 2016. The dependent variable is % *PF* in columns 1 to 4, defined as the percentage of AUM allocated to a stock. In column 1, we exclude firms and institutions headquartered in Tri-State (NY, NJ, CT) area and California. The dependent variable in column 2 is % *PF* winsorizing the top 1%. In column 3, we calculate institutions and firms to county levels. In column 5, we replace the continuous dependent variable with a dummy equal to one if % *PF* > 0 and zero otherwise. *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)*. The first three regressions are estimated using Poisson Pseudo Maximum Likelihood (PPML) and the results in columns 4 and 5 are estimated using OLS. Panel B provides further robustness tests at the pairwise level using PPML specifications. Column 1 includes three indicators (Populous County Pair FE), controlling for whether both the institution and the firm counties are populous areas (i.e., whether both counties have USDA Continuum Codes $= 1, \leq 3$, and both counties are in the top 100 most populous counties). Columns 2 to 4 include controls for 10, 50, and 100 distance-tile indicators. We interact social connectedness with an indicator for whether an institution-firm pair is closer than 100 miles in columns 5 and 6. Panel C reports pairwise robustness test results using panel PPML regressions. We consider institution holdings between Jun 2007 and Dec 2016. Column 1 includes controls for institution style in Bushee types. Columns 2 to 4 include controls for firms' economic presences in institution states. In Panel A and B, we also consider Firm, Institution × Industry, Distance 500-tile, Same County, and Same State fixed effects. Same County (Same State) is a dummy equal to one if the institution and the firm are located in the same county (same state) and zero otherwise. Distance 500-tile fixed effects indicate the quantile of the distance between the firm and the institution based on all firm-institution pairs. Refer to Appendix A.1 for detailed variable definitions. Industry classification is based on Fama-French 48 industries. In Panel C, we additionally consider Firm × Quarter, Institution × Quarter, Firm × Quarter × Style, and Institution × Quarter × Industry fixed effects. Standard errors are clustered by institution and firm in Panel A and B, and we further cluster by quarter in Panel C. t-statistics are reported in parentheses. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

Panel A: Robustness Tests for Cross-sectional Pairwise Regressions

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	% PF	**% PF Winsorized**	**% PF County**	**% PF**	**I(% PF > 0)**
Log Social Connectedness	0.398***	0.161***	0.441***	11.624***	0.554***
	(10.85)	(12.03)	(7.68)	(8.92)	(11.10)
Model	PPML	PPML	PPML	OLS	OLS
Firm FE	YES	YES	NO	YES	YES
Institution × Industry FE	YES	YES	NO	YES	YES
Distance 500-tile FE	YES	YES	YES	YES	YES
Same State FE	YES	YES	YES	YES	YES
Same County FE	YES	YES	YES	YES	YES
Institution County FE	NO	NO	YES	NO	NO
Firm County FE	NO	NO	YES	NO	NO
N	3,429,572	8,694,060	193,224	8,694,060	8,694,060
Pseudo R^2	0.562	0.526	0.451	0.073	0.403

Table IA.2: (Continued)

Panel B: Robustness Test for Pairwise Regressions: Metro Status and Physical Distance

	(1)	(2)	(3)	(4)	(5)	(6)
Log Social Connectedness	0.301***	0.259***	0.284***	0.294***		
	(11.69)	(10.56)	(10.64)	(11.03)		
× I(<=100 Mile)					0.347***	0.342***
					(6.12)	(3.38)
× I(>100 Mile)					0.106***	0.108***
					(6.13)	(4.41)
Firm FE	YES	YES	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES	YES	YES
Same State FE	YES	YES	YES	YES	YES	YES
Same County FE	YES	YES	YES	YES	YES	YES
Populous County Pair FE	YES	NO	NO	NO	NO	NO
Distance Tile FE	500	10	50	100	500	NO
N	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060
Pseudo R^2	0.508	0.506	0.506	0.507	0.548	0.550

Table IA.2: (Continued)

Panel C: Robustness Tests for Pairwise Panel Regressions

	(1)	(2)	(3)	(4)	(5)
Log Social Connectedness	0.274***	0.269***	0.263***	0.274***	0.274***
	(11.51)	(10.97)	(10.80)	(11.16)	(11.16)
Citation Share		0.059**			
		(2.47)			
I(Cited)			0.093***		
			(6.85)		
Revenue Share				0.367**	
				(2.10)	
I(Revenue State)					0.164*
					(1.76)
Firm $\times$ Quarter FE	NO	YES	YES	YES	YES
Institution $\times$ Quarter FE	YES	YES	YES	YES	YES
Institution $\times$ Industry FE	YES	YES	YES	YES	YES
Distance 500-tile FE	YES	YES	YES	YES	YES
Same State FE	YES	YES	YES	YES	YES
Same County FE	YES	YES	YES	YES	YES
Firm $\times$ Quarter $\times$ Style FE	YES	NO	NO	NO	NO
N	2.881e+08	2.670e+08	2.670e+08	2.881e+08	2.881e+08
Pseudo R^2	0.4589	0.450	0.450	0.442	0.442

Table IA.3: Two-Dimensional Heterogeneity Based on Firm Size and Analyst Coverage

This table examines if the heterogeneity in firm size reported in Panel A of Table 5 and analyst coverage reported in Panel B of Table 5 subsume each other. We alter the specification in the column 3 in Panel B of Table 2 by interacting *Log Social Connectedness* with 3×3 double-sorted indicator variables. All institution-firm pairs are assigned into one of the 3×3 independently sorted groups on firm size and analyst coverage. We report the coefficients of the interacted coefficients of these indicators with *Log Social Connectedness*. Standard errors are double clustered by institution and firm, and t-statistics are reported in parentheses. The number of observations in each cell are reported below standard errors. F statistics for whether the coefficients of the high coverage and low coverage are statistically different for each market capitalization tercile. We also report the F statistics of the joint test on whether the coefficients in of high coverage and low coverage firms are different in all three market capitalization categories. ***, **, and * indicate significance levels of 10%, 5%, and 1%, respectively.

	Small Cap	**Mid Cap**	**Large Cap**
	0.527***	0.603***	0.666***
Low Coverage	(5.12)	(6.13)	(3.42)
	1,304,109	746,086	101,739
	0.380*	0.485***	0.441***
Mid Coverage	(1.82)	(8.32)	(7.83)
	437,786	1,865,215	727,588
	-1.748	0.381***	0.216***
High Coverage	(-1.17)	(4.01)	(8.06)
	43,162	779,999	2,688,376
F Test (Low=High)	2.30	2.63	5.26**
F Test (Joint Low=High)		10.04**	

Table IA.4: Social Proximity to Capital, Physical Proximity to Capital, and Firm Outcomes

We examine how social proximity to capital affects firms' outcomes across firms with different levels of physical proximity to capital (PPC). The dependent variables are total institutional ownership (TIO) in columns 1 and 2, log market to book in columns 3 and 4, log Tobin's Q in columns 5 and 6, log effective spread in columns 7 and 8, and log illiquidity in columns 9 and 10. The key independent variables are Log Social Connectedness, interacted with an indicator for whether the firm has high Physical Proximity to Capital. Refer to Table 8 for controls in columns 1 and 2, Table 9 for controls in columns 3 to 6, and Table 10 for controls in columns 7 to 10. Standard errors are double clustered by quarter and firm, and t-statistics are reported in parentheses below each estimate. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	TIO%		Log (M/B)		Log (Tobin's Q)		Log (Effective Spread)		Log (Illiquidity)	
Log Social Proximity to Capital	1.089	1.049	0.099***	0.101***	0.047***	0.049***	-0.076***	-0.074***	-0.229***	-0.219***
× Low Physical Proximity to Capital	(1.13)	(1.10)	(3.46)	(3.55)	(2.74)	(2.82)	(-3.31)	(-3.23)	(-3.27)	(-3.16)
Log Social Proximity to Capital	0.779	0.835	0.063**	0.058**	0.029*	0.025	-0.019	-0.021	-0.080	-0.081
× High Physical Proximity to Capital	(0.96)	(1.03)	(2.43)	(2.24)	(1.79)	(1.51)	(-0.89)	(-0.96)	(-1.30)	(-1.32)
Control	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Quarter FE × Split	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO
Industry FE × Split	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO
State FE × Split	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Quarter FE × Industry FE × Split	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
F Test (No Heterogeneity)	0.066	0.032	0.954	1.382	0.631	1.094	3.841*	3.362*	3.055*	2.680
N	99,555	99,555	96,762	96,762	96,761	96,761	100,502	100,502	101,008	101,008
R^2	0.368	0.389	0.301	0.332	0.339	0.367	0.775	0.786	0.798	0.806

Table IA.5: Portfolio Social Connectedness and Performance by Institution Type

This table reports portfolio returns, volatility, and Sharpe ratios of institutional investors with different propensities to hold socially connected stocks by their Bushee (1998) classification. The propensity to hold socially connected stocks for each institutional investor is estimated using the following equation: using the following equation:

$$\%PF_{i,j} = \exp[\beta_{SC,i} \textit{Log Social Connectedness}_{i,j} + \beta_2 \textit{Same County}_{i,j} + \beta_3 \textit{Same State}_{i,j} + \textit{Firm FE} + \textit{Institution} \times \textit{Industry FE} + \textit{Distance 500-tile FE}] \cdot \epsilon_{i,j}.$$

We sort institutional investors into deciles based on their propensity to hold socially connected stocks ($\beta_{SC,i}$). The portfolio is rebalanced at the end of each quarter using institutions' previous quarter end holdings. We report results based on portfolios' percentage excess returns over the risk-free rate, CAPM, and Fama-French 5-factor adjusted returns in the first three columns. In columns 4 to 6, we report the standard deviation of the excess or residual returns from factor models. We first calculate return volatility for each institution in each quarter using daily return data. Then, we report the average returns in a given decile. Columns 7 to 9 report Sharpe ratios (or information ratios) of institutions. We compute average returns or risk-adjusted returns using daily data in each quarter. Then we divide the returns by standard deviations of the excess residual returns. For these analyses, we compute t-statistics using quarter and firm clustered standard errors. T-statistics are reported in parentheses. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

		Return			σ(Return)			Sharpe Ratio		
Bushee Type	$\beta_{SC,i}$	(1) Excess	(2) CAPM	(3) FF5	(4) Excess	(5) CAPM	(6) FF5	(7) Excess	(8) CAPM	(9) FF5
	Low	0.061*	0.022	0.026**	1.723***	1.260***	1.101***	0.056***	0.018	0.040***
		(1.91)	(1.64)	(2.47)	(13.16)	(15.25)	(15.37)	(3.87)	(1.27)	(3.47)
Dedicated	High	0.043	-0.000	0.005	1.947***	1.321***	1.151***	0.040***	-0.005	0.009
		(1.26)	(-0.00)	(0.51)	(9.37)	(7.26)	(7.22)	(3.09)	(-0.65)	(1.38)
	High-Low	-0.018	-0.022	-0.021	0.224	0.062	0.050	-0.016*	-0.023*	-0.031**
		(-1.33)	(-1.64)	(-1.57)	(1.27)	(0.34)	(0.31)	(-1.73)	(-1.81)	(-2.41)
	Low	0.042	0.006	0.005**	1.243***	0.516***	0.418***	0.061***	0.010	0.022***
		(1.57)	(1.61)	(2.50)	(12.43)	(15.50)	(15.62)	(4.22)	(1.68)	(5.18)
Quasi-Indexer	High	0.039	0.003	0.003	1.200***	0.467***	0.400***	0.062***	0.012	0.016***
		(1.49)	(1.29)	(1.31)	(11.94)	(13.55)	(13.80)	(4.35)	(1.65)	(2.84)
	High-Low	-0.003	-0.003	-0.002	-0.044	-0.049	-0.018	0.001	0.002	-0.005
		(-1.00)	(-0.82)	(-1.00)	(-1.31)	(-1.34)	(-0.57)	(0.63)	(0.36)	(-1.02)
	Low	0.044	0.003	0.006	1.421***	0.703***	0.576***	0.059***	0.009	0.030***
		(1.42)	(0.52)	(1.12)	(13.03)	(15.05)	(15.42)	(3.80)	(0.93)	(4.21)
Transient	High	0.050*	0.012*	0.016***	1.506***	0.814***	0.687***	0.056***	0.009	0.022***
		(1.70)	(1.74)	(3.34)	(13.90)	(11.45)	(10.79)	(3.98)	(1.36)	(3.69)
	High-Low	0.006	0.008	0.010*	0.085	0.110	0.110	-0.003	0.000	-0.007
		(1.04)	(1.49)	(1.83)	(1.30)	(1.46)	(1.62)	(-1.14)	(0.02)	(-1.04)

Table IA.6: Performance of Socially Connected Holdings: Local vs Non-Local Firms

This table reports portfolio returns ranked based on stock performance of socially connected and unconnected holdings, split by whether the firm is considered a local firm. Local firms are those located within 100 miles from institutions' headquarters. We first split holdings into local and non-local firms. We then sort the firms based their social connectedness to an institution into tercile as in 12. We report excess returns, CAPM, and FF-5 adjusted returns. Differences between low connectedness stocks and high connectedness stocks are reported in columns 10 to 12. Newey-West adjusted t-statistics are reported in parentheses. ***, **, and * indicate significance level of 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Low Connectedness				Med		High Connectedness				High-Low	
	Excess	CAPM	FF-5	Excess	CAPM	FF-5	Excess	CAPM	FF-5	CAPM	CAPM	FF-5
Non-Local	0.039*	0.003	0.002	0.039*	0.002	0.001	0.039*	0.003	0.004**	-0.001	-0.000	0.001
	(1.66)	(0.83)	(0.45)	(1.67)	(0.74)	(0.48)	(1.70)	(1.64)	(2.06)	(-0.20)	(-0.08)	(0.31)
Local	0.044*	0.010**	0.008*	0.045**	0.009***	0.010***	0.043*	0.007**	0.008***	-0.002	-0.003	-0.001
	(1.86)	(2.13)	(1.81)	(1.99)	(3.52)	(3.67)	(1.79)	(2.29)	(3.01)	(-0.47)	(-0.74)	(-0.13)
Local - Non-Local	0.006	0.007	0.008*	0.007*	0.007*	0.009***	0.004	0.002	0.005***	-0.003	-0.005	-0.004
	(1.23)	(1.44)	(1.68)	(1.77)	(1.94)	(2.61)	(1.53)	(1.12)	(2.62)	(-0.65)	(-1.14)	(-0.92)

Figure IA.1: Coefficients of 500-Tile Distance Indicators

We display the coefficient estimates for the 500-tile distance indicators in our baseline regressions. The left panel shows the plot after including *Log Social Connectedness*, and the right panel shows the result excluding *Log Social Connectedness* as a regressor.