NBER WORKING PAPER SERIES

SOCIAL PROXIMITY TO CAPITAL: IMPLICATIONS FOR INVESTORS AND FIRMS

Theresa Kuchler
Yan Li
Lin Peng
Johannes Stroebel
Dexin Zhou

Working Paper 27299
http://www.nber.org/papers/w27299

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020

We thank Elsa Allman, Yakov Amihud, Zhi Da, Phil Dybvig, Abhinav Gupta, Zhiguo He, Xing Huang, Chi-Ching Hung, Narasimhan Jegadeesh, Wenlan Qian, Charles Lee, Cameron Peng, Oliver Rehbein, Simon Rother, Wei Xiong, Hongjun Yan, Jun Yang, as well as seminar participants at Baruch, NYU Stern, the New York Fed, and the International Macroeconomics and Finance Conference for helpful comments. Lin Peng acknowledges the Krell research grant and the Wassarman research grant for financial support. For part of this project, Stroebel and Kuchler were research consultants at Facebook. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Social Proximity to Capital: Implications for Investors and Firms
Theresa Kuchler, Yan Li, Lin Peng, Johannes Stroebel, and Dexin Zhou
NBER Working Paper No. 27299
June 2020
JEL No. G2,G3,G4

ABSTRACT

We use social network data from Facebook to show that institutional investors are more likely to invest in firms from regions to which they have stronger social ties. This effect of social proximity on investment behavior is distinct from the effect of geographic proximity. Social connections have the largest influence on investments of small investors with concentrated holdings as well as on investments in firms with a low market capitalization and little analyst coverage. We also find that the response of investment decisions to social connectedness affects equilibrium capital market outcomes: firms in locations with stronger social ties to places with substantial institutional capital have higher institutional ownership, higher valuations, and higher liquidity. These effects of social proximity to capital on capital market outcomes are largest for small firms with little analyst coverage. We find no evidence that investors generate differential returns from investments in locations to which they are socially connected. Our results suggest that the social structure of regions affects firms' access to capital and contributes to geographic differences in economic outcomes.

Theresa Kuchler
Stern School of Business
New York University
44 West 4th Street
New York, NY 10012
and NBER
tkuchler@stern.nyu.edu

Johannes Stroebel
Stern School of Business
New York University
44 West 4th Street
New York, NY 10012
and NBER
johannes.stroebel@nyu.edu

Yan Li
Baruch College
Yan.Li@baruch.cuny.edu

Dexin Zhou
Baruch College
Dexin.zhou@baruch.cuny.edu

Lin Peng
Department of Economics and Finance
Baruch College, City University of New York
lin_peng@baruch.cuny.edu

1 Introduction

There are substantial 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 located in 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 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 (even though we find that investments in socially proximate regions do not generate excess returns). We then document that firms in regions that are socially connected to institutional capital have higher liquidity and higher valuations, and argue that this relationship is likely caused by the social proximity to capital instead of omitted variables.

To measure 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 headquarter counties to be connected via a friendship link on Facebook.[1]

Our focus on the role of social connections in shaping public equity investments by institutional investors is driven by three considerations. First, institutional investors play a key role in providing capital, market liquidity, and corporate governance for U.S. firms (see Gompers and Metrick, 2001; Blume and Keim, 2012; Aghion et al., 2013). Access to institutional capital can therefore be crucial for firms to finance their operations and to grow. Second, data on both the 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 document that institutional investors invest more in firms located in regions that they are socially connected to. In our baseline specification, we control for firm and institution $\times$ industry fixed

[1] Present-day friendship links between counties are determined by a large number of factors, including historical migration movements. 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) for a detailed discussion of the determinants of social connectedness in the United States.

effects to alleviate concern that our results might be picking up other confounding factors such as investors locating in regions that are socially connected to industry clusters important to the investors' mandates. 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 1.9% increase in the weight of the firm in the investor's portfolio.

An extensive literature has documented that investors have a preference for geographically proximate firms, a feature often referred to as "home bias" (e.g., Coval and Moskowitz, 1999; Baik et al., 2010; Bernile et al., 2015). Consistent with this literature, we find that the shorter the geographic distance between a firm and an investor, the more investment occurs. However, controlling for the social proximity between firms and investors renders the effect of physical distance statistically insignificant or even changes its sign in some specifications. Moreover, the inclusion of controls for physical distance does not affect the estimated effect of social distance on investments. These findings suggest that in many of these prior studies, physical proximity may have served as a proxy for social proximity.

Next, we explore heterogeneity in the effect of social connectedness on investment decisions along characteristics of investors and firms. The effect of social connections on investments is largest for firms that are small or have low analyst coverage. Investors are likely to be less familiar with these firms, but stronger social connections may allow them to be more aware of their existence, or to have better true or perceived information about them. In addition, the effects are larger for institutions that rely more heavily on non-financial and intangible factors rather than quantitative accounting measures. Smaller institutional investors with fewer resources are also disproportionately more likely to invest in firms from locations they are socially connected to.[2] We also explore how managers' characteristics relate to the tendency of investing in socially-proximate stocks. We find that older fund managers are generally more likely to invest in firms to which they are socially connected; we do not find significant differences based on the managers' gender or business education background (i.e., having an MBA degree).

The previous cross-sectional analyses control for both firm and institution $\times$ industry fixed effects. Despite these tight controls, one might worry about possible omitted variables at the firm-investor-pair level that correlate with social connectedness between firm and investor locations, but that also independently affect the investor's propensity to hold stocks of a particular firm. While any such variable would be unlikely to generate the heterogeneities shown above, we next provide further identification using panel data of institutional holdings from June 2007 through December 2016. We focus on the subsample of firms that changed their headquarters' location during the sample period, and include institution $\times$ firm fixed effects to capture any time-invariant determinants of an institution's preference for holding a particular stock. The within-institution-firm-pair variation in social connectedness 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

[2]This finding is consistent with the results in Pool et al. (2012), who document that managers from smaller fund families particularly 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, and generate more immediate price reactions and weaker post announcement drifts. Those results are also stronger for smaller firms and firms with low institutional ownership.

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.

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 does social proximity to institutional capital affect other 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-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 an 18.7 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 institutional 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, Scheinkman and Xiong (2003) show that, in a dynamic setting with short-sale constraints and time-varying investors disagreement, valuations disproportionately reflect the assessments of the most optimistic investors. Since firms known to more investors in general are also more likely to attract the attention of particularly optimistic investors, valuations might increase with social proximity to capital.

Consistent with these potential mechanisms, we find that firms located in regions that are socially closer to institutional investors have higher valuations. In particular, a 10% increase in social proximity to capital is associated with a 1% increase in a firm's market-to-book ratio and a 0.57% increase in its Tobin's Q. These results are robust to including a large number of firm and county controls, such as the physical proximity to capital and state $\times$ industry fixed effects. We also show that the effect of social proximity to capital on firm valuation is generally the highest 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 have been shown to 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 effects are larger for smaller firms with lower analyst coverage. These results suggest that social connectedness improves a firm's access to institutional capital and leads to beneficial firm outcomes such as higher valuation and secondary market liquidity.

Our results so far exploit cross-sectional variation in social proximity to capital to estimate its effects on secondary market liquidity and valuations. 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 these counties (or firms in those counties) other than social proximity to capital that leads them to have higher liquidity and valuations (independent of their ownership structures). For example, firms in places with high social proximity to capital might just happen to be less risky, for reasons that are not accounted for by the comprehensive set of control variables used in the prior literature. One might also speculate that places with high social proximity to capital have more well-known firms, conditional on industry and other controls, and thus will have higher liquidity provision from *all* investors. 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.

However, 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 stocks' heterogeneous social connectedness to areas affected by the shock. Specifically, we study the effects of Hurricane Sandy, the second-costliest Hurricane in the U.S. history. Hurricane Sandy's landfall in the Mid-Atlantic region on October 22, 2012 resulted in $71.4 billion in damages, led to major flooding in New York city's transportation systems, and left six million people without power and thousands of homes destroyed. 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 a high social proximity to institutional capital in the Mid-Atlantic states experienced a relative reduction in their stock liquidity. Regular seasonal patterns in the relative liquidity of firms with social proximity to capital in the Mid-Atlantic states do not account for our findings. This provides further support that our results are not driven by omitted firm-level characteristics that affect liquidity provision by all investors. Instead, they reinforce 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.

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; Rehbein et al., 2020). On the other hand, if investments in socially proximate firms are driven by investor awareness of such firms without proprietary information, we do not expect investors to outperform, and suboptimally diversified portfolios could even lead them to underperform.

We examine institutional investors' performance along three dimensions: (i) across-institution comparisons of returns for investors with a differential general propensity to invest in connected firms, (ii) comparisons at the institutional holding level, where we compare an institution's high-connectedness holdings to its low-connectedness holdings, and (iii) comparing an institution's high-connectedness holdings to high-connectedness stocks the institution chooses not to hold. We first estimate an 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.[3] 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.

2 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 are socially connected to. We begin by describing our measures of social connectedness between U.S. counties, as well as the construction of our investment variables.

Measuring Social Connectedness. To measure social connectedness between U.S. counties, we use the *Social Connectedness Index* provided 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 social 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), Kuchler et al. (2020), Rehbein et al. (2020) and Bali et al. (2019) provide evidence that friendships observed on Facebook are a good proxy for real-world U.S. social connections.

[3] The diverging conclusion in the literature on whether interactions through social networks can help improve investment performance may be due to the fact that different social networks might be differentially able to convey real information.

To measure social connectedness between geographic areas, Bailey et al. (2018a) map Facebook users to their respective county locations using information such as the users' regular IP addresses. They then construct a measure of the relative number of friendship links between each county pair, $Friendships_{i,j}$. Our measure of social connectedness between two counties corresponds to the (relative) probability that a Facebook user in county i is friends with a Facebook user in county j:

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

where $Population_i$ corresponds to the total population in county i. Figure 1 shows heatmaps of our measure of $Social\ Connectedness_{i,j}$ for San Francisco, 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 location. San Francisco 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. 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.[4] 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:

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

[4]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.

We obtain institutional investors' headquarter locations from Bernile et al. (2015, 2019), who collect this information from Nelson's Directory of Investment Managers and by searching SEC filings.[5] We obtain firms' historical headquarter locations from Compustat.

Overall, we have information on 3,083 firms and 2,820 fund families. Panel A of Table **1** presents summary statistics on institution-firm pairs (see also 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]

2.1 Empirical Analysis

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

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

This functional form is motivated by the binscatter plots in Figure 2, which suggest a linear relationship between *Log %PF* and $Log\ Social\ Connectedness_{i,j}$, both with and without controlling for the geographic distance between firms and investors.[6] 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.[7] We cluster standard errors by firm and investor.

[Insert Figure 2 near here]

We report results from regression 3 in Table **2**. The first column shows our baseline estimates. The coefficient on $Log\ Social\ Connectedness_{i,j}$ is positive and statistically significant, consistent with institutional investors overweighting firms that are headquartered in counties to which the investors are socially connected to. 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 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 concern is that social connectedness may be correlated with economic linkages or business ties between two counties that might also affect investments directly. For example, investors may be located

[5] 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.

[6] 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.

[7] This regression 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). We use the estimation procedure implemented in Correia et al. (2019).

in regions that are socially proximate to industry clusters important to the investors' mandates. To rule out such concerns, we next also include institution × Fama-French 48 industry fixed effects to account for institutions' preferences to locate in areas that are socially connected to a particular set of industries. Column 2 shows that the inclusion of institution × industry fixed effects increases the regression R^2 from 32% to 51%, suggesting that the industry preference explains a substantial portion of cross-sectional variation in funds' holding choices. Nevertheless, the coefficient of *Log Social Connectedness*$_{i,j}$ remains unchanged, suggesting that our results are robust with respect to such industry-related ties.

[Insert Table 2 near here]

In column 3, we examine the role of physical distance, on investment choices. Consistent with the home bias literature, we find that institutional investors tend to overweight firms that are headquartered in close geographic proximity (e.g., Coval and Moskowitz, 1999).[8] In column 4, we include both *Log Social Connectedness* and *Log Distance* as explanatory variables, two variables that 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 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 home bias, which often interprets geographic home bias as a result of strong connection 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 social networks than physical distance is.

The next specifications explore whether our measure of social connectedness picks up a potentially non-linear relationship between physical distance and institutional investment decisions. In column 5, we control for the geographic distance between county pairs using 500-tile dummies (Appendix Figure IA.1 reports the coefficients on these dummies). The effect of social connectedness on investments even increases somewhat in this specification. 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 (e.g., Hong et al., 2004). 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 the first quintile of social connectedness. There is a monotonic increase in portfolio weights across quintiles of social connectedness.[9]

[8] We also calculate the percentage home-bias measure used in Coval and Moskowitz (1999). Institutions in our sample exhibit a percentage bias of 6.54% using the equal-equal weighting scheme and 1.2% in the value-value weighting scheme. These measures are qualitatively consistent with, but quantitatively lower than those reported in Coval and Moskowitz (1999), who report 9.32% and 11.2% using the equal-equal and value-value weighting schemes, respectively. These difference are consistent with Bernile et al. (2019), who document a declining trend in home bias in recent years. Additionally, our analyses are at the fund-family level, while Coval and Moskowitz (1999) analyze individual funds, which are likely to be less diversified.

[9] We also conduct a number of tests to ensure the robustness of our baseline results to alternative specifications. Column 1 of Appendix Table IA.2 reports the result when we drop firms and institutions located in New York and California, two states with heavy presence of both investors and firms. Column 2 reports the result when we winsorize our dependent variable at the 99th percentile to remove extreme outliers. Columns 3 and 4 report the result using OLS regressions: in column 3, the dependent variable is $\%PF$ and in column 4, it is a dummy variable indicating whether there is non-zero institutional holding as the dependent variable. Our results are robust using all the aforementioned specifications.

The key takeaway from these findings is that institutional investors tend to overweight firms they are socially connected to. As discussed in the introduction, our interpretation of this result is that institutional investors either have a (perceived) information advantage by being socially connected to a firm's county, or that institutional managers are prone to familiarity biases with respect to the firms located in socially connected counties. We next provide evidence for this interpretation by exploring both firm and investor heterogeneity.

Heterogeneity by Investor Characteristics. We first split institutional investors based on their Bushee (2001) 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, one would expect that our proposed channels (i.e., information advantage or familiarity bias) would make the investments of dedicated institutions most sensitive to social ties to firms.

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 **3**. 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 they 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, the investments of dedicated investors react four times more strongly to variation in social connectedness than the investments of transient investors do.

[Insert Table 3 near here]

In Panel B of Table **3**, 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 investments of large institutions. 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, and that their managers are thus 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. In particular, we split firms based on size and analyst coverage as proxies for the availability of information through other channels (see Hong et al., 2000). There are at least two reasons why investments in small and informationally opaque firms may be disproportionately affected by social connectedness. First, managers may perceive having an information advantage for opaque firms they are socially connected to. Alternatively, to the extent that social connectedness increases investor awareness of firms, these effects are likely to be less important for large firms with national brands.

The first test presented in Panel A of Table 4 is based on splitting firms by their size (i.e., their market equity at the end of the prior quarter). We group firms into size terciles, labeled as Small Cap, Mid Cap, and Large Cap, respectively. The baseline specification in column 1 shows that the relationship between social connectedness and investments is indeed highest for small firms, and lowest for the large firms. The differences are large 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 basic results are not affected by the addition of these control variables.

[Insert Table 4 near here]

We also split firms based on analyst coverage, as analysts are an important information intermediary in financial markets.[10] These results are reported in Panel B of Table 4. 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. Our results are robust to including controls for geographic distance interacted with analyst-coverage-tercile indicators. Taken together, our firm heterogeneity tests suggest that investment in small firms with little analyst coverage are most strongly affected by social connectedness to potential investors.[11]

2.2 Fund Manager Characteristics, Social Connectedness, and Fund Investment

Our next analysis turns to a sample of actively managed U.S. domestic equity mutual funds.[12] 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.[13]

[10] Analyst coverage is based on the number of analysts issuing an earnings forecast on the current fiscal year, following Hong et al. (2000). 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.

[11] Since firm size and analyst coverage are positively correlated, we further explore whether analyst coverage has explanatory power in the loading on *Log Social Connectedness* beyond 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, investment sensitivity to social connectedness is highest for firms with low analyst coverage.

[12] The mutual fund holding data are from Thomson Reuters Mutual Fund Holdings Data. This sample differs from the institution sample, as institutions also include entities such as banks, pension funds, hedge funds, and insurance companies.

[13] The fund manager location data are collected from funds' Form ADV filings. See Chang (2019) for detailed data description. We are grateful for Suzanne Chang for sharing this dataset. Fund manager characteristics are obtained from Morningstar and public records. See Chung (2018) for detailed data description. We are grateful for 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 367 unique mutual funds, of which 341 have fund manager age information. The median age of young management teams is 47 years, and of old management teams is 56 years. Among the 367 funds, 302 have only male managers and 65 have at least one female manager. Additionally, 144 funds are managed by a team with less than 50% managers with an MBA degree, while 223 funds have more than 50% managers with MBA degrees.

We report this set of analyses in Table 5. The regression specifications correspond to regression 3. Column 1 reports the baseline regression. We find that, similar to our analysis of institutional holdings, there is a positive and significant relationship between *Log Social Connectedness* and $\%PF$, even after including fixed effects for distance (500-tile), same county, same state, firm, and fund $\times$ industry. Column 2 reports the results when splitting fund managers into equal-sized age groups. The investments of funds managed by relatively young managers do not seem to be influenced by social connectedness to firms' headquarter locations. Instead, it is the older managers who invest significantly more in firms connected to the funds' location. Column 3 explores if the relationship between social connectedness and investment decisions differs for funds managed only by male managers versus funds with at least one female manager. We find that both types of funds exhibit similar tendencies to hold socially connected stocks. Column 4 analyzes how fund managers' education affects their tendency to hold stocks 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 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 5 near here]

2.3 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 tight controls in our cross-sectional analysis, one might worry about possible omitted variables at the firm-investor-pair level that correlate with social connectedness between firm and investor locations, but that 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 use quarterly panel data of institutional 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.

We then focus on the subsample of 556 firms that changed their headquarter locations during the sample period. For this sample, we can include firm $\times$ investor fixed effects while retaining time-series variation in social connectedness.[14] These fixed effects allow us to capture any time-invariant determinant of an investor's preference for holding a particular stock. The within-firm-investor-pair variation in social connectedness continues to affect investment patterns: when a firm moves its headquarter 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. Column 4 establishes this result in the full panel; the effect of social connectedness on holdings is statistically significant but smaller than in

[14]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 rely on the time-varying institutional investor locations for identification purpose.

column 3. This is likely related to the fact that investor portfolios do not adjust immediately following the headquarter change. Similarly, attention to firms does not adjust discretely around the headquarter change, for example because individuals in the original headquarter county do not immediately 'forget' about the firm the moment it moves to another county. To address this, in column 5 of Table **6**, we include, for every firm-investor pair, only the holdings in the three years before and the three years after the headquarter change; this allows sufficient time for an adjustment of both investments and awareness around the headquarter move. The coefficient on *Log Social Connectedness* is 0.381, which is similar to the magnitude in the full panel analysis in column 3 of Table **6**.

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.

3 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 higher "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.

3.1 Data and Measurement

Our main explanatory variable in this section, the *Social Proximity to Capital* 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 asset under management by fund families 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.[15] Figure 3 shows a heat map of this measure 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

[15]Since we found that social connectedness had 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.

of *Social Proximity to Capital*, while counties in the middle of the country tend to have lower *Social Proximity to Capital.* However, consistent with the evidence above 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.

3.2 Social Proximity to Capital and Firms' Institutional Ownership

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

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

where $\%TIO_{i,t}$ represents the total institutional ownership share of firm i in quarter t, and $X_{i,t-1}$ includes firm-level control variables that have been shown to affect a firm's institutional ownership share (Baik et al., 2010; Green and Jame, 2013) as well as controls for county characteristics.[16] Our baseline specification, reported in column 1 of Table **7**, also includes quarter, state, and industry fixed effects. We find that *Social Proximity to Capital* is significantly related to firms' institutional ownership shares: a 10% increase in *Social Proximity to Capital* is associated with a 18.7 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.

[Insert Table 7 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

[16] 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 total assets (R&D), dividend yield (Yield), an S&P 500 index dummy (S&P), firm age (Age), advertising expenditures over total assets (Advertising), and exchange dummies (Exchange). Our county controls include log physical proximity to capital (Physical Proximity), educational attainment (Education), income per capita (Income), and unemployment rate (Unemployment). The detailed description 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.

coverage. Consistent with those results, columns 3 and 4 of Table 7 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 5 and 6, we find that the effect of *Social Proximity to Capital* on institutional investment share is most pronounced for firms with the lowest analyst coverage.

3.3 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 (i.e., 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. This theory has found strong support in subsequent empirical work (e.g., Fang and Peress, 2009; Lehavy and Sloan, 2008). We therefore hypothesize that as a result of higher investor recognition, firms with larger social proximity to capital that are more widely held by institutional investors 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. While they might find firms to be both overvalued or undervalued, short-sale constraints mean that prices disproportionately reflect the views of the most optimistic investors (see Scheinkman and Xiong, 2003). Valuations might therefore increase in the number of investors that consider a firm, since this likely leads to higher valuations of the most optimistic investors. The resulting overvaluation could persist when the investors' beliefs oscillate with the future arrival of new information. To explore whether social proximity to capital indeed affects firm valuations, we run the following regression:

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

where $Valuation_{i,t}$ 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.[17] 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.[18] The fixed effects and standard errors correspond to those in the specifications in Table 7.

The results form regression 7 are reported in Table 8. 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

[17] 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.

[18] We include the following firm-level control variables: Log Assets, Profitability, Sales Growth, Asset Turnover, R&D, Advertising, book leverage (Leverage), dividend payout (Payout), an S&P, firm age, Exchange, Physical Proximity. We also control for county-level measures of education, income, and unemployment. The detailed description of these variables can be found in Appendix A.1.

associated with a statistically significant 1.1% increase in the market-to-book ratio. The rest of Table 8 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 8 near here]

Our second measure of firm valuation is Tobin's Q. Columns 1 and 2 of Panel B of Table 8 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, R&D expenditures scaled by sales. In our sample, a one-standard-deviation increase in firm R&D expenditures is associated with a 7.7% increase in Tobin's Q, consistent with the findings in Habib and Ljungqvist (2005). Similarly, our estimates imply that a one-standard-deviation increase in *Log Social Proximity to Capital* (which corresponds to a 1.13 log point increase) is associated with a 6.4% increase in Tobin's Q.

In columns 3 and 4, 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 5 and 6 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.

3.4 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 two measures of secondary market liquidity of firm i in quarter t: the effective percentage spread and the Amihud (2002) illiquidity measure.[19] 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.[20]

[Insert Table 9 near here]

[19] 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 is that a liquid stock can allow a high trade volume passing through in any given day without significant change in price.

[20] We include the following firm-level controls: Log Assets, Log M/B, Volatility, Momentum, Turnover, Price, R&D, Yield, S&P, Age, Yield, Advertising, Exchange, Physical Proximity. We also control for county-level measures of education, income, and unemployment. The detailed description of these variables can be found in Appendix A.1.

The regression results are reported in Panel A of Table 9. Column 1 shows that a 10% increase in *Social Proximity to Capital* — equivalent to a 0.08 standard deviation increase in that number — is associated with a 0.86% 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 6% decrease in the effective spread, while a one-standard-deviation increase in *Log Social Proximity to Capital* is associated with a 9.7% decrease in the effective spread. The coefficient remains significant at the 10% level when we include Quarter × Industry fixed effect in column 2.

Next, we estimate the effect of social proximity to capital separately for firms of different sizes and with different analyst coverage. Columns 3 and 4 show that the relationship between *Social Proximity to Capital* and effective spreads is concentrated among small firms. Columns 5 and 6 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 9, 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 2.7% decrease in illiquidity. This effect is robust to including 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 13.9% reduction in the Amihud illiquidity measure, while a one-standard-deviation increase in *Log Social Proximity to Capital* is associated with a 30.5% decrease in the Amihud measure. As before, the rest of Table 9 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 firms that have relatively opaque information environments. 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 disproportionate investment in those firms by institutional investors in socially close counties relative to investments by institutional investors in socially distant counties that we documented in Section 2.1. 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 informationally opaque 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' 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 the weeks after the storm, many employees were unable to physically come into their offices due to the disruption in roads and public transportation. As a result, the ability of liquidity provision by the institutions in the affected area is likely impaired. 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,' Redler said. 'If people thought volume was thin recently, Monday could be the Wild West for low liquidity...'"* (Russolillo, 2014). As a result, the competition in liquidity provision in firms with high connectedness to East Coast-based investors is reduced, which could lead to lower levels of liquidity (Liu and Wang, 2016). We next analyze whether secondary market liquidity did indeed drop more for firms with higher social connectedness to areas affected by Hurricane Sandy.

In our baseline analysis, we define affected areas as 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 regression, 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:

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

The dependent variable, $Spread_{i,t}$, is firm i's percentage effective spread on day t. We choose spread as the main dependent variable, as it allows us to track firm-level liquidity at a very high frequency.

The key dependent variable, $Affected\ Ratio_i$, is defined as the ratio of socially proximate capital in the affected area (i.e., the Mid-Atlantic states) to the overall socially proximate capital, as measured in the quarter before Hurricane Sandy:

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

In other words, *Affected Ratio* measures the cross-sectional exposure of firms to institutional capital in affected areas. $I(Sandy)_t$ is an indicator variable that equals to one during the Sandy period, defined as October 22, 2012, to November 7, 2012.[21] In our baseline specification, we also include day fixed effects

[21] 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.

as well as firm fixed effects (the latter absorbing any direct effect of *Affected Ratio* on liquidity). $X_{i,t}$ captures control variables, which are discussed below.

[Insert Table 10 near here]

Our results are reported in Table 10. Our coefficient of interest is β, which captures how the cross-sectional variation in exposure to institutional capital affected by Sandy leads to differential changes in the effective spread during the Sandy period. The sample period covers January 2012 to July 2013. In column 1, we include day and firm fixed effects. The coefficient is positive and significant at the 1% level. This result shows that the effective spreads of firms with higher social proximity to institutional capital affected by Sandy widened more during the Sandy period compared to the spreads of other firms. The economic magnitude of the effect is large: firms with a one standard deviation higher *Affected Ratio* experienced a 5% additional increase in their effective spreads.

In column 2 of Table 10, we add industry and state fixed effects to the specification in column 1. This ensures that our results are not affected by time-variation in the industry classification of a given firm or the possibility of firms moving their headquarters. In column 3, we further control for day$\times$ industry fixed effect. The addition of this control allows us to rule out that our results are affected by time-varying liquidity differences across industries. In columns 4 and 5, we include the additional firm-level and county-level control variables from Table 9, each interacted with $I(Sandy)_t$, to the specifications in columns 2 and 3, respectively.[22] These controls rule out that our results are affected by any time-varying effect on liquidity of other firm or county characteristics that might be correlated with social proximity to capital in the areas affected by Sandy. Our baseline effects are highly robust across all specifications.

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 3.4, 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 6 to 9 of Table 10, we show that it is generally the liquidity of those same small firms with social proximity to capital in the affected areas that falls the most during Hurricane Sandy.

In summary, we show that 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 small firms with low analyst coverage. These findings provide strong evidence that the positive cross-sectional relationship between *Social Proximity to Capital* and firm liquidity is not driven by omitted firm-level or county-level variables that correlate with both firm liquidity and *Social Proximity to Capital*. In particular, our findings rule out the possibility that our liquidity results are driven by firm or county characteristics that affect liquidity provision by all investors.

4 Implications for Institutional Investors

We previously showed that institutional investors tend to overweight firms located in counties to which they are socially connected. We now examine the implications of this behavior for investor performance.

[22]We include the following controls, each interacted with $I(Sandy)_t$: Log Assets, Log M/B, Volatility, Momentum, Turnover, Price, R&D, Yield, S&P, Age, Yield, Advertising, Exchange, Physical Proximity, Education, Income, Unemployment, as well as Affected Ownership Ratio, which is the ratio of the ownership by institutions located in the Mid-Atlantic Area and the total institutional ownership.

This analysis will shed some light on the extent to which overweighting is the result of information advantages arising from the social connections. Specifically, if 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 are likely to 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 recognition or familiarity biases, investors with larger holdings of socially connected stocks should not expect to outperform, and may even underperform.[23]

We examine measures of investor performance along two dimensions: across institutions and within institutions. In the across-institution test, we sort investors by their propensity to hold 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.

4.1 Across-Institution Performance Comparison

For the across-institution test, we first obtain daily excess returns and the CAPM- and Fama and French (2015) five-factor-model-adjusted returns for each institution.[24] To measure risk, we compute the standard deviation of an institution's daily returns for a given quarter. We define an institution's Sharpe Ratio for a quarter as its average (excess or risk-adjusted) return divided by the corresponding return's standard deviation for the quarter.

We next estimate each institution's propensity to hold connected stocks using the following regression:

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

This cross-sectional regression corresponds to the specification of column 6 in Table 2, except that we allow the propensity to disproportionately hold socially connected stocks to vary across institutions – in other words, we let the loading on $\mathit{Log\ Social\ Connectedness}_{i,j}$ vary by institution, giving us a β_j coefficient for each investor. We will refer to this beta as the social-connectedness beta or β_{SC}. We then sort institutions into deciles of β_{SC} and form portfolios of institutions.

For each decile portfolio, we then compute the equally-weighted averages of the institutions' performance measures: returns, return standard deviation, and Sharpe Ratio. The results are presented in Table 11. Columns 1 through 3 present the excess return, CAPM-adjusted returns, and FF5-adjusted returns, respectively. We do not find evidence that investors that overweight stocks from areas they are socially connected to outperform. We also investigate whether firms that invest more in socially connected counties are under-diversified and therefore more risky. We report the average volatility of daily returns for institutions in each decile portfolio. We do not find that those with a higher propensity to

[23]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 results in portfolio under-diversification and therefore a deviation from the efficient portfolio.

[24]We compute institution i's daily holding period return as $Return_{i,t} = \sum_j w_{i,j,t-1} Return_{j,t}$, where $w_{i,j,t-1}$ is i's holding of stock j at end of the prior quarter and $Return_{j,t}$ is stock j's daily return.

hold socially connected stocks have higher volatilities. Finally, columns 7 through 9 report institutions' Sharpe ratios. We do not find any significant difference in Sharpe ratios between institutions in the top decile and the institutions in the bottom decile of β_{SC}.[25] Overall, there is no evidence that the propensity to hold socially connected stocks lead to a differential performance among institutional investors.

[Insert Table 11 near here]

4.2 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. Therefore, we conduct a second test to examine whether, for the same institution, holdings with high connectedness to the investor perform better than holdings with low 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 headquarter 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 headquarter 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, hedged, portfolio by taking a long position in the 'high' social connectedness portfolio and a short position in the 'low' social connectedness portfolio. 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.[26] In Table **12**, we report average returns for sub-portfolios across institutions. Panel A shows that there is no significant return difference between the 'low' and 'high' connectedness portfolios, suggesting that an institution's connected holdings do not significantly outperform its non-connected holdings.

Additionally, we also 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. We do not find that 'high' connectedness portfolios held by institutions 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_{SC} decile, institutions with small AUM, and dedicated institutions). In Panel B of Table **12**, we report the average performance difference between high connectedness and low connectedness sub-portfolios for each of these institution types, as well as between connected stocks held by institutions and connected stocks not held by institutions. Columns 1 through 3 show that there is no significant difference in excess returns, CAPM alpha, or Fama-French 5-factor adjusted

[25] Table IA.4 shows that while there is some heterogeneity across Bushee (2001) types, these differences are not economically meaningful.

[26] Similar to Section 4.1, 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.

alpha, respectively. Columns 4 through 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 12 near here]

In summary, we find that institutions do not systematically benefit from investing in socially connected stocks. Unlike the home bias literature (e.g., Coval and Moskowitz, 2001), which shows that mutual funds may have an information advantage for local stocks, we find no evidence that institutional investors have an information edge through social connections as measured by Facebook relationships. As a result, our results are more consistent with a story in which institutional investors' investments in socially connected firms is primarily driven by awareness of firms rather than by superior information.[27]

5 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.

References

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

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.

[27] 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 formed based on factors such as shared neighbourhood 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.

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, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, 2019c, Peer effects in product adoption, *Working Paper* .

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

Bali, Turan, David Hirshleifer, Lin Peng, and Yi Tang, 2019, Attention, social interaction, and investor attraction to lottery stocks, Technical report, Working Paper.

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

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* .

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

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.

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.

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.

Habib, Michel A, and Alexander Ljungqvist, 2005, Firm value and managerial incentives: a stochastic frontier approach, *The Journal of Business* 78, 2053–2094.

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

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, 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.

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.

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.

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

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.

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.

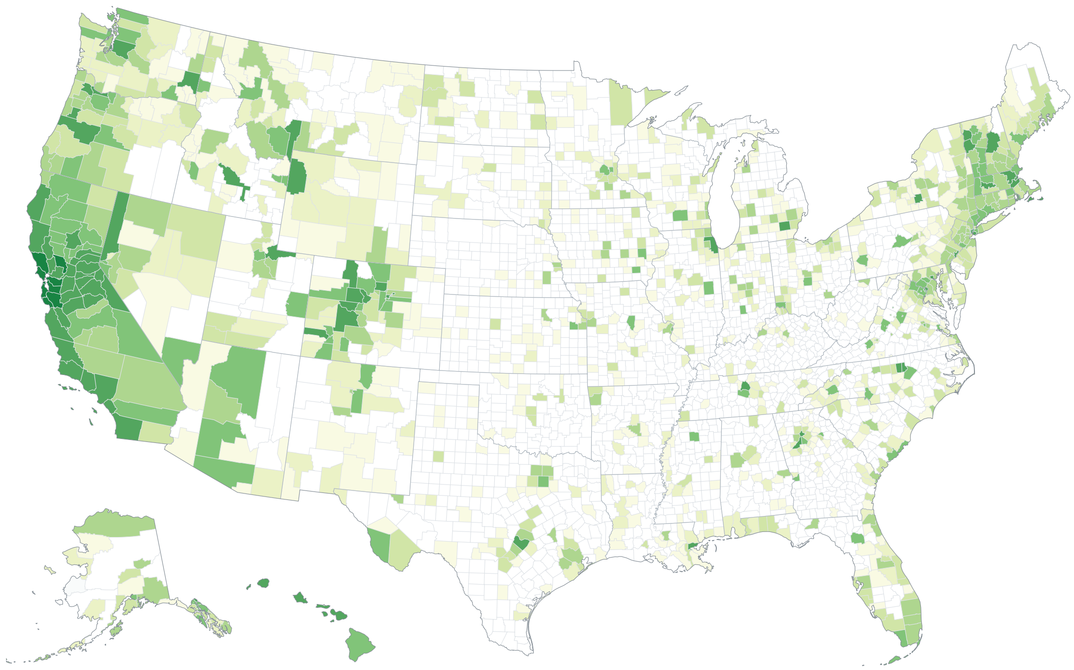
Russolillo, Steven, 2014, Traders brace for 'frankenstorm', *Wall Street Journal* .

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

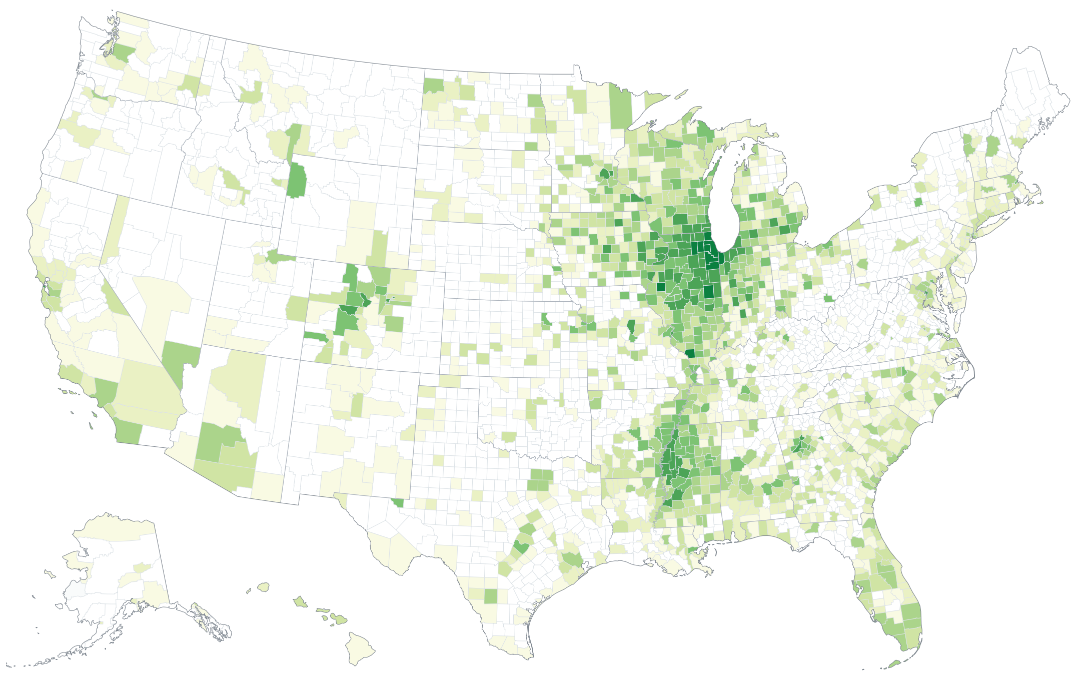
Silva, JMC Santos, and Silvana Tenreyro, 2006, The log of gravity, *The Review of Economics and statistics* 88, 641–658.

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 with the indicated 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 × 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.

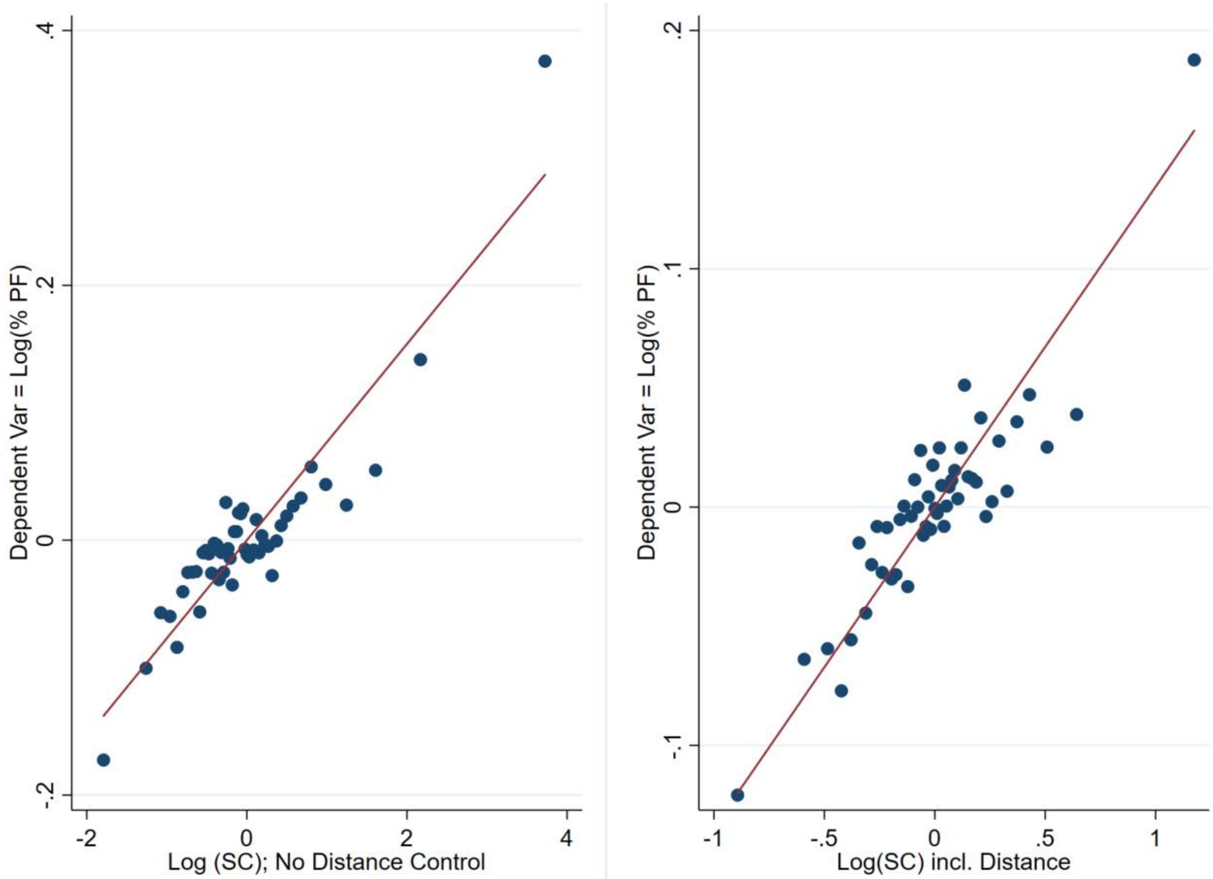


Figure 3: Heat Map of Social Proximity to Capital

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*.

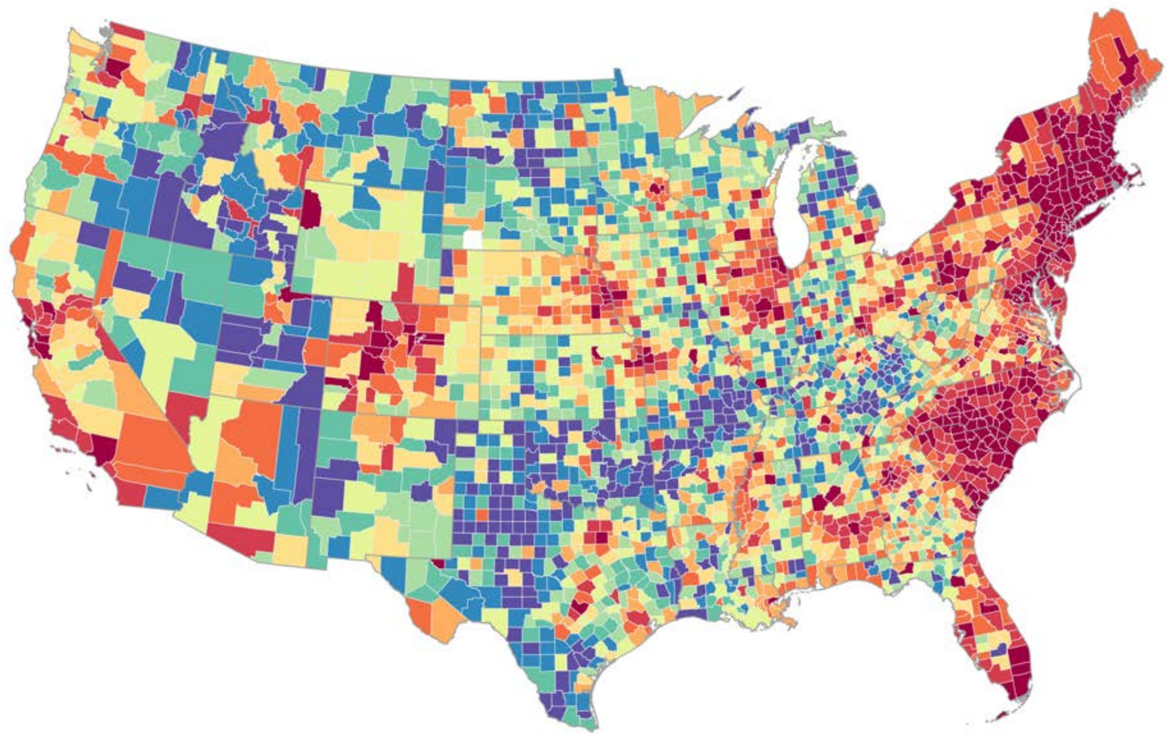


Table 1: Summary Statistics

This table reports summary statistics for our key variables. Summary statistics for the firm-institution level variables 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}). *Log Social Connectedness* is defined as *log(Social Connectedness)*. *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 sample 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. 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. Physical Proximity to Capital is $log(\sum(\frac{County\ AUM}{1+Distance}))$ summed across all U.S. counties. *ILLIQ* is the Amihud (2002) illiquidity measure at the quarterly horizon (scaled up by 10^6). *Effective Spread* is the dollar-weighted percentage effective spread (scaled up 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). The statistics reported in this panel are based on the cross section of June 2016. We report the number of institutions in each institution type, as well as the distribution institutional investors' AUM. Refer to Appendix A.1 for detailed variable definitions. Our sample includes institutions and firms located in the contiguous states of the U.S. (i.e., excluding those located in PR, VI, AK, and HI). We also 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 Jun 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

Table 1: (Continued)

Panel C: Institution Characteristics, by Institution Type

		AUM (Million USD)							
Institution Type	**N**	**Mean**	**ST. DEV**	**P5**	**P10**	**Median**	**P90**	**P95**
Dedicated	75	4215.57	16124.16	52.61	85.05	565.09	6982.40	8345.77
Quasi-Indexer	1741	5624.85	44621.73	28.98	58.00	292.18	4626.11	14066.13
Transient	724	4182.23	47348.04	15.28	31.97	344.69	4308.45	9305.58
Not Identified	543	272.289	1107.78	7.47	15.315	88.42	373.35	721.91

Table 2: Social Connectedness and Institutional Investment

This table shows the results on how social connectedness affects institutional investors' portfolio decisions. They are obtained using Poisson Pseudo Maximum Likelihood (PPML) estimators. Our sample includes all firm-institution pairs in June 2016. 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* 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 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×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. 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)	(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 × 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: Social Connectedness and Institutional Investment: by Institution Characteristics

This table shows the result on how the effect of social connectedness varies by institutional investors' characteristics. The dependent variables 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 3: (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 4: Social Connectedness and Institutional Investment: by Firm Characteristics

This table shows the result on how the effect of social connectedness varies with firm characteristics. The dependent variables is % *PF*, defined as the percentage of institutional AUM allocated to a stock. In Panel A, we report results on the heterogeneity across firms 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 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 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 4: (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 5: Social Connectedness and Individual 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 367 actively managed U.S. domestic equity mutual funds and 2,701 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. In columns 2 to 4, we interact *Log Social Connectedness* with various fund manager characteristics. Column 2 reports results on the heterogeneity across fund managers' age, where we split the 341 funds with age information into two groups by median age of the management team. In column 3, we split the 367 funds into two groups by whether the fund is (co-)managed by a female manager. In column 4, we split the 367 funds into two groups by whether more than 50% of the management team has an MBA degree. We consider Firm, Fund×Industry, 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. 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)
	Whole		**Age**		**Gender**		**Education**
Log Social Connectedness	0.126***						
	(2.96)						
		× **Young**	0.052	× **Male only**	0.106**	× **MBA Minority**	0.100**
			(1.23)		(2.27)		(2.15)
Log Social Connectedness		× **Old**	0.132**	× **Female on Team**	0.170**	× **MBA Majority**	0.100**
			(2.45)		(2.34)		(2.38)
Fund × Industry FE	YES		YES		YES		YES
Firm FE	YES		YES × AGE SPLIT		YES × GENDER SPLIT		YES × MBA SPLIT
Distance 500-TILE FE	YES		YES × AGE SPLIT		YES × GENDER SPLIT		YES × MBA SPLIT
Same County FE	YES		YES × AGE SPLIT		YES × GENDER SPLIT		YES × MBA SPLIT
Same State FE	YES		YES × AGE SPLIT		YES × GENDER SPLIT		YES × MBA SPLIT
F Test (No Heterogeneity)			1.38		0.55		0.38
N	991,267		921,041		991,267		991,267
Pseudo R^2	0.409		0.430		0.425		0.429

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 3 represents institutional holdings from June 2007 to December 2016. We eliminate firms that have not changed headquarters during the sample period in column 4. We further restrict the sample to holdings at three year (i.e., 12 quarters) before and after a change in headquarters 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)*. 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***	0.069**	0.381***
	(15.86)	(12.48)	(11.21)	(2.11)	(3.51)
Firm × Quarter FE	YES	YES	YES	NO	NO
Institution × Quarter FE	YES	YES	YES	YES	YES
Firm × Institution FE	NO	NO	NO	YES	YES
Institution × Industry FE	YES	YES	YES	NO	NO
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	2.881×10^8	2.881×10^8	2.881×10^8	34,867,427	1,162,997
Pseudo R^2	0.441	0.442	0.442	0.857	0.997

Table 7: 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 intentional 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*, *Education*, *Income*, and *Unemployment*. Refer to Appendix A.1 for detailed variable definitions. The regressions also include Quarter, State, Industry, and Quarter $\times$ Industry fixed effects. Industry fixed effects are based on the Fama-French 48 industry classification. Columns 3 and 4 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 5 and 6 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)
	Whole Sample		**Split by Size**		**Split by Analyst Coverage**	
Log Social Proximity to Capital	1.872**	1.870**				
	(2.13)	(2.14)				
Low $\times$ Log Social Proximity to Capital			2.121**	2.145**	2.203**	2.134**
			(2.28)	(2.32)	(2.33)	(2.28)
Mid $\times$ Log Social Proximity to Capital			-0.041	-0.114	0.763	0.718
			(-0.04)	(-0.11)	(0.83)	(0.78)
High $\times$ Log Social Proximity to Capital			-0.488	-0.418	0.411	0.525
			(-0.50)	(-0.44)	(0.42)	(0.55)
Firm Controls	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
Industry FE	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
State FE	YES	YES	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE
Quarter $\times$ Industry FE	NO	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE
F Test (Null: Low=High)			7.438***	7.256**	3.279*	2.698
N	97,976	97,967	97,976	97,622	97,975	97,603
R^2	0.357	0.371	0.431	0.454	0.438	0.461

Table 8: 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*, *Asset Growth*, *Asset Turnover*, *R&D*, *Advertising*, *Leverage*, *Payout*, *S&P*, *Age*, and *Exchange* dummies. Additional county controls include *Log Physical Proximity to Capital*, *Education*, *Income*, and *Unemployment*. Refer to Appendix A.1 for detailed variable definitions. The regressions also include Quarter, State, Industry, and Quarter $\times$ Industry fixed effects. Industry fixed effects are based on the Fama-French 48 industry classification. Columns 3 and 4 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 5 and 6 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)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Log Social Proximity to Capital	0.111***	0.108***				
	(4.11)	(4.03)				
Low $\times$ Log Social Proximity to Capital			0.078***	0.078***	0.088***	0.086***
			(3.10)	(3.12)	(3.19)	(3.18)
Mid $\times$ Log Social Proximity to Capital			0.054**	0.055**	0.095***	0.093***
			(2.28)	(2.34)	(3.48)	(3.49)
High $\times$ Log Social Proximity to Capital			0.062**	0.058**	0.048	0.039
			(2.20)	(2.09)	(1.62)	(1.34)
Firm Controls	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
Industry FE	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
State FE	YES	YES	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE
Quarter $\times$ Industry FE	NO	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE
F Test (Null: Low=High)			0.283	0.485	1.665	2.346
N	95,183	95,172	95,183	94,827	95,183	94,800
R^2	0.294	0.315	0.496	0.522	0.394	0.428

Table 8: (Continued)

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

	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Log Social Proximity to Capital	0.057***	0.056***				
	(3.48)	(3.41)				
Low × Log Social Proximity to Capital			0.036**	0.036**	0.047***	0.045***
			(2.49)	(2.50)	(2.96)	(2.94)
Mid × Log Social Proximity to Capital			0.034**	0.035**	0.044***	0.043**
			(2.29)	(2.34)	(2.74)	(2.71)
High × Log Social Proximity to Capital			0.037**	0.033*	0.029	0.023
			(2.08)	(1.91)	(1.57)	(1.28)
Firm Controls	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES × TERCILE	NO	YES × TERCILE	NO
Industry FE	YES	NO	YES × TERCILE	NO	YES × TERCILE	NO
State FE	YES	YES	YES × TERCILE	YES × TERCILE	YES × TERCILE	YES × TERCILE
Quarter × Industry FE	NO	YES	NO	YES × TERCILE	NO	YES × TERCILE
F Test (Null: Low=High)			0.006	0.017	0.931	1.494
N	95,182	95,171	95,182	94,826	95,182	94,799
R^2	0.333	0.354	0.552	0.576	0.437	0.469

Table 9: 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 (scaled up by 10^6). 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*, *Education*, *Income*, and *Unemployment*. Refer to Appendix A.1 for detailed variable definitions. The regressions also include Quarter, State, Industry, and Quarter $\times$ Industry fixed effects. Industry fixed effects are based on the Fama-French 48 industry classification. Columns 3 and 4 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 5 and 6 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)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Log Social Proximity to Capital	-0.086***	-0.083***				
	(-4.17)	(-4.02)				
Low $\times$ Log Social Proximity to Capital			-0.075***	-0.072***	-0.093***	-0.088***
			(-3.47)	(-3.39)	(-4.28)	(-4.09)
Mid $\times$ Log Social Proximity to Capital			-0.046**	-0.044**	-0.072***	-0.069***
			(-2.57)	(-2.42)	(-4.01)	(-3.84)
High $\times$ Log Social Proximity to Capital			0.025	0.025	-0.000	0.002
			(1.36)	(1.39)	(-0.02)	(0.10)
Firm Controls	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
Industry FE	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE	NO
State FE	YES	YES	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE	YES $\times$ TERCILE
Quarter $\times$ Industry FE	NO	YES	NO	YES $\times$ TERCILE	NO	YES $\times$ TERCILE
F Test (Null: Low=High)			20.280***	19.893***	15.068***	14.326****
N	98,908	98,899	98,908	98,559	98,907	98,537
R^2	0.771	0.780	0.832	0.843	0.825	0.836

Table 9: (Continued)

Panel B: Social Proximity to Capital and Illiquidity

	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Log Social Proximity to Capital	-0.271***	-0.265***				
	(-4.68)	(-4.62)				
Low × Log Social Proximity to Capital			-0.243***	-0.241***	-0.276***	-0.266***
			(-3.52)	(-3.51)	(-4.09)	(-3.95)
Mid × Log Social Proximity to Capital			-0.098**	-0.097**	-0.177***	-0.171***
			(-2.14)	(-2.13)	(-3.88)	(-3.77)
High × Log Social Proximity to Capital			-0.010	-0.007	-0.083*	-0.077*
			(-0.22)	(-0.16)	(-1.92)	(-1.80)
Firm Controls	YES	YES	YES	YES	YES	YES
County Controls	YES	YES	YES	YES	YES	YES
Quarter FE	YES	NO	YES × TERCILE	NO	YES × TERCILE	NO
Industry FE	YES	NO	YES × TERCILE	NO	YES × TERCILE	NO
State FE	YES	YES	YES × TERCILE	YES × TERCILE	YES × TERCILE	YES × TERCILE
Quarter × Industry FE	NO	YES	NO	YES × TERCILE	NO	YES × TERCILE
F Test (Null: Low=High)			12.691	12.939	7.902	7.590
N	99,406	99,397	99,406	99,058	99,405	99,037
R^2	0.794	0.800	0.850	0.856	0.846	0.853

Table 10: 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 Jan, 2012 to Jul, 2013. The dependent variable is daily *Log Effective Spread*. The key variable of interest is the interaction between the Sandy indicator and *Affected Capital Ratio*. *I(Sandy)* is an indicator variable equal to one during the affected period defined as Oct 22, 2012 to Nov 07, 2012. *Affected Capital Ratio* is defined as the *Social Proximity to the Mid Atlantic Capital*, divided by *Social Proximity to Capital*. We fix this ratio as of the quarter prior to Sandy (i.e. the third quarter of 2012). We consider the control variables included in Table 9 and the *Affected Ownership Ratio* and their interaction with the Sandy indicator. Refer to Appendix A.1 for detailed variable definitions. We also control for firm fixed effects. In addition, we consider Day, Industry, State, Day × Industry fixed effects. Columns 6 and 7 exhibit the result on heterogeneity across firm size terciles. Firm size terciles are based on firms' market capitalization at the end of the prior quarter. Columns 8 and 9 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. Firms in Eastern states are excluded. 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)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
	Whole Sample					**Split by Size**		**Split by Analyst Coverage**	
I(Sandy) × Affected Capital Ratio	0.240***	0.240**	0.220**	0.203**	0.200**				
	(2.64)	(2.63)	(2.64)	(2.30)	(2.56)				
Low × I(Sandy) × Affected Capital Ratio						0.353	0.343*	0.472**	0.495***
						(1.58)	(1.82)	(2.21)	(2.81)
Med × I(Sandy) × Affected Capital Ratio						0.190	0.054	0.034	-0.042
						(1.17)	(0.34)	(0.38)	(-0.36)
High × I(Sandy) × Affected Capital Ratio						0.041	0.088	0.092	0.178**
						(0.36)	(0.81)	(1.24)	(2.26)
Control × Sandy FE	NO	NO	NO	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	NO	YES	NO	YES × TERCILE	NO	YES × TERCILE	NO
Industry FE	NO	YES	NO	YES	NO	YES × TERCILE	NO	YES × TERCILE	NO
State FE	NO	YES	YES	YES	YES	YES × TERCILE	YES × TERCILE	YES × TERCILE	YES × TERCILE
Day × Industry FE	NO	NO	YES	NO	YES	NO	YES × TERCILE	NO	YES × TERCILE
F Test (Low=High)						1.426	1.223	2.559	2.300
N	545,295	545,295	544,803	545,295	544,803	545,295	540,707	545,295	541,108
R^2	0.754	0.754	0.765	0.756	0.766	0.761	0.783	0.760	0.782

Table 11: 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 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 12: 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 for each 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 12: (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 (log(SC))	The logarithm of Social Connectedness.
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
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 $\lvert R\rvert_{i,t}/V_{i,t}$, 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.
Tobin's Q	The ratio of market value and replacement cost of firm assets. Market value of a firm is defined as total assets minus book value of equity plus book value of 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.
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.
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.
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.
Momentum	Cumulative monthly return from month t-2 to t-12 before the end of the quarter.

Appendix A.1: (Continued)

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.
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.
Panel C: Institution (Mutual Fund)-Level Variables	
AUM	Asset 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	
Education Attainment (Education)	Percent of adults in a county that obtain a high school education. Obtained from American Community Survey.
Income Per Capita (Income)	County-level income per capita. Obtained from American Community Survey.
Unemployment Rate (Unemployment)	County-level unemployment rate. Obtained from American Community Survey.

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 Baseline Regressions
- Table IA.3: Two-Dimensional Heterogeneity Based on Firm Size and Analyst Coverage
- Table IA.4: Portfolio Social Connectedness and Performance by Institution Types
- 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 statistic of *Affected Ratio* used in Hurricane Sandy analysis is presented in Panel C. Summary statistics for 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	0.06	0.24	0	0	0	0	1
Same County	0.02	0.13	0	0	0	0	0

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	1,129.11	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	0.16	0.36	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	0.00	0.00	0.00	0.04	0.05
R&D Expense	1.26	80.20	0.00	0.00	0.00	0.15	0.25
Advertising Expense	0.01	0.08	0.00	0.00	0.00	0.03	0.05
Profitability	0.09	0.17	-0.09	0.01	0.10	0.22	0.28
Education Attainment	85.36	5.90	74.60	77.60	86.40	92.30	93.10
Income Per Capital	55,121	22,902	33,306	36,081	48,791	78,955	110,487
Unemployment	6.54	2.26	3.60	3.90	6.20	9.70	10.60

Table 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	0.444	0.237	0.132	0.183	0.397	0.859	0.908

Panel D: Summary Statistics for Mutual Fund-Firm Pair Variables as of Jun 2016							
Variables	**MEAN**	**ST. DEV**	**P5**	**P10**	**MEDIAN**	**P90**	**P95**
Log Social Connectedness	6.151	1.253	4.595	4.853	5.939	7.612	8.487
Log Distance	6.501	1.357	3.96	5.096	6.815	7.784	7.841
% PF	0.037	0.319	0	0	0	0	0

Table IA.2: Robustness Tests for Baseline Regressions

This table exhibits the robustness tests for our baseline results in Table 2. Our sample is based on the cross section of institutional holdings as of June 2016. In column 1, we exclude firms and institutions headquartered in New York and California. The dependent variable is % *PF* in columns 1 to 3, defined as the percentage of AUM allocated to a stock. The dependent variable in column 2 is % *PF* winsorizing the top 1%. In column 4, 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 regressions are estimated using Poisson Pseudo Maximum Likelihood (PPML) and the results in columns 3 and 4 are estimated using OLS. We 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. 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)**
	%PF	**% PF Winsorized**	**% PF**	**I(% PF > 0)**
Log Social Connectedness	0.308***	0.161***	11.624***	0.554***
	(9.02)	(12.03)	(8.92)	(11.10)
Model	PPML	PPML	OLS	OLS
Firm FE	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES
Distance 500-tile FE	YES	YES	YES	YES
Same State FE	YES	YES	YES	YES
Same County FE	YES	YES	YES	YES
Institution FE	NO	NO	NO	NO
N	4,180,932	8,694,060	8,694,060	8,694,060
R2	0.555	0.526	0.073	0.403

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

This table examines if the heterogneity in firm size reported in Panel A of Table 4 and analyst coverage reported in Panel B of Table 4 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. 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: 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 risk-free rate, CAPM, and Fama-French 5-factor adjusted returns in the first three columns. In columns 4 to 6, we report 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 of 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		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Bushee Type	$\beta_{SC,i}$	Excess	CAPM	FF5	Excess	CAPM	FF5	Excess	CAPM	FF5
Dedicated	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)
	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)
Quasi-Indexer	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)
	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)
Transient	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)
	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)

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.