Social Proximity to Capital: Implications for Investors and Firms*

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

PRELIMINARY DRAFT

Abstract

We use social network data from Facebook to show that institutional investors are more likely to invest in firms that are located in areas the investor is more socially connected to. This effect is particularly large for investments in small and informationally opaque firms. After controlling for social connectedness, investments do not vary with the geographic distance between investors and firms. We find no evidence that investors achieve higher returns when investing in firms located in areas that the investor is socially connected to. In contrast, we show that firms in locations that are socially proximate to capital have higher valuations and higher liquidity, again with the largest effects for smaller firms with lower analyst coverage. In addition to these cross-sectional results, we document that, during Hurricane Sandy, there was a disproportionately adverse effect on the liquidity of firms that were socially connected to regions affected by Sandy. Our results suggest that the geographic structure of social networks affects firms' access to capital and can contribute to geographic differences in economic outcomes.

JEL Codes: G2, G3, G4

Keywords: Social networks, Social Connectedness Index, Institutional Investors

---

*This draft: November 18, 2019. We thank seminar participants at Baruch, NYU Stern, and the New York Fed for helpful comments.

†Stern School of Business, New York University, CEPR: theresa.kuchler@stern.nyu.edu.
‡Baruch College: yan.li@baruch.cuny.edu.
§Baruch College: lin.peng@baruch.cuny.edu.
¶Stern School of Business, New York University, NBER, CEPR: johannes.stroebel@stern.nyu.edu.
||Baruch College: dexin.zhou@baruch.cuny.edu.

1 Introduction

Within the United States, there are substantial regional differences in economic outcomes. For example, firms in metropolitan and coastal regions are often more productive, innovative, and valuable than those in other parts of the country (Baum-Snow and Pavan, 2011; Moretti, 2012; Dougal et al., 2018). A growing literature explores a variety of explanations for these disparities: for example Dougal et al. (2018) argue that coastal cities with better weather are able to attract more high-skilled workers, and that firms are able to capture some of this value added. In this paper, we argue that the geographic structure of social networks of different regions influences the allocation of capital to firms, and thereby contributes to the differences in firm outcomes. Specifically, we first show that — conditional on physical distance and other controls — institutional investors invest more in firms located in regions to which they have stronger social ties. As a result, firms in regions that are more socially proximate to institutional capital have higher liquidity and higher valuations.

Our measure of social connectedness between firms and investors builds on the *Social Connectedness Index* introduced by Bailey et al. (2018a). Specifically, for each county-pair in the U.S., we compute the relative probability of two Facebook users in those counties being connected via a friendship link on Facebook, the world's largest online social networking service with 239 million active users in the United States and Canada as of the end of 2017. Given Facebook's scale, the relative representativeness of Facebook's user body, and the fact that Facebook is primarily used to connect with real-world friends and acquaintances, these data provide a comprehensive measure of the regional structure of U.S. social networks. 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. Using this data, we then construct the social connectedness between each investor and firm as the connectedness between the location of the firm headquarter and the location of the institutional investor: for example, the investor Northern Trust, based in Chicago, is disproportionately connected firm Trustmark Corporation, based in Mississippi.

We focus on the role of social connections in shaping public equity investments by institutional investors. The focus on institutional investors is driven by two considerations. First, these investors play an important role in providing capital, market liquidity, and corporate governance for U.S. firms (see Gompers and Metrick, 2001; Gompers et al., 2003; Blume and Keim, 2012; Aghion et al., 2013). Access to institutional capital is therefore crucial for firms to finance their operations and to grow. Second, there is substantial geographic variation in the location of institutional investors across the United States, with the Tri-State Area (New York, New Jersey and Connecticut) accounting for only about one-third of total U.S. institutional assets under management. The large variation in the location of institutional investors, combined with substantial differences in the geographic structure of social networks across U.S. counties, creates meaningful 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, which controls for firm and investor $\times$ industry fixed effects as well as the geographic distance between firm and investor, a 10% increase in the social connectedness between firm and investor location is associated with a 0.12 basis point increase in the share of that firm in the investor's portfolio. This effect is large relative to a sample average portfolio share of 4 basis points. From the firms' perspective, a 10% increase in social connectedness between firm and investor county increases the share of a firm held by a given institutional investor by 0.03 basis points (relative to a baseline of 2 basis points). These results are not driven by funds locating in regions that are socially proximate to industry clusters important to the investors' mandate.

An extensive literature has documented that investors invest more in geographically proximate firms (Coval and Moskowitz, 1999; Baik et al., 2010; Bernile et al., 2015). Consistent with that literature, we find that a larger geographic distance between a firm and an investor location decreases investments substantially. However, after controlling for the social proximity between firms and investors, the effect of physical distance becomes insignificant. On the other hand, the inclusion of physical distance does not affect the estimated effect of social distance on investments. This suggests that in many of these prior studies, physical proximity served as a proxy for social proximity.

Next, we explore heterogeneity in the effect of social connectedness on investment behavior. We find the largest economic effects are for investments in firms that are small and that have low analyst coverage. These firms are likely to be less well-known, and institutional investors with strong social connections to the headquarter locations of these firms might be more aware of their existence or might have better true or perceived information about their fundamentals. We also investigate the heterogeneity across investor characteristics. We find the largest effects of social connectedness on the investments of institutions that rely less on quantitative accounting measures and are more likely to use non-financial and intangible factors to make investment decisions (i.e., investors classified as "dedicated" by Bushee, 2001). We also find that smaller institutions with fewer resources are more likely to invest in firms that they are socially connected to. This is consistent with findings from Pool et al. (2012), who argued that these smaller funds who document that managers of smaller funds and funds from smaller fund families tend to overweight their home-state more than the average fund.

Overall, the heterogeneity in the effect of social connectedness on investments across characteristics of both firms and investors highlight that our baseline relationship is not primarily driven by omitted variables that affect the general propensity of institutional investors in one location to invest in firms in a connected region through channels other than social connectedness. The findings also suggest that investors' (perceived) information or familiarity advantage plays a significant role in explaining the importance of social connection in the investment process.

In the second part of the paper, we show that being located in a county that is socially proximate to capital affect firm-level capital market outcomes such as valuation and liquidity. Our measure of social proximity to capital is higher for firms in those counties with stronger social connections to other counties that are home to institutional investors that have large assets under management. We first validate our institution-firm level regression results at the firm level, and find that firms that are located in counties with high social proximity to capital tend to have more institutional ownership.

We then examine if firms located in counties with higher social proximity to capital have higher valuation. There are two possible mechanisms for such a relationship. First, firms that are located in counties with higher connection to capital are likely to be more broadly held by institutional investors. As a result, investors can better share the risks in these firms and would thus demand a lower required rate of return, which is associated with higher valuation (see Merton, 1987). Second, in a world where investors disagree about firm valuation, short-sale constraints imply that valuations reflect the assessments of the relatively optimistic investors (see Scheinkman and Xiong, 2003). As a result, firms that are considered by more unique investors have more opportunities to attract investments from particularly optimistic investors, and are thus more likely to have a high valuation.

Consistent with this prediction, we find that firms that are socially proximate to institutional investors have higher valuations as measured by market-to-book ratio and Tobin's Q. These results are unaffected by the inclusion of a large number of controls, including the physical proximity to capital and state $\times$ industry fixed effects. The economic magnitude of the relationship is large. A 10% increase in *Social Proximity to Capital* is associated with a 1% increase in the market-to-book ratio and a 0.8% increase in Tobin's Q. The effects on firm valuation is particularly large for small firms with low analyst coverage, precisely those firms where social connectedness had the largest effect on investment flows.

We find also that firms with higher social proximity to capital tend to have higher liquidity, using both effective spread and the Amihud (2002) illiquidity measure. A 10% increase in *Social Proximity to Capital* is associated with a 0.6% reduction in effective spreads and a 1.8% reduction in the Amihud (2002) illiquidity measure. This result is consistent with findings from a large literature that has documented that institutional investors are important providers of liquidity (Blume and Keim, 2012; Rubin, 2007). As before, the effects of *Social Proximity to Capital* on liquidity are larger for smaller firms with lower analyst coverage.

Despite our ability to control for many firm and county characteristics in these regressions, one concern with interpreting the firm-level results presented above is the potential for omitted variables that are correlated with both a firm's liquidity and valuation as well as its social proximity to capital. To address such concerns, we next exploit time variation in different institutional investors' abilities to provide liquidity. Specifically, we argue that Hurricane Sandy had a substantial temporary effect on the ability of institutional investors in the Mid-Atlantic region to participate in financial markets. Indeed, for a short period following the Hurricane's landfall on October 29, 2012, schools were closed and public transportation was severely impaired. We find that during this period, firms with a high social proximity to institutional capital from the Mid-Atlantic states experienced a relative reduction in market liquidity. This finding is not driven by a direct effect of Hurricane Sandy on these firms, nor by regular seasonal movements in the relative liquidity of firms with social proximity to institutional capital affected by Sandy. These findings suggest that our baseline result is not primarily driven by omitted variables that are correlated with both a firm's liquidity as well as its *Social Proximity to Capital*.

In the final section of the paper, we explore whether investors achieve higher returns by focusing their investments in regions they are socially connected to. If institutional investors obtain a helpful information advantage over their social connections, we should observe that institutions that hold more socially connected stocks are likely to outperform (see Cohen et al., 2008; Hong and Xu, 2019). On the other hand, if investments in socially proximate firms are driven by familiarity bias, we expect no outperformance or even underperformance (see Cohen et al., 2008; Pool et al., 2012).[1] We conduct

[1] Institutional investors may perceive to have informational advantage in firms that they are familiar with, even if they do not possess actual information edge.

two tests. First, we sort investors into deciles based on their propensity to invest in socially proximate firms. When comparing across these deciles, we find no significant variation in investor performance, measured by excess returns, CAPM, Fama and French (1996) three-factor, Fama and French (2015) five-factor alphas, or Sharpe Ratios. Second, we look within investors and find no differential performance between investments in firms that investors are differentially socially connected to. Thus, our results do not support the idea that institutional investors obtain valuable information through social connections as measured by Facebook friendship links. Instead, our results are more consistent an interpretation in which institutional investments in socially proximate firms are driven by familiarity biases.

We contribute to a large literature that has studied the role of social networks in investment decisions. A number of papers have studied whether investors' investments in firms they might be more familiar with allows them to outperform their competitors. For example, Cohen et al. (2008) find that mutual fund managers can trade profitability in firms with a connected top manager or a connected board member through their educational network. Consistent with word-of-mouth interaction, Hong et al. (2004) find that fund managers in the same city tend to place similar trades even for stocks that are far away. Pool et al. (2015) find that fund managers may be able to obtain information from their peer managers residing in the same neighborhood. Hong and Xu (2019) use a statistical method to infer socially connected holdings and find that stocks that are held from highly socially connected institutions tend to outperform. However, there is also evidence against the information view. For example, Pool et al. (2012) find that there is no evidence that fund managers obtain superior information from their connection to hometown firms. Bailey et al. (2011) find mutual fund managers are subject to a range of biases, including familiarity bias. These biases tend to harm mutual fund performances. Heimer (2016) studies the social interaction of individual investors, and find that social network may transmit behavioral biases. Bali et al. (2019) find that social network amplifies investor demand for lottery stocks and contributes to the overvaluation of such stocks. The conclusions from this literature are mixed. Cookson and Niessner (2018) study the disagreement using a social network dataset. Chen et al. (2014) show that stock analyses contributed to an investor social network contains value-relevant information.

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 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.

Social Connectedness. To measures the social connectedness between U.S. counties, we use information from the *Social Connectedness Index* provided by Bailey et al. (2018a). This measure was created using aggregated and 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 such 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 shows that Facebook usage rates among U.S.-based online adults were relatively constant across income groups, education levels, and race, and among urban, rural, and suburban residents; usage rates were slightly declining in age (from 88% of individuals aged 18 to 29, to 62% of individuals aged 65 and older). In the United States, 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 (Jones et al., 2013; Gilbert and Karahalios, 2009; Hampton et al., 2011). 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, such as celebrities, are common (see Bailey et al., 2018a,b, 2019, for additional evidence that friendships observed on Facebook serve as a good proxy for real-world U.S. social connections).

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

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

where $Population_i$ corresponds to the total population in county i. Figure 1 shows heatmaps of the resulting measure of $\mathit{Social\ Connectedness}_{i,j}$ to San Francisco, CA (Panel A) and Cook County, IL (Panel B), which is home to the city of Chicago. Darker colors correspond to stronger social connections to the origin 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 distance. 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 the connections from college graduates moving from college towns to San Francisco for career opportunities. Cook County, IL, 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 may have very different social connectedness to the institutional investors located in San Francisco and Chicago. This pattern is even more salient among physically distant counties. 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 Thomson Reuter's Institutional (13F) Holdings data set. This data is reported at the level of the fund family, not the level of the individual fund. In our baseline analysis, we use institutional investments data from June 2016, which most closely corresponds to the time when we observe social connectedness. In the

Appendix we show that our results remain robust when we use panel regression using the holdings data from 2007 to 2016. We combine institutional investors' holdings data with information on stock prices and shares outstanding from CRSP to construct measures of the total investment by each fund in each firm.[2] For each institution-firm pair, we construct two measures of institutional investments. The first measure, $\%PF_{i,j}$, corresponds to the share of firm i in investor j's portfolio:

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

where an institution's AUM is the sum of the equity values held by investor j. Our second measure, $\%IO_{i,j}$, corresponds to the share of firm i that is owned by investor j:

$$\%IO_{i,j} = \frac{\text{Ownership (\$) of institution } i \text{ in firm } j}{\text{Firm } j \text{ value (\$)}}. \tag{3}$$

We obtain the location of each institutional investor from Bernile et al. (2015, 2019), who collect this information from Nelson's Directory of Investment Managers and by searching SEC filings.[3] We obtain firms' headquarter location data from Compustat.

Overall, we have information on 3,083 firms and 2,820 fund families. We present summary statistics for variables based on institution-firm pairs. The mean of *%PF* is 0.04%, indicating that at the average firm-institution pair, the stock constitutes 0.04% of the institutional public U.S. equity portfolio. Most institution-firm 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 a highly diversified portfolio. The average value of *%IO* is 0.02%, suggesting that at the average firm-fund pair, the fund holds about 0.02% of the firm's equity.[4] Again, these holdings are right-skewed, with most investors holding positions in relatively few firms.

[Insert Table 1 near here]

2.1 Empirical Analysis

We use the following regression 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 decisions to invest in firm i:

$$Investment_{i,j} = \alpha + \beta \log(Social\ Connectedness_{i,j}) + \gamma X_{i,j} + \psi_i + \xi_{j\times ind(i)} + \epsilon_{i,j}. \tag{4}$$

In this specification, $Investment_{i,j}$ corresponds to either $\%Portfolio_{i,j}$ or $\%IO_{i,j}$ as defined above. The choice to control for social connectedness in this log-linear functional form is motivated by the binscatter

[2] 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.

[3] 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.

[4] The difference in the means of *%PF* and *%IO* is due to the construction of two variables. When constructing *%PF*, we only consider equity positions of institutional investors, so each portfolio by default adds up to holding 100% of equity positions. In contrast, firms are not completely held by institutional investors, so summing a given firm's *%IO* would not add to 100%.

plots presented in Figure 2. These plots suggest a linear relationship between investment variables and $\log(\mathit{Social\ Connectedness}_{i,j})$, both with and without including other controls. The vector $X_{i,j}$ includes controls for various measures of the geographic distance between the headquarter counties of firm i and investor j, as well as indicator variables for whether the firm and investor are in the same county or the same state. We include firm fixed effect so that our results are not driven by any firm characteristics. We also include investor $\times$ industry fixed effect to avoid spurious results driven by the desire to locate specialized funds in areas that are socially connected to the firms the fund invests in (we use the Fama-French 48-industries classification). We cluster the standard errors by firm and investor.

[Insert Figure 2 near here]

We report results from regression 4 in Table **2**; the regression coefficients are scaled up by a factor of 1,000 for better readability. Panel A shows results with $\%PF_{i,j}$ as the dependent variable. The first column shows the baseline estimates. The coefficient on $\log(\mathit{Social\ Connectedness}_{i,j})$ is positive and statistically significant, consistent with institutional investors overweighting firms that are headquartered in counties the investors are socially connected to. The coefficient estimate implies that a 10 percent increase in social connectedness is associated with a 0.12 bps increase in $\%PF$, relative to an unconditional average of $\%PF$ of 4 bps. Importantly, the inclusion of firm fixed effects ensures that this finding is not driven by the any characteristics that might make firms that are located in counties connected with institutional investors on average more prone to attracting institutional investments on average.

[Insert Table 2 near here]

The key independent variable in column 2 is $\log(\mathit{Distance}_{i,j})$. 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).[5] In column 3, we control for both social connectedness and geographic distance. Social connectedness remains positively related to firms' weight in institutional portfolios, with a coefficient estimate close to that from column 1. The effect of geographic distance becomes statistically insignificant and even changes sign. This finding contributes to 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 $\mathit{Social\ Connectedness}_{i,j}$ is a richer proxy for social networks than physical distance.

The next specifications explore whether our measures of social connectedness pick up a potentially non-linear relationship between physical distance and institutional investment decisions. In column 4, we control for the geographic distance between county pairs using 500-tile dummies (see Internet Appendix Figure IA.1 for coefficients on these dummies). The effect of social connectedness on investments even increases somewhat in this specification. In column 5, we further add indicators to control

[5]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.

for whether the firm-institution pair is located in the same county or the same state. These additional controls, which allow us to control for the word-of-mouth interaction within a short distance (e.g., Hong et al., 2004), do not change the economic or statistical significance of our results. Finally, column 6 explores the monotonicity of the relationship between social connectedness and investments by including Social Connectedness as quintile indicators as independent variables. Coefficients for quintiles 2 to 5 indicators 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.

We report the regression results with *%IO* as the dependent variable in Panel B of Table 2. The key difference in the depending variable is that *%IO* scales institutional holdings by the firms' overall market values rather than by investors' total equity holdings. We first conduct a univariate regression with $\log(\mathit{Social\ Connectedness}_{i,j})$ as the key explanatory variable and firm and investor $\times$ industry fixed effects as controls. The coefficient is positive and significant, and implies that all else equal, investors located in a county that is 10% more connected to a firm hold a 0.03 bps higher share of a firm's market capitalization. To put the magnitude in context, our unconditional mean of $\%IO_{i,j}$ in firms is 2 bps. The results in the remaining columns mirror those from Panel A: geographic distance is negative related to $\%IO_{i,j}$, but its effect disappears when also controlling for social connectedness. The effect of social connectedness on investments is also robust to more flexible controls for geographic distance, as well as to the inclusion of same state and same county fixed effects.

The key takeaway from these findings is that institutional investors tend to overweight firms located in counties that they are socially connected to. Our interpretation of this result is that institutional investors either have an (perceived) information advantage by being socially connected to a firm's county, or that institutional managers are prone to familiarity biases to the firms located in socially connected counties. We next provide evidence for this interpretation by exploring heterogeneity in the baseline effect along characteristics of both the firms in the investors.

Heterogeneity by Investor Characteristics. We first split institutional investors based on their Bushee (2001) classification. Bushee (2001) groups different fund families into three categories. *Transient* fund families are short-term-focused investors with high levels of portfolio turnover and diversification. *Quasi-index* families are characterized by low portfolio turnover and high diversification. *Dedicated* fund families tend to take large stakes in firms and tend to 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 this description of investment strategies, one would expect that our proposed channels would make the investments of dedicated institutions are likely to be the most sensitive to social connectedness to firms.

To explore this prediction, we interact the Bushee-type indicators with $\log(\mathit{Social\ Connectedness}_{i,j})$. 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 investor $\times$ industry fixed effects. We find that, consistent with our hypothesis, the dedicated institutions tend to have the highest propensity to invest in firms headquartered in socially connected counties. The investments of transient investors appear to be least affected by social connectedness, and the differences between the investor types are statistically significant. This finding is unaffected by including controls for physical distance between firm and fund, even when those are interacted with Bushee type. Indeed, in the specification with the richest controls, the investments of dedicated investors react six times as 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 the smallest institutions respond more than twice as much to social connectedness than investments by the largest institutions do. Panel C of Table **1** shows that this is not just the result of "dedicated" investors having lower AUM; the split by institution size therefore provides additional information about the heterogeneity in institutions' tendency 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 may have fewer resources to conduct large-scale research, and thus individual managers may be more prone to act based on their familiarity bias (e.g., Pool et al., 2012).

Heterogeneity by Firm Characteristics. Next, we explore heterogeneity of the investment sensitivity to social connectedness across various firm characteristics. Our proposed channel suggests that that investors would disproportionately invest in socially connected firms when these firms are small and informationally opaque. There are two possible mechanisms for such a relationship. First, investments might increase more in social connectedness for more opaque firms because socially connected managers might perceive to have an information advantage at those firms (e.g., Pool, Stoffman, Yonker, 2012). Alternatively, to the extent that social connectedness increase investors familiarity with firms, these affects are likely to be less strong for large firms with national brands.

Thus, we split the firms based on firm size and analyst coverage as measures of the availability of information through other channels (see Hong, Lim, and Stein, 2000). The first test presented in Panel A of Table **4** is based on firm size, and splits firms into terciles based on their market capitalization. Since we focus on investigating heterogeneity across firm characteristics, we report our results with $\%IO$ as the dependent variable. The baseline specification in column 1 shows that relationship between social connectedness and investments is indeed highest for the smallest firms, and smallest for the largest firms. Those differences are highly statistically significant. In columns 2-4, we add our standard controls for the geographic distance between the locations, interacted with fixed effects for firm size tertiles. Our basic results are not affected by the addition of these control variables.

[Insert Table 4 near here]

We also split firms based on their analyst coverage, as analysts are an important information intermediary in financial markets.[6] These results are reported Panel B of Table **4**. Column 1 shows the baseline

[6]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.

results. We find that investments in low coverage respond twice as much to social connectedness as investments in high coverage firms. After including controls for geographic distance, interacted with coverage indicators, we find that our results remain robust. Taken together, our firm heterogeneity tests suggest that the investment in small firms with little analyst coverage are most strongly affected by the social connectedness to potential investors.[7]

3 Capital Market Implication for Firms

In the previous section, we established that institutional investors are more likely to invest in firms that are located in counties that the investors are socially connected to. We next show that this effect is large enough to generate differences in the equilibrium capital market outcomes of firms that are located in counties that are socially connected to regions with many large institutional investors. We explore three sets of capital market outcomes. We first show that firms that are more socially proximate to capital have more institutional ownership in equilibrium. We then document effects of social proximity to capital on firm valuation and liquidity. We also explore firm-level heterogeneity and show that these effects are larger for smaller and more informationally opaque firms.

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:

$$\textit{Social Proximity to Capital}_{i,t} = \sum_{j} AUM_{j,t} \times \textit{Social Connectedness}_{i,j}, \tag{5}$$

where $AUM_{j,t}$ is the total asset under management by fund families headquartered in county j in quarter t, and $\textit{Social Connectedness}_{i,j}$ is the social connectedness between counties i and j as defined in equation 1. Increases in this measure reflect that county i (and therefore any firm headquartered in that county) has close social connections to counties that are home to institutional investors with high AUM.[8] Figure 3 shows a heat map of this measure of *Social Proximity to Capital* across U.S. counties.[9] As expected, counties located on the East coast, especially those in the Northeast, have the highest levels of *Social Proximity to Capital*, while counties in the mid of the country tend to have lower *Social Proximity to Capital.* However, consistent with the evidence above that neighboring counties can have very differently structure social networks, we find that the *Social Proximity to Capital* can also vary across physically proximate counties – this allows us to include state fixed effects in our regressions below, and thereby only exploit within-state variation in *Social Proximity to Capital*.

[7] 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 sizes. 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 Table IA.6. We find that even controlling for firm size, firms with low analyst coverage still seem to have their investments vary more with social connectedness.

[8] Since we found that social connectedness had the largest effect on the investments of dedicated investors, we also explored separate measures of social proximity to dedicated capital. However, funds of the different classifications are located in similar counties. 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.

[9] A data set with each county's social proximity to capital can be found on the authors' websites.

[Insert Figure 3 near here]

Analogously, we construct a measure of a county's for the physical distance to capital:

$$\textit{Physical Proximity to Capital}_{i,t} = \sum_j AUM_{j,t} / \textit{Distance}_{i,j}, \tag{6}$$

where $\textit{Distance}_{i,j}$ is the physical distance between counties i and j. We divide by distance in this equation to reflect that firms' access to institutional capital reduces with distance. Due to the strong relationship between social connectedness and physical distance shown above, we also find that *Social Proximity to Capital* and *Physical Proximity to Capital* have a correlation coefficient 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 equilibrium effect on firms' overall institutional ownership. We analyze this question using the following panel regression:

$$TIO_{i,t} = \beta \log(\textit{Social Proximity to Capital}_{i,t-1}) + \gamma X_{i,t-1} + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \tag{7}$$

where $TIO_{i,t}$ represents the total institutional ownership share in firm i quarter t, and $X_{i,t-1}$ is a set of standard control variables that have been shown to affect a firm's institutional ownership (Baik et al., 2010; Green and Jame, 2013). These include the physical proximity to capital, book-to-market ratio, log equity market capitalization, 12-month momentum, equity turnover, the price at the beginning of the quarter, R&D expenses (scaled by market capitalization), dividend yield, an S&P 500 Index dummy, and a dummy indicates if a firm has been public for more than 5 years. We also include time fixed effect, state fixed effect, and industry fixed effect. The standard errors are double clustered by quarter and firm.

The results are reported in Table 5. In Columns 1 and 2, we examine the relationship using the whole sample. The key difference is that column 2 includes industry fixed effects to control for across-industry differences in institutional ownership. In both specifications, *Social Proximity to Capital* is significantly related to firms' institutional ownership: a 10% increase in *Social Proximity to Capital* is associated with a 15-25 bps increase in overall institutional ownership, relative to a baseline mean of 60%.

[Insert Table 5 near here]

We next investigate whether this relationship differs across firms characteristics. We documented above that social connectedness is particularly important for attracting investments to small firms and firms with low analyst coverage. Consistent with our previous results, columns 3 and 4 show that *Social Proximity to Capital* has the largest effect on institutional ownership among smaller firms, while institutional ownership in the largest tercile of firms are not significantly affected by social proximity to capital. Similarly, in columns 5 and 6 we find that the effect of *Social Proximity to Capital* on institutional investment share is most important for firms with low analyst coverage.

3.2.1 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 of 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 narrow investor base facilitates less risk sharing, and the higher levels of risk lead to lower valuation and higher cost of capital. This theory has found strong support in the subsequent empirical studies (e.g., Fang and Peress, 2009; Lehavy and Sloan, 2008). We hypothesize that because of higher investor recognition, firms with higher social proximity to capital that are more widely held by institutional investors are likely to have higher valuation. A second mechanism through which social proximity to capital might raise valuations is driven by short-sale constraints (e.g., Scheinkman and Xiong, 2003). In this story, investors are more likely to look at firms they are aware of – while they might be equally likely to find the firm to be overvalued or undervalued, short-sale constraints mean that the views of the more pessimistic investors are less likely to be reflected in market prices, and valuations are therefore higher the more investors consider a firm. Additionally, the resulting overvaluation could persist when the investors' beliefs oscillate with the future arrival of new information. To explore whether social proximity to capital affects firm valuations, we run out the following regression:

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

where $Valuation_{i,t}$ represents the market valuation of firm i in quarter t. We consider two measures of valuation, Tobin's Q and market-to-book to ratio.[10] The dependent variables are in log form following Green and Jame (2013). We include the same set of controls as in Table **5**.

The results are reported in Table **6**. The dependent variable in Panel A is the log 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. In the baseline specification reported in column 1, a 10% increase in *Social Proximity to Capital* is associated with a 0.5%-1% increase in the market-to-book ratio. This effect is statistically significant at the 1% level. The effects of *Social Proximity to Capital* on market-to-book ratios appear to be concentrated in the small and mid market cap firms. Our back-of-envelope calculation suggests that if Caterpillar Inc. moves from Deer Field, IL to Chicago, IL, its market capitalization will increase by 11 percent to 22 percent.[11]

[Insert Table 6 near here]

Our second measure of firm valuation is Tobin's Q. Columns 1 and 2 of Panel B report the whole sample results. We find that in the baseline specification, a 10% increase in *Social Proximity to Capital* is

[10] 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.

[11] This is likely an upper bound, since we use the unconditional coefficient for this calculation. However, we show that the effect is likely weaker among larger firms like Caterpillar. Using the coefficient for large firms, the estimated increase in firm value is between 2 percent and 9 percent.

associated a 0.4% - 0.8% increase in Tobin's Q. To get a sense of magnitude, a one standard deviation increase in firm R&D is associated with a 7.7% increase in Tobin's Q, consistent with the findings reported in Habib and Ljungqvist (2005). In columns 3 and 4, we report differential effects of social proximity to capital for firms of different sizes. We find that the effect of social proximity of capital is significant among low and mid market cap firms. This effect is essentially nonexistent among the largest firms. The differences between the effects on small and large firms are statistically significant. Columns 5 and 6 report the results separately for firms with differential analyst coverage. Again, we find that the effect of social proximity to capital is generally stronger among firms with lower analyst coverage, though the differences across terciles of analyst coverage are less statistically significant.

3.3 Social Proximity to Capital and Firm Liquidity

We next examine the impact of social proximity to capital on firms' liquidity. We also explore the effect of social proximity to capital on the liquidity of firms. 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 capital will have higher liquidity. This prediction is based on prior work that shows that stocks of firms with higher investor recognition (for example, because of more fluent names) are more liquid (see Green and James, 2013). We first show that social proximity to capital does indeed raise firms' liquidity on average. To provide further evidence for our story, we also show that the effect of social proximity to capital varies with shocks to the locations of socially proximate investors. In particular, we show that those firms that were socially proximate to capital in the Tri-State area had relatively lower liquidity during the period when Hurricane Sandy prevented many investors in that region to show up to work. We conduct the following regression analysis:

$$\log(Liquidity)_{i,t} = \beta \log(Social\ Proximity\ to\ Capital_{i,t-1}) + \gamma X_{i,t-1} + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \quad (9)$$

where $Liquidity_{i,t}$ represents secondary market liquidity measures of firm i in quarter t. The dependent variables are in log form following Green and Jame (2013). The other controls follow the specifications in Tables 5 and 6. We consider two firm-level measures of liquidity: the effective percentage spread and the illiquidity measure proposed by Amihud (2002).[12]

[Insert Table 7 near here]

The regression results are reported in Panel A of Table 7. Column 1 shows that *Social Proximity to Capital* is indeed strongly negatively associated with the Effective Spread. The coefficient suggests that a 10% increase in *Social Proximity to Capital* is associated with a 0.6% reduction in percentage effective spread. The coefficient remains significant at the 10% level when we add industry fixed effects in column 2. To put these coefficients in perspective, we compare the economic magnitude of the coefficient of Log Social Proximity to Capital with the coefficient for turnover. Turnover is the second strongest

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

determinant of effective spread after market capitalization. We find that one standard deviation increase in turnover leads to 20% decrease in effective spread, as indicated in Column 1 of the regression, while one standard deviation change in *Log Social Proximity to Capital* is associated with 13% decrease in percentage effective spread.

Next, we interact Log Social Proximity to Capital with firm split by size and by analyst coverage. The interaction with size tercile indicators are reported in Columns 3 and 4. We find that the relation between social proximity to capital and spreads are concentrated in small firms than it is for the largest firms. Similarly, we find that the effect of Social Proximity to Capital on liquidity strongest among the small firms. In the specification without industry fixed effect control, we also find that the effect is significant for mid size firms at the 5% significance level. However, in both Columns 5 and 6, we find no evidence that high analyst coverage firms are affected. We report the associated F-statistics on the heterogeneity of the interaction coefficients. We find that in both specifications, we strongly reject the null of no heterogeneity.

We repeat the same set of tests with Amihud (2002) illiquidity measure as the dependent variable in Panel B of Table 7. In the whole sample analyses reported in Column 1, we find that social proximity to capital is significantly negatively related to firms' illiquidity. A 10% increase in Social Proximity to Capital is related to a 1.8% decrease in illiquidity. This effect is robust after adding industry fixed effect as additional controls. The magnitude of the effect rmains strong: a one standard deviation change in turnover leads to 40% decrease in the Amihud illiquidity measure, while one standard deviation change in Log Social Proximity to Capital 18% decrease in the Amihud measure.

We next explore interaction between Log Social Proximity to Capital and two measures of firm opacity. The interaction with firm size indicators are reported in Columns 3 and 4. We find that the effect is most significant in small firms. Meanwhile, we cannot find statistical support that the effect of social proximity to capital exists for large cap firms. In both specifications, we find significant support for heterogeneous effect in the social proximity to capital. We next examine the heterogeneity across analyst coverage terciles. These results are reported in Columns 5 and 6. Similar to the size split, we find that the liquidity increase associated with high social proximity to capital is mainly present in the the low and mid coverage firms. We report the F-statistics associated with these heterogeneity test and strongly reject the null of no heterogeneity.

In summary, we find that firms' social proximity to capital is negatively associated with spread and illiquidity. Additionally, these effects are strongest for small firms and firms with low analyst coverage. These results suggest that institutional attention in firms with high social proximity may lead to higher liquidity, with the strongest effects for firms that have relatively opaque information environment.

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 control variables at the firm level that have been shown to influence our outcome variables of interest. We also include county-level controls that might be correlated with both social distance to capital and the characteristics of local firms. In the end, we are not aware of any omitted variables that can jointly explain our findings, in particular the various heterogeneities results. For example, if firms in counties that were socially more proximate to capital were better 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 above. It is also unclear why being socially proximate to capital would affect the fundamental quality of only small and informationally opaque firms. Nevertheless, we next provide additional evidence for our proposed explanation.

Hurricane Sandy and Stock Liquidity: A Case Study. One concern with the specifications above 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 does so more for more informationally opaque firms). To provide further evidence for our interpretation of the findings, we next explore the response of firms' liquidity to a shock to investors in socially connected counties.

Specifically, we explore the temporary shock to East Coast-based investors during Hurricane Sandy in 2012, which caused nearly 70 billion US dollars. 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 area. The most direct impact to financial markets was a two-day closure of NYSE and Nasdaq. They reopened on October 31, 2012. However, 6 million people were still without power in 15 States and the District of Columbia as of that date. In the aftermath of the storm, many employees are still 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. Because a large number of important capital providers are located in the the affected area, one might expect changes to the liquidity of firms that are particularly connected to the affected areas.

In our baseline analysis, we define the area affected by Sandy to be the Mid-Atlantic states (NY, NJ, CT, DC, PA, DE, MD, VA, and WV), though our results are robust to broader or narrower definitions. In our 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:

$$\log(Spread_{i,t}) = \beta Affected\ Ratio_i \times I(Sandy)_t + \gamma X_{i,t} + \epsilon_{i,t}. \tag{10}$$

The dependent variable, $Spread_{i,t}$, is firm i's percentage effective spread on day t. The key dependent variable, *Affected Ratio* is defined as the ratio of Social Proximity to Capital in the affected area (i.e., the Mid-Atlantic states) to the overall Social Proximity to Capital, as measured in the quarter before Hurricane Sandy:

$$Affected\ Ratio_i = \frac{\sum_{k \in Mid-Atlantic} AUM_j \times Social\ Connectedness_{i,j}}{\sum_j AUM_j \times Social\ Connectedness_{i,j}}, \tag{11}$$

In other words, *Affected Ratio* measures the cross-sectional exposure of firms to the capital in the affected area. $I(Sandy)_t$ is an indicator variable that equals 1 during the Sandy period, defined as the Sandy period as Oct, 22 to Nov, 07.[13] X represents a vector of control variables, including those controls included in the Table 7. Additionally, we include firm fixed effects and day fixed effects, which remove

[13]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.

any direct effect of *Affected Ratio* on liquidity. In addition, our specifications include day $\times$ state $\times$ industry fixed effects, $I(Sandy)_t \times Control$, and $I(Sandy)_t \times Calendar\ Month\ FE$.

[Insert Table 8 near here]

Our results are reported in Table 8. Our variable of interests is $Affected\ Ratio_i \times I(Sandy)_t$, which captures how the cross-sectional difference in exposure to institutional capital affected by Sandy leads to differential changes in the spread during the Sandy period. We start with a sample from Jan 2012 to July 2013. In Column 1, we only control for day and firm fixed effects. The coefficient is positive and significant at the 5% level. This result is consistent with spreads of firms with higher exposure to institutional capital affected by Sandy widening more during the Sandy period than the other firms. The economic magnitude of the effect is large. For example, as implied by the coefficient reported in Column 1, firms with a one standard deviation increase in Affected Capital Ratio experienced a 5% additional increase in their effective spreads.

In Column 2, we additionally include day$\times$ state $\times$ industry FE. This control helps us rule out the possibility that our result is purely driven by geographic proximity to affected area or our result is driven by certain industries having strong connection to the affected area. Our result remains robust after including these fixed effects. In Column 3, while controlling for Day and Firm FE, we further include additional control variables from Table 7, each interacted with $I(Sandy)_t$. This specification allows us to rule out other firm characteristics that might be correlated with social distance to capital are driving our results. The coefficient is essentially unaffected by these controls. In Column 4, we modify the specification in Column 2 by adding additional controls and their interaction with Sandy indicator and find that the result remains qualitatively similar.

Next, we extend our sample period to Jan 2011 to Dec 2014. The purpose of this extension is that it allows us to further control for seasonality effects. For example, we would like to rule out the possibility that firms in locations that are socially connected to the areas affected by Sandy always trade at higher spreads every October. Extending the sample allows us to better estimate the seasonality effect. We first repeat our analysis in Column 4 in the extended sample. As shown in Column 5, our result does not change materially as we extend the sample. Next, we include based on the specification in Column 5, we include Calendar Month FE$\times$Affected Capital Ratio. We report this result in Column 6. In this model, our point estimate is 0.215 and significant at the 5% level.

In summary, we show that shocks to institutional capital widens the spread of socially proximate firms. This provides strong evidence that the cross-sectional relationship between social proximity to capital and firm liquidity that we documented above is not driven by omitted firm-level variables that correlate with both firm liquidity and social proximity to capital.

4 Implications for Institutional Investors

In the final section of the paper, we examine if institutional investors' overweighting of socially connected stocks are driven by information or familiarity bias. If these overweighting are driven by institutional investors' information advantage through the social network, we expect that institutional investors that hold more socially connected stocks are likely to outperform. On the other hand, if the overweighting is driven by familiarity bias, we expect that institutional investors that put more weight on socially connected stocks will not outperform, or even underperform.[14]

We explore the implications for institutional investors of investing in socially connected stocks in two ways. In the first approach we split investors by their propensity to hold socially connected stocks. In the second approach, we look within investors, and explore the extent to which their holdings in areas they are socially connected to outperform otherwise similar holdings in areas that they are not connected to. We estimate an institution's preference for socially connected stocks using the following regression:

$$Investment_{i,j} = \alpha + \beta_j \log(Social\ Connectedness_{i,j}) + \gamma X_{i,j} + \psi_i + \xi_{j \times ind(i)} + \epsilon_{i,j}.$$

This regression corresponds to the specification of column 5 in Table **2**, except that we allow the propensity of disproportionately hold socially connected stocks to vary across institutions. Our estimates of β_j goes from, on average, -11.26 in the lowest decile to 42.37 in the highest; this suggests that, while most investors disproportionately invest in firms located in areas they are socially connected to, there is substantial heterogeneity in this propensity. We then sort institutions into deciles based on their β_j. We calculate the fund-level returns as follows: $Return_{i,t} = \sum_j w_{i,j,t-1} Return_{j,t}$, where the weighting are based on the value of the holdings. We then aggregate these portfolio returns into decile portfolio returns using both equal-weighted and value-weighted averages. In table **9** Panel A, We report a range of metrics for institutional investors' returns based on position-level measures $Return_{j,t}$ that include excess returns, CAPM alpha, Fama-French three-factor alpha, Fama-French five-factor alpha, as well as their volatility and Sharpe ratios.

Overall, there is no strong evidence that institutional investors that overweight socially connected stocks outperform. None of the alphas of the long-short portfolio by taking a long position in decile 10 and a short position in decile 1 are statistically significant from zero. This result suggests that institutional investors do not gain a systematic information advantage in their investment in socially connected firms. We also investigate whether firms that put more weight into socially connected stocks are under diversified. We report the corresponding volatility of excess returns for each decile portfolio. We find that those with portfolios with a higher propensity to hold socially connected stocks tend to have lower return volatility.

In the second test, we split all institutions into their Bushee types. As shown in Section 2.1, dedicated institutions are more likely to invest in firms located in socially connected firms, which is consistent with Bushee (2001). Thus, these institutions may be able to better use the information from social connection to generate returns. In Panel B of Table **9**, we report the long-short portfolio performance for three Bushee institution types. We do not find significant alpha in any of the classification.

[Insert Table 9 near here]

We conduct a third test to examine whether institutional investors are better at picking firms that they are more socially connected with. This test is motivated by Coval and Moskowitz (2001), who

[14] Underperformance may be due to institutional investors substituting their value creating stock picking ability with relying on the non-value-creation activity of buying socially connected stocks.

show that investors are better at picking stocks of the firms that are geographically close by. In the same spirit of Coval and Moskowitz (2001), we sort the firm county-institution county pairs based on their *Social Connectedness* into tercile in Table **10**. We report the performance metrics, including excess returns, CAPM, Fama and French (1996) three-factor, Fama and French (2015) five-factor alphas for each social connectedness tercile. We also report the performance metrics for the long-short portfolio by taking long positions in socially connected stocks and short positions in stocks with social connectedness. Overall, we do not find any evidence that socially connected stocks picked by institutions outperform the their non-connected stock picks.

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), who show that mutual funds may have information advantage in local stocks, we find no evidence that institutional investors have an information edge through social connections. As a result, our results suggest that institutional investors' investment in socially connected firms is primarily driven by their familiarity bias.

5 Conclusion

There is a growing literature that explores a variety of explanations for geographic disparities of economic outcomes within the United States. We attempt to contribute to this literature by investigating how the geographic structure of social networks of different regions shapes the allocation of capital to firms, and thereby contributes to the differences in firm outcomes (Baum-Snow and Pavan, 2011; Moretti, 2012; Dougal et al., 2018). We find that institutional investors invest more in firms located in regions to which they have stronger social ties, even after controlling for physical distance. As a result, firms in regions that are more socially proximate to institutional capital have higher liquidity and higher valuations. These effects are more prevalent among small and informationally opaque firms. We argue that differences in the social proximity to capital can be an important channel through which regional characteristics affect the 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.

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

Bailey, Warren, Alok Kumar, and David Ng, 2011, Behavioral biases of mutual fund investors, *Journal of Financial Economics* 102, 1–27.

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

Chen, Hailiang, Prabuddha De, Yu Jeffrey Hu, and Byoung-Hyoun Hwang, 2014, Wisdom of crowds: The value of stock opinions transmitted through social media, *The Review of Financial Studies* 27, 1367–1403.

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.

Cookson, J Anthony, and Marina Niessner, 2018, Why don't we agree? evidence from a social network of investors, *Evidence from a Social Network of Investors (August 6, 2018)* .

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.

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, 1996, Multifactor explanations of asset pricing anomalies, *The journal of finance* 51, 55–84.

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.

Gilbert, Eric, and Karrie Karahalios, 2009, Predicting tie strength with social media, in *Proceedings of the SIGCHI conference on human factors in computing systems*, 211–220, ACM.

Gompers, Paul, Joy Ishii, and Andrew Metrick, 2003, Corporate governance and equity prices, *The quarterly journal of economics* 118, 107–156.

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.

Hampton, Keith, Lauren Sessions Goulet, Lee Rainie, and Kristen Purcell, 2011, Social networking sites and our lives, *Pew Internet & American Life Project* 16, 1–85.

Heimer, Rawley Z, 2016, Peer pressure: Social interaction and the disposition effect, *The Review of Financial Studies* 29, 3177–3209.

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, 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.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of finance* 48, 65–91.

Jones, Jason J, Jaime E Settle, Robert M Bond, Christopher J Fariss, Cameron Marlow, and James H Fowler, 2013, Inferring tie strength from online directed behavior, *PloS one* 8, e52168.

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

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).

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.

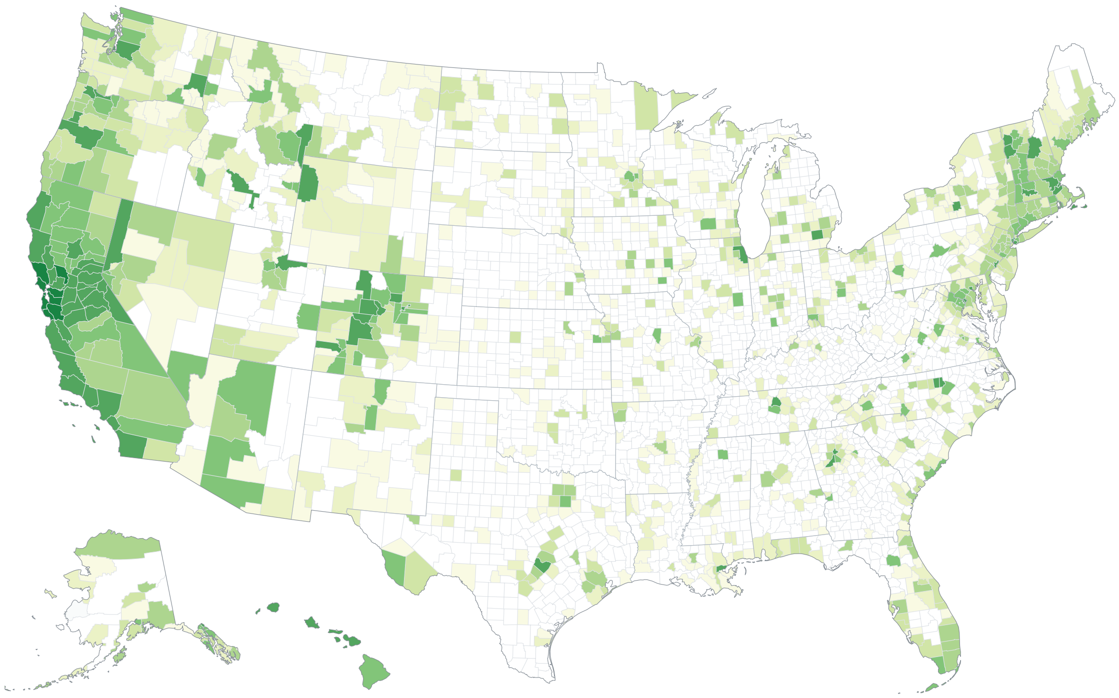
Rehse, Dominik, Ryan Riordan, Nico Rottke, and Joachim Zietz, 2019, The effects of uncertainty on market liquidity: Evidence from hurricane sandy, *Journal of Financial Economics* .

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

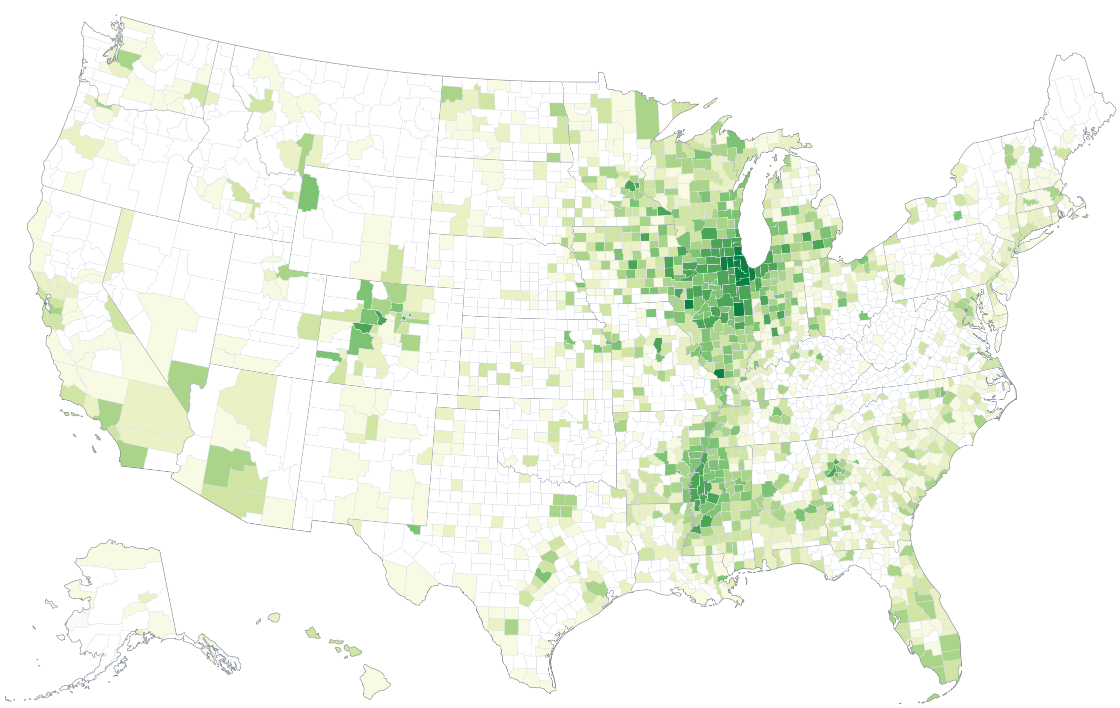
Scheinkman, Jose A, and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of political Economy* 111, 1183–1220.

Figure 1: Examples of Social Connectedness

Panel A: San Francisco County, CA



Panel B: Cook County, IL (Chicago)



Note: Figure shows heat maps of the social connectedness to San Francisco County, CA, and Cook County, IL. Darker colors correspond to more socially connected regions.

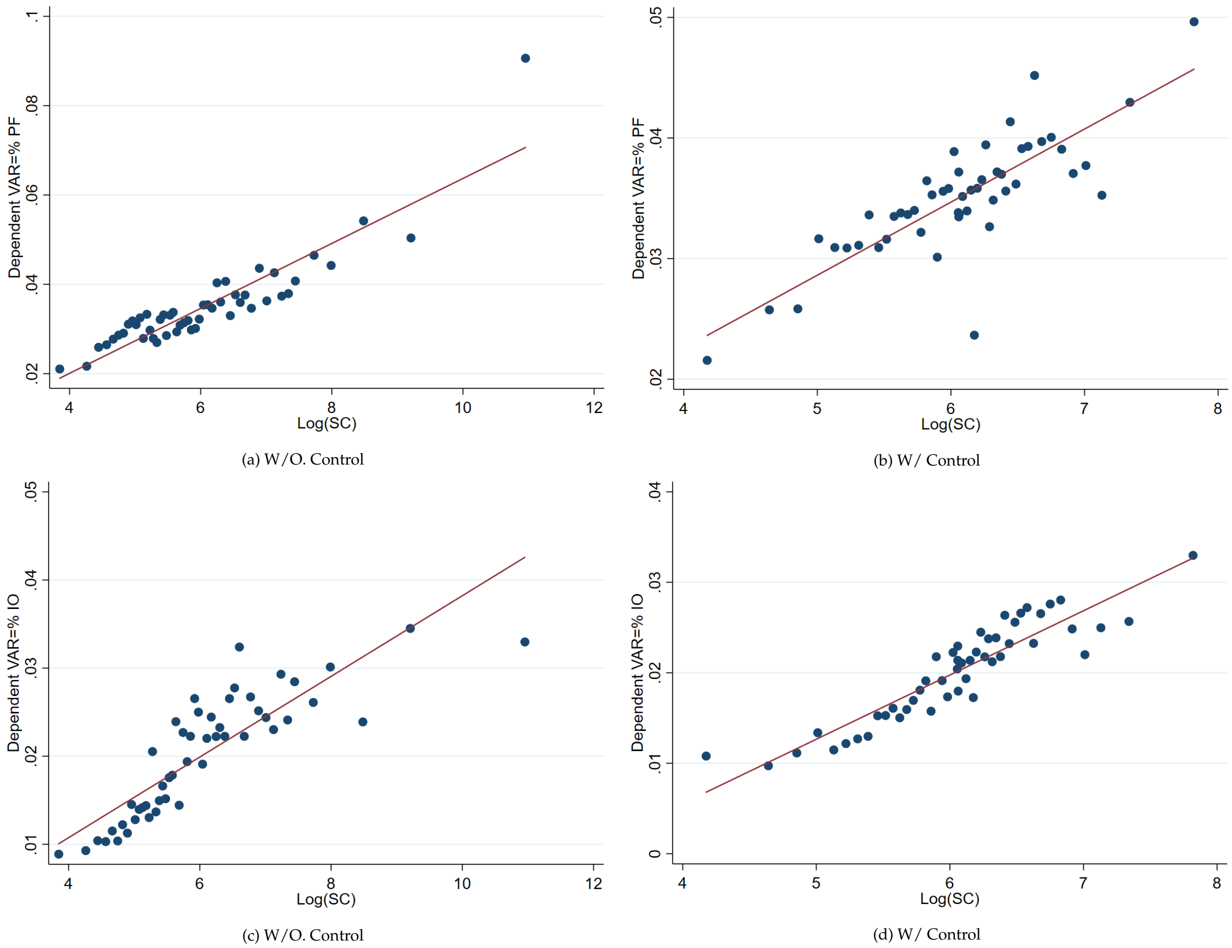


Figure 2: Bin Scatter Plot of % PF and % IO on SC on Social Connectedness

The independent variable log of Social Connectedness (SC), as of Jun 2016, is independently sorted into 50 bins. For each bin, the mean of Log(SC) and conditional mean of the dependent variable, % PF in (a) and (b) and % IO in (c) and (d), is plotted as a scatter point. And each figure also plot OLS-fitted line. The figure in the left column shows univariate analysis and figure in the right controls same county dummy, same state dummy, and 500-tile distance dummies.

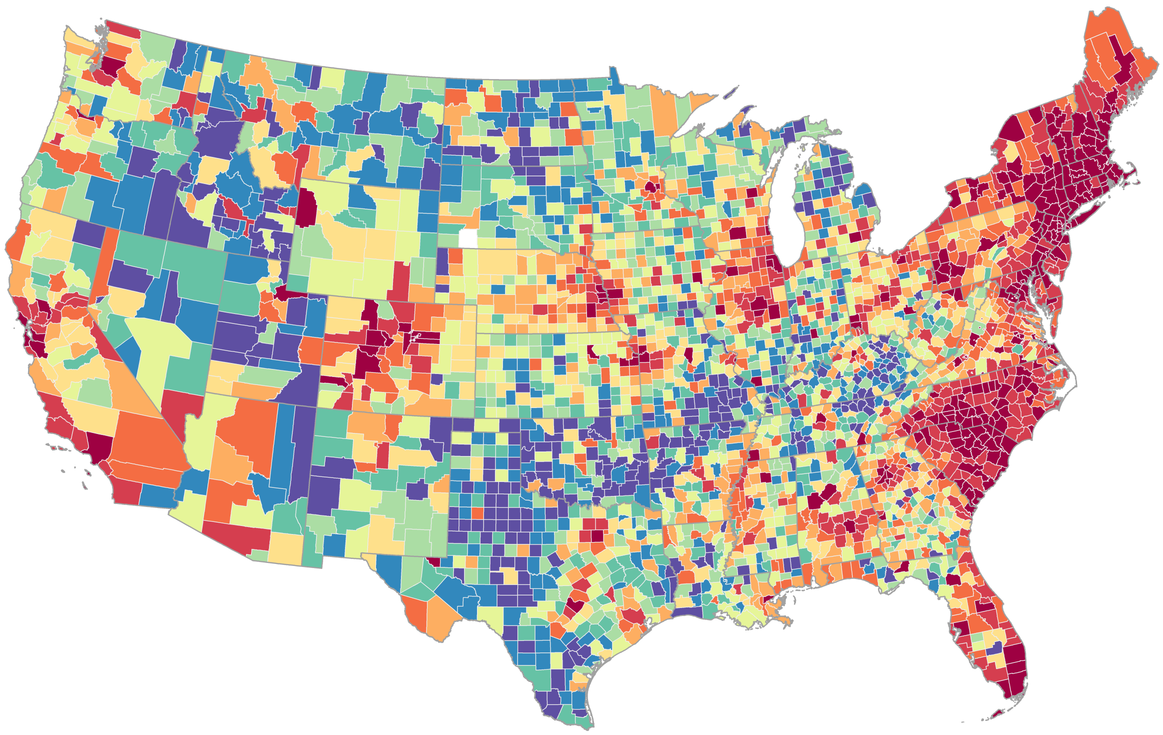


Figure 3: Heat Map of Social Proximity to Capital

Note: Social Proximity to Capital of county j is defined as $\sum_j \mathit{County\ AUM}_i \times \mathit{Social\ Connectedness}_{i,j}$. This figure plots the heat map of Social Proximity to Capital across all the U.S. counties as of Jun, 2016. Red indicates high levels of Social Proximity to Capital and blue indicates low Social Proximity to Capital

Table 1: Summary Statistics

We report summary statistics for key variables in this table. Variables at the firm-institution level are exhibited in Panel A. These variables are captured in June 2016. Social Connectedness is defined as number of Facebook links between a firm's headquarter county and an institution's headquarter county, scaled by the product of the populations in these two counties (multiplied by 10^{12}). Log Social Connectedness is defined as $Log(1 + Social\ Connectedness)$. Distance is the distance in miles between two counties. Log Distance is defined as $Log(1 + Distance)$. % Portfolio (% PF) is the percentage of AUM allocated to a stock, where AUM is measured by the value of the institution's equity holdings. % IO measures the percentage ownership by an institution, defined as number of shares owned by an institution, divided by the firm's total shares outstanding. Panel B includes the firm-level variables. Our firm-level sample covers 2007:Q2 to 2016:Q4. Social Proximity to Capital is $Log(\sum(CountyAUM \times Social\ Connectedness))$, where County AUM is measured in million US dollars and the summation is taken across all U.S. counties. Physical Proximity to Capital is $Log(\sum(\frac{County\ AUM}{1+Distance})$, where Distance is the number of miles between a firm's headquarter county and a U.S. county and the summation is taken across all U.S. counties. The Excess Return is the quarterly average of monthly return over the risk-free rate. ILLIQ is the Amihud (2002) illiquidity at the quarterly horizon (scaled up by 10^6). Effective Spread is dollar-weighted percent effective spread (scaled up by 10^3). Tobin's q is the ratio between the market value of the firm over the replacement cost of its assets. M/B is defined as the ratio of market value of equity over the book value of equity. Size is the firm market cap. in million dollars. Number of Analysts is the number of analysts who provides the one-year ahead earnings estimates. Panel C includes information on institutional investors. Institutional Classification is from Bushee (2001). The statistics reported in this Panel are cross-sectional summary statistics averaged across all quarters. Our sample includes institutions and firms located in the mainland U.S. excluding those located in State PR, VI, AK, and HI. We also require the institution to at least hold five different stocks. We study common stocks listed on NYSE, NASDAQ, and NYSE MKT (formerly AMEX). We exclude penny stocks (Price$<$ \$5).

Panel A: Summary Statistics for Institution-Firm Pair Variables 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
% IO	0.02	0.32	0	0	0	0	0.00

Panel B: Firm Level Summary Statistics from 2007 to 2016

Variables	MEAN	ST. DEV	P5	P10	MEDIAN	P90	P95
Excess Return	0.51	7.49	-11.73	-7.90	0.71	8.37	11.57
Log(ILLIQ)	-5.38	3.16	-9.75	-9.06	-5.79	-1.07	1.06
Log(Effective Spread)	1.13	1.16	-0.48	-0.26	0.97	2.81	3.31
Log(Tobin's Q)	0.44	0.56	-0.15	-0.06	0.30	1.22	1.55
Log(M/B)	0.70	0.89	-0.49	-0.25	0.57	1.79	2.25
Log(Social Proximity to Capital)	23.08	1.13	21.49	21.82	22.94	24.51	25.65
Log(Physical Proximity to Capital)	10.93	1.47	9.09	9.35	10.65	13.01	14.37
Size	5,545.63	21,605.36	51.47	88.25	805.06	10,142.54	22,180.95
Number of Analysts	7.73	7.40	0	0	6	18	23

Table 1: (Continued)

Institution Type	N	AUM (Millions) Mean	ST. DEV	P5	P10	Median	P90	P95
Dedicated	63.82	4296.18	14210.15	62.37	100.11	665.07	8128.38	14725.09
Quasi-Index	1551.89	4521.17	31326.02	29.90	56.21	252.87	4193.07	13161.83
Transient	733.39	2795.77	23819.30	19.43	38.06	309.04	3799.65	8497.47
Not Identified	179.05	2563.48	14564.79	7.78	14.50	88.59	2528.70	9010.64

Table 2: **Social Connectedness and Institutional Investment**

This table shows the results on how social connectedness affects institutional investors' portfolio decisions. Our sample includes all firm-institution pairs at June 2016. The dependent variables are % PF (Panel A) defined as the percentage of AUM allocated to a stock and % IO (Panel B), which measures the percentage ownership by an institution. If an institution does not report holding in a given firm, boith % PF and % IO are assigned to be 0. Social Connectedness is defined as number of Facebook links between a firm's headquarter county and an institution's headquarter county, scaled by the product of the populations in these two counties (multiplied by 10^{12}). Log Social Connectedness is defined as $Log(1 + Social\ Connectedness)$. We also consider the Social Connectedness Quintiles indicators as independent variables. Log Distance is $Log(1 + Distance)$, where Distance measures the distance in miles between the firm's headquarter county and the institution's headquarter county. We consider Firm, Institution$\times$Industry, Distance 500-tile, Same County, and Same State fix effects. Same County (Same State) is a dummy equal to one if the institution and the firm are located in the same county (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. 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: % PF (Scaled by Fund AUM)

	(1)	(2)	(3)	(4)	(5)	(6)
Log Social Connectedness	8.158***		9.171***	11.986***	11.624***	
	(10.71)		(7.85)	(9.44)	(8.92)	
Log Distance		-4.960***	0.909			
		(-8.69)	(1.06)			
Social Connectedness Quintile = 2 (Low)						1.410*
						(1.88)
Social Connectedness Quintile = 3						4.519***
						(4.51)
Social Connectedness Quintile = 4						7.844***
						(4.93)
Social Connectedness Quintile = 5 (High)						15.081***
						(6.77)
Firm FE	YES	YES	YES	YES	YES	YES
Institution $\times$ Industry FE	YES	YES	YES	YES	YES	YES
Distance 500-tile FE	NO	NO	NO	YES	YES	YES
Same State FE	NO	NO	NO	NO	YES	YES
Same County FE	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
R2	0.072	0.072	0.072	0.073	0.073	0.072

Table 2: (Continued)

Panel B: % IO (Scaled by Firm Value)

	(1)	(2)	(3)	(4)	(5)	(6)
Log Social Connectedness	2.185***		2.894***	3.481***	3.488***	
	(8.79)		(6.92)	(8.10)	(7.82)	
Log Distance		-1.216***	0.636**			
		(-6.93)	(2.20)			
Social Connectedness Quintile = 2 (Low)						0.093
						(0.31)
Social Connectedness Quintile = 3						1.801***
						(4.36)
Social Connectedness Quintile = 4						3.098***
						(5.52)
Social Connectedness Quintile = 5 (High)						6.006***
						(6.97)
Firm FE	YES	YES	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES	YES	YES
Distance 500-tile FE	NO	NO	NO	YES	YES	YES
Same State FE	NO	NO	NO	NO	YES	YES
Same County FE	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
R2	0.305	0.305	0.305	0.305	0.305	0.305

Table 3: Heterogeneity across Institutional Type

This table shows how the effect of social connectedness varies with institutional characteristics. The dependent variables is % PF, defined as the percentage of AUM allocated to a stock. In Panel A, we report the heterogeneity across institution types based on Bushee (2001) classification. We interact Log Social Connectedness with Bushee classification dummies (Dedicated, Quasi-Index, and Transient). Institutions without a Bushee classification are dropped from this sample. In Panel B, we report the heterogeneity across institution sizes. We interact Log Social Connectedness with dummy variables based on institution AUM tercile. Log Social Connectedness is defined as $Log(1 + Social\ Connectedness)$, where Social Connectedness is defined as number of Facebook links between a firm's headquarter county and an institution's headquarter county, scaled by the product of the populations in these two counties (multiplied by 10^{12}). We consider Firm, Institution×Industry, Distance 500-tile, Same County, and Same State fix effects, interacted with the respective institutional characteristics dummies. Same County (Same State) is a dummy equal to one if the institution and the firm are located in the same county (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. 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: Bushee (2001) Classification

	(1)	(2)	(3)	(4)
Heterogeneity across Institutional Type	% PF			
Transient×Log Social Connectedness	4.586***	3.933***	5.438***	5.110***
	(5.72)	(3.26)	(4.31)	(3.87)
Quasi-Index×Log Social Connectedness	9.178***	10.754***	13.161***	12.921***
	(11.26)	(8.95)	(8.91)	(8.46)
Dedicated×Log Social Connectedness	18.515***	21.569*	29.725**	31.743**
	(2.87)	(1.73)	(2.00)	(2.01)
Institution Type FE × Firm FE	YES	YES	YES	YES
Institution × Industry FE	YES	YES	YES	YES
Institution Type FE × Log Distance	NO	YES	NO	NO
Institution Type FE × Distance 500- tile FE	NO	NO	YES	YES
Institution Type FE × Same County FE	NO	NO	NO	YES
Institution Type FE × Same State FE	NO	NO	NO	YES
F stat (Null: No heterogeneity across institutional type)	11.863***	11.048***	9.740***	9.480***
N	7,162,800	7,162,800	7,162,800	7,162,800
R2	0.083	0.083	0.083	0.083

Table 3: (Continued)

Panel B: AUM

Heterogeneity across Institutional Type	**(1)**	**(2)** % PF	**(3)**	**(4)**
Small AUM×Log Social Connectedness	11.962***	11.239***	14.639***	14.125***
	(9.47)	(6.16)	(6.92)	(6.51)
Mid AUM×Log Social Connectedness	7.755***	9.203***	11.714***	11.092***
	(8.53)	(6.27)	(6.80)	(6.37)
Large AUM×Log Social Connectedness	4.619***	6.724***	9.017***	9.048***
	(5.78)	(4.95)	(5.37)	(5.07)
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 stat (Null: No heterogeneity across institutional type)	16.281***	2.943**	2.636*	2.014
N	8,694,060	8,694,060	8,694,060	8,694,060
R2	0.075	0.075	0.076	0.076

Table 4: **Heterogeneity across Firm Types**

This table shows how the effect of social connectedness varies with firm characteristics. The dependent variables is % IO, which measures the percentage ownership by an institution. In Panel A, we report the heterogeneity across firms size. We interact Log Social Connectedness with dummy variables based on firm size tercile. In Panel B, we report the heterogeneity across firm analyst coverage. We interact Log Social Connectedness with dummy variables based on firm analyst coverage tercile. Log Social Connectedness is defined as $Log(1 + Social\ Connectedness)$. Social Connectedness is defined as number of Facebook links between a firm's headquarter county and an institution's headquarter county, scaled by the product of the populations in these two counties (multiplied by 10^{12}). We consider Firm, Institution×Industry, Distance 500-tile, Same County, and Same State fix effects, interacted with the respective firm characteristics dummies. Same County (Same State) is a dummy equal to one if the institution and the firm are located in the same county (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. 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: Firm Size

	(1)	(2)	(3)	(4)
Heterogeneity across Firm Type		% IO		
Small Cap×Log Social Connectedness	2.457***	3.227***	3.889***	3.830***
	(7.48)	(6.25)	(6.38)	(6.08)
Mid Cap×Log Social Connectedness	2.307***	2.638***	3.219***	3.258***
	(6.24)	(4.29)	(4.93)	(4.82)
Large Cap×Log Social Connectedness	1.025***	1.498***	1.680***	1.593***
	(3.72)	(3.36)	(3.22)	(2.95)
Firm FE	YES	YES	YES	YES
Firm Size Tercile×Institution×Industry FE	YES	YES	YES	YES
Firm Type FE × Log Distance	NO	YES	NO	NO
Firm Type FE × Distance 500-tile FE	NO	NO	YES	YES
Firm Type FE × Same County FE	NO	NO	NO	YES
Firm Type FE × Same State FE	NO	NO	NO	YES
F stat (Null: No heterogeneity across firm type)	8.048***	4.142**	4.434**	4.523**
N	8,660,147	8,660,147	8,660,147	8,660,147
R2	0.348	0.348	0.348	0.348

Table 4: (Continued)

Panel B: Analyst Coverage

Heterogeneity across Firm Type	(1)	(2) % IO	(3)	(4)
Low Coverage×Log Social Connectedness	2.747***	3.315***	4.492***	4.533***
	(7.16)	(5.93)	(5.97)	(5.89)
Mid Coverage×Log Social Connectedness	2.117***	2.565***	2.765***	2.650***
	(6.38)	(4.76)	(5.05)	(4.64)
High Coverage×Log Social Connectedness	1.251***	1.523***	1.891***	1.774***
	(4.45)	(3.06)	(3.65)	(3.38)
Firm FE	YES	YES	YES	YES
Firm Type FE×Institution FE × Industry FE	YES	YES	YES	YES
Firm Type FE × Log Distance	NO	YES	NO	NO
Firm Type FE × Distance 500- tile FE	NO	NO	YES	YES
Firm Type FE × Same County FE	NO	NO	NO	YES
Firm Type FE × Same State FE	NO	NO	NO	YES
F stat (Null: No heterogeneity across firm type)	7.03***	3.954**	4.389**	4.717**
N	8,653,981	8,653,981	8,653,981	8,653,981
R2	0.351	0.351	0.351	0.351

Table 5: Table: Firms' Social Proximity to Capital and Institutional Ownership

We study how firms' social connectedness to institutional capital affects Their institutional ownership. The dependent variable is institutional ownership (% TIO), which is percentage of shares held by institutional investors over total shares outstanding . The main independent variable is Social Proximity to Capital, defined as $Log(\sum(County\ AUM \times Social\ Connectedness))$, where County AUM is measured in million U.S. dollars and the summation is taken across all U.S. counties. We control Physical Proximity to Capital, defined as $Log(\sum \frac{County\ AUM}{1+Distance})$, where Distance is the number of miles between a firm's headquarter county and a U.S. county and the summation is taken across all U.S. counties. Additional controls include log of Book-to-Market ratio, the log of firm equity market capitalization, 12-month momentum, equity turnover, the price at the beginning of the quarter, R&D expenses, dividend yield, a Standard & Poor's (S&P) 500 Index dummy, and a firm age dummy. The definitions of control variables are detailed in Table A.1. The regressions also include Quarter, State, and Industry fixed effects. Industry fixed effect is based on Fama and French (1997)'s 48 industry classification. Columns (3) and (4) exhibit heterogeneity across firm size tercile. Firm size terciles are based on NYSE 30 and 70 percentile. Columns (5) and (6) exhibit heterogeneity across analyst coverage terciles, where we rank the firms into terciles based on the number of analysts covering the firm. the 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.

Dep: % TIO	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Social Proximity to Capital	2.565***	1.556**				
	(3.16)	(2.05)				
Low×Social Proximity to Capital			2.571***	1.508*	3.123***	1.729**
			(3.04)	(1.94)	(3.72)	(2.22)
Mid×Social Proximity to Capital			2.590***	1.713**	1.898**	1.275
			(3.02)	(2.08)	(2.30)	(1.61)
High×Social Proximity to Capital			1.545	1.099	1.707*	1.15
			(1.59)	(1.16)	(2.00)	(1.37)
Physical Proximity to Capital	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
QTR FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Industry FE	NO	YES	NO	YES	NO	YES
F-stat (Null: No heterogeneity across terciles)			1.095	0.361	2.969*	0.504
N	96,644	96,644	96,644	96,644	96,644	96,644
R-Squared	0.305	0.366	0.326	0.383	0.344	0.399

Table 6: **Firms' Social Proximity to Capital and Firm Value**

We study how firms' social proximity to institutional capital affects their valuation. The dependent variables are $Log(M/B)$ in Panel A and $Log(Tobin's\ Q)$ in Panel B. The main independent variable is Social Proximity to Capital, defined as $Log(\sum(County\ AUM \times Social\ Connectedness))$, where County AUM is measured in million U.S. dollars and the summation is taken across all U.S. counties. We control Physical Proximity to Capital, defined as $Log(\sum \frac{County\ AUM}{1+Distance})$, where Distance is the number of miles between a firm's headquarter county and a U.S. county and the summation is taken across all U.S. counties. Additional controls include the log of Amihud (2002) illiquidity, the log of firm equity market capitalization, 12-month momentum, equity turnover, the price at the beginning of the quarter, R&D expenses, a Standard & Poor's (S&P) 500 Index dummy, and a firm age dummy. The definitions of control variables are detailed in Table A.1. The regressions also include Quarter, State, Industry, and State×Quarter fixed effects. Industry fixed effect is based on Fama and French (1997)'s 48 industry classification. Columns (3) and (4) exhibit heterogeneity across firm size tercile. Firm size terciles are based on NYSE 30 and 70 percentile. Columns (5) and (6) exhibit heterogeneity across analyst coverage terciles, where we rank the firms into terciles based on the number of analysts covering the firm. the 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: M/B

Dep: Log(Market-to-Book)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Whole Sample		**Split by Size**		**Split by Analyst Coverage**	
Social Proximity to Capital	0.102***	0.049**				
	(4.22)	(2.22)				
Low×Social Proximity to Capital			0.108***	0.053**	0.104***	0.047**
			(4.42)	(2.36)	(4.26)	(2.10)
Mid×Social Proximity to Capital			0.120***	0.062**	0.118***	0.065**
			(4.27)	(2.41)	(4.34)	(2.61)
High×Social Proximity to Capital			0.042	0.010	0.069**	0.025
			(1.21)	(0.33)	(2.28)	(0.91)
Physical Proximity to Capital	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
QTR FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Industry FE	NO	YES	NO	YES	NO	YES
F-stat (Null: No heterogeneity across terciles)			3.631**	2.111	2.369*	2.156
N	102,415	102,415	102,415	102,415	102,415	102,415
R-Squared	0.215	0.327	0.216	0.328	0.217	0.328

Table 6: (Continued)

Panel B: Tobin's Q

Dep: Log(Tobin's Q)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Whole Sample		**Split by Size**		**Split by Analyst Coverage**	
Social Proximity to Capital	0.083***	0.040***				
	(5.13)	(2.81)				
Low $\times$ Social Proximity to Capital			0.082***	0.037**	0.079***	0.033**
			(5.07)	(2.59)	(4.91)	(2.34)
Mid $\times$ Social Proximity to Capital			0.106***	0.058***	0.097***	0.052***
			(5.48)	(3.39)	(5.37)	(3.31)
High $\times$ Social Proximity to Capital			0.043*	0.019	0.068***	0.032*
			(1.92)	(0.99)	(3.31)	(1.81)
Physical Proximity to Capital	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
QTR FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Industry FE	NO	YES	NO	YES	NO	YES
F-stat (Null: No heterogeneity across terciles)			6.269***	3.393**	2.516*	2.258
N	102,414	102,414	102,414	102,414	102,414	102,414
R-Squared	0.211	0.352	0.213	0.354	0.214	0.354

Table 7: Firms' Social Proximity to Capital and Transaction Cost

We study how firms' social connectedness to institutional capital affects their stock liquidity. The dependent variable is Log(Effective Spread) in Panel A, where Effective Spread is the quarterly average of percent effective spread. The dependent variable is Log(ILLIQ) in Panel B, where ILLIQ is the quarterly average of Amihud (2002) illiquidity measure. ILLIQ is defined as daily absolute return, divided by dollar trading volume. The main independent variable is Social Proximity to Capital, defined as $Log(\sum(County\ AUM \times Social\ Connectedness))$, where County AUM is measured in million U.S. dollars and the summation is taken across all U.S. counties. We control Physical Proximity to Capital, defined as $Log(\sum \frac{County\ AUM}{1+Distance})$, where Distance is the number of miles between a firm's headquarter county and a U.S. county and the summation is taken across all U.S. counties. Additional controls include the log of Book-to-Market ratio, the log of firm equity market capitalization, 12-month momentum, equity turnover, the price at the beginning of the quarter, R&D expenses, dividend yield, a Standard & Poor's (S&P) 500 Index dummy, and a firm age dummy. The definitions of control variables are detailed in Table A.1. The regressions also include Quarter, State, Industry, and State×Quarter fixed effects. Industry fixed effect is based on Fama and French (1997)'s 48 industry classification. Columns (3) and (4) exhibit heterogeneity across firm size tercile. Firm size terciles are based on NYSE 30 and 70 percentile. Columns (5) and (6) exhibit heterogeneity across analyst coverage terciles, where we rank the firms into terciles based on the number of analysts covering the firm. the 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: Effective Spread

Dep: Log(Effective Spread)	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Social Proximity to Capital	-0.060***	-0.030*				
	(-3.68)	(-2.02)				
Low×Social Proximity to Capital			-0.089***	-0.053***	-0.095***	-0.055***
			(-5.18)	(-3.39)	(-5.42)	(-3.44)
Mid×Social Proximity to Capital			-0.030*	-0.009	-0.036**	-0.016
			(-1.84)	(-0.60)	(-2.27)	(-1.09)
High×Social Proximity to Capital			0.021	0.030	0.003	0.017
			(0.98)	(1.56)	(0.16)	(1.04)
Physical Proximity to Capital	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
QTR FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Industry FE	NO	YES	NO	YES	NO	YES
F-stat (Null: No heterogeneity across terciles)			19.904***	13.71***	17.142***	11.32***
N	80,046	80,046	80,046	80,046	80,046	80,046
R-Squared	0.816	0.836	0.823	0.842	0.827	0.845

Table 7: (Continued)

Panel B: Amihud (2002) Illiquidity

Dep: Log(ILLIQ)	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample		Split by Size		Split by Analyst Coverage	
Social Proximity to Capital	-0.179***	-0.113***				
	(-4.93)	(-3.25)				
Low×Social Proximity to Capital			-0.222***	-0.147***	-0.246***	-0.160***
			(-5.67)	(-3.95)	(-5.80)	(-3.97)
Mid×Social Proximity to Capital			-0.118***	-0.067**	-0.121***	-0.076**
			(-3.70)	(-2.14)	(-3.93)	(-2.49)
High×Social Proximity to Capital			-0.051	-0.023	-0.081**	-0.042
			(-1.29)	(-0.61)	(-2.48)	(-1.35)
Physical Proximity to Capital	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
QTR FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Industry FE	NO	YES	NO	YES	NO	YES
F-stat (Null: No heterogeneity across terciles)			14.246***	9.034***	12.914***	7.576***
N	101,158	101,158	101,158	101,158	101,158	101,158
R-Squared	0.894	0.902	0.899	0.907	0.900	0.908

Table 8: The Effect of Social Proximity to Capital during Hurricane Sandy

We analyze the cross-sectional difference in the impact of Hurricane Sandy on the liquidity of firms with various levels of social proximity to the institutional capital in the Mid-Atlantic area. The dependent variable is daily percentage effective spread. The key variable of interest is the interaction between Sandy indicator and Affected Capital Ratio. Sandy Indicator equals 1 during the affected period defined as Oct 22, 2012 to Nov 07, 2012. Affected Capital Ratio is defined as the Social Proximity to the institutional capital located in the Mid-Atlantic Area, divided by the Social Proximity to Capital. We fix this ratio in the quarter (2012:Q3) prior to Sandy. We consider the control variables included in Table 7 and their interaction with Sandy indicator. We also control firm fixed effects. In addition, we consider Day, Day×State×Industry fixed effects. The sample for first four columns range from 1/2012 to 7/2013. We use an extended sample fro 1/2011 to 12/2014 in Columns 5 and 6. We also consider the interaction between Affected Ratio and calendar month as additional controls. 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)
Sample Period:		Jan 2012 to Jul 2013			Jan 2011 to Dec 2014	
Dep: Log(Effective Spread)						
I(Sandy)×Affected Capital Ratio	0.225**	0.268**	0.197**	0.237**	0.257**	0.215**
	(2.63)	(2.14)	(2.20)	(2.03)	(2.06)	(2.10)
Control	NO	NO	YES	YES	YES	YES
Sandy FE× Control	NO	NO	YES	YES	YES	YES
Month FE× Affected Capital Ratio	NO	NO	NO	NO	NO	YES
Day FE	YES	NO	YES	NO	NO	NO
Firm FE	YES	YES	YES	YES	YES	YES
Day×State×Industry FE	NO	YES	NO	YES	YES	YES
Cluster	FIRM and WEEK	FIRM and WEEK	FIRM and WEEK	FIRM and WEEK	FIRM and WEEK	FIRM and WEEK

Table 9: Portfolio Return Alpha and Social Connectedness Beta

This table reports portfolio returns 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:

$$\%PF_{i,j} = \beta_{SC,i} Log\ Social\ Connectedness_{i,j} + \beta_2 SameCounty_{i,j} + \beta_3 SameState_{i,j} + \\ Firm\ FE + Industry\ FE + Distance\ 500 - tile\ FE$$

In Panel A, 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 the institutions' quarter-end holdings. The dollar value-weighted monthly returns, Sharpe ratio, and the standard deviation of the institutional portfolios are computed for each institution every month and then averaged across all institutions (both equally weighted and value-weighted by the AUM) in the same decile. column Excess reports the raw portfolio returns in excess of the three-month treasury bill rate. CAPM, FF-3, and FF-5 report portfolio alphas based on CAPM, Fama and French (1992) three-factor, and Fama and French (2015) five-factor models, respectively. In Panel B, we report these performance metrics for the long-short portfolios constructed within each Bushee institution types. t-statistics are reported in parentheses.

Panel A: Performance Metric for Institutions with Different Propensity to Hold Socially Connected Stocks

Groups:10	Sharpe R.	Std(Excess Ret.)	Equal-Weighted				Value-Weighted			
			Excess	CAPM	FF-3	FF-5	Excess	CAPM	FF-3	FF-5
LOW	0.19	5.65	0.79	0.04	0.03	0.11	0.74	0.02	0.00	0.08
	(28.79)	(56.72)	(1.61)	(0.40)	(0.47)	(1.65)	(1.58)	(0.25)	(0.02)	(1.30)
2	0.2	5.12	0.76	0.07	0.05	0.09	0.72	0.03	-0.00	0.05
	(31.93)	(60.22)	(1.69)	(0.98)	(1.28)	(2.28)	(1.60)	(0.37)	(-0.03)	(1.12)
3	0.19	4.99	0.70	0.01	0.00	0.02	0.66	-0.03	-0.04	-0.02
	(34.05)	(23.52)	(1.58)	(0.28)	(0.15)	(0.66)	(1.49)	(-0.76)	(-1.35)	(-0.84)
4	0.2	4.76	0.72	0.03	0.02	0.03	0.64	-0.05	-0.05	-0.04
	(31.24)	(57.21)	(1.61)	(0.66)	(0.59)	(0.93)	(1.45)	(-1.85)	(-2.07)	(-1.50)
5	0.19	4.49	0.67	0.01	0.01	0.00	0.65	-0.02	-0.03	-0.02
	(31.52)	(65.94)	(1.58)	(0.44)	(0.31)	(0.14)	(1.50)	(-1.11)	(-1.36)	(-1.04)
6	0.19	4.56	0.65	-0.00	-0.00	-0.01	0.67	-0.00	-0.01	0.00
	(32.31)	(59.05)	(1.55)	(-0.11)	(-0.04)	(-0.42)	(1.55)	(-0.16)	(-0.39)	(0.04)
7	0.2	4.4	0.65	-0.02	-0.01	-0.01	0.69	-0.00	-0.01	0.01
	(35.07)	(66.59)	(1.52)	(-0.66)	(-0.51)	(-0.53)	(1.55)	(-0.04)	(-0.25)	(0.31)
8	0.2	4.41	0.68	0.03	0.04	0.02	0.63	-0.03	0.00	-0.01
	(37.58)	(63.08)	(1.62)	(0.95)	(1.46)	(0.62)	(1.48)	(-0.54)	(0.06)	(-0.25)
9	0.21	4.58	0.68	0.03	0.04	0.04	0.66	-0.03	-0.00	0.00
	(35.73)	(41.92)	(1.62)	(0.87)	(1.38)	(1.27)	(1.49)	(-0.55)	(-0.04)	(0.01)
HIGH	0.2	4.95	0.69	0.03	0.05	0.04	0.63	-0.06	-0.04	-0.03
	(28.50)	(39.58)	(1.63)	(0.80)	(1.55)	(1.18)	(1.41)	(-1.10)	(-0.81)	(-0.58)
H-L	0.02	-0.7	-0.09	-0.01	0.02	-0.07	-0.11	-0.08	-0.04	-0.11
	(1.76)	(-4.42)	(-0.85)	(-0.07)	(0.23)	(-0.86)	(-1.24)	(-0.94)	(-0.55)	(-1.57)

Table 9: (Continued)

Panel B: Hoterogeneity of Institution Performance through Bushee Classification

Dep: H-L Bushee Class	Sharpe R.	Std(Excess Ret.)	Equal-Weighted Excess	 CAPM	 FF-3	 FF-5	Value-Weighted Excess	 CAPM	 FF-3	 FF-5
Dedicated	-0.09	1.52	-0.12	-0.12	-0.13	-0.20	-0.41	-0.41	-0.33	-0.59
	(-2.01)	(1.60)	(-0.56)	(-0.55)	(-0.57)	(-0.87)	(-1.20)	(-1.18)	(-1.01)	(-1.78)
Quasi-Index	0.03	-0.86	-0.04	0.05	0.05	0.00	-0.03	0.01	0.01	0.00
	(2.47)	(-4.03)	(-0.40)	(0.51)	(0.67)	(0.01)	(-0.33)	(0.08)	(0.15)	(0.04)
Transient	-0.01	0.22	-0.02	-0.04	-0.00	0.00	-0.01	-0.03	-0.00	0.01
	(-0.65)	(0.80)	(-0.21)	(-0.31)	(-0.02)	(0.02)	(-0.09)	(-0.29)	(-0.02)	(0.12)

Table 10: Portfolio Return Alpha and Social Connectedness Beta

This table reports institutional portfolio returns of socially connected holdings vs. non-connected holdings, using monthly panel data from 2007 to 2016. We sort county pairwise (institional county by firm county) social connectedness into terciles and designate firms located in the top (bottom) tericle as connected (non-connected). We report the portfolio alphas of the highly socally connected, modestly connected, and non-connected stocks. The portfolios are rebalanced at the end of each quarter using institutions' quarter-end holdings. The value-weighted Monthly returns, Sharpe ratio, and the standard deviation of the institutional portfolios are computed for each institution-month and then averaged across all institutions (we report both metrics based on equal-weight and AUM-weight) in the same groups. Excess column reports the raw portfolio returns in excess of the three-month treasury bill rate. CAPM, FF-3, and FF-5 report portfolio alphas based on CAPM, Fama and French (1992) three-factor, and Fama and French (2015) five-factor models. t-statistics are reported in parentheses.

	Equal-Weighted				Value-Weighted			
Social Connectedness	**Excess**	**CAPM**	**FF-3**	**FF-5**	**Excess**	**CAPM**	**FF-3**	**FF-5**
LOW	0.71	0.01	0.02	-0.01	0.66	-0.03	-0.03	-0.06
	(1.55)	(0.13)	(0.31)	(-0.14)	(1.47)	(-0.47)	(-0.48)	(-1.01)
MED	0.72	0.04	0.04	0.00	0.70	0.03	0.01	-0.02
	(1.64)	(0.75)	(0.89)	(0.03)	(1.62)	(0.62)	(0.34)	(-0.47)
HIGH	0.68	-0.00	0.01	0.03	0.65	-0.03	-0.03	-0.01
	(1.55)	(-0.04)	(0.33)	(1.20)	(1.48)	(-1.29)	(-1.37)	(-0.43)
H-L	-0.03	-0.01	-0.01	0.04	-0.01	-0.00	-0.00	0.05
	(-0.42)	(-0.16)	(-0.18)	(0.53)	(-0.14)	(-0.01)	(-0.06)	(0.82)

Appendix A.1 Variable List

Table A.1: Variable Definitions

Variables	Definition
Panel A: Institution-Firm Level Variables	
Social Connectedness Index (SCI)	Number of Facebook friends linked between two counties
Social Connectedness	SCI divided by the product of two county's population. We scale this variable by 10^{12})
Log Social Connectedness	Log(Social Connectedness)
Log Distance	Log(1+Distance)
% PF	(Shares Held$\times$Price/AUM)*100
% IO	(Shares Held$\times$Price/Market Capitalization)*100
Panel B: Firm Level Variables	
Excess Return	the quarterly average of monthly returns (%) over the 3-month T-bill rate.
ILLIQ	Amihud (2002) illiquidity at the quarterly horizon (scaled up by 10^6)
Effective Spread	Dollar-weighted percent effective spread (scaled up by 10^3).
Tobin'Q	the ratio between the market value of the firm over the replacement cost of its assets
M/B	the ratio of market value of equity over the book value of equity
Social Proximity to Capital	County j's social proximity to capital is defined as $\log(\sum_j County\ AUM_i \times Social\ Connectedness_{i,j})$
Physical Proximity to Capital	County j's physical proximity to capital is defined as $\log(\sum_j County\ AUM_i / Distance_{i,j})$
Price	Historical Stock Price at the end of quarter
Size	Firm market capitalization in millions of dollars.
Industry	Fama-French 48 industry indicators.
S&P500	An indicator variable that equals 1 if the firm is included in S&P500 and 0 otherwise
Young	An indicator variable that equals 1 if the firm is listed in the recent 5 years and 0 otherwise
Turnover	Average trading volume scaled by total shares outstanding
B/M	Ratio of book value of equity and market value of equity
Momentum	Cumulative monthly return from t-2 to t-12 constructed following Jegadeesh and Titman (1993)
DY	Dividend yield
R&D	R&D Expenses Scaled by Market Capitalization
Panel C: Fund Level Variables	
AUM	Fund family's asset asset under management in Millions, based on the total market capitalization of their equity holdings.
Return	The value-weighted returns based on fund family's equity holdings reported in the prior quarter.

Social Proximity to Capital: Implications for Investors and Firms

Internet Appendix

Theresa Kuchler, Yan Li, Lin Peng, Johannes Stroebel, and Dexin Zhou[1]

[1]Theresa Kuchler and Johannes Stroebel are with NYU Stern's Finance Department. Yan Li, Lin Peng, and Dexin Zhou are with Baruch College, CUNY.

Table IA.1: Extending Snapshot to Panel Data from 2007 to 2016

This table test the baseline regression with an extended panel data from 2007 to 2016. Column (1) to column (5) replicate the Colum (1) to (5) of Panel A in Table 2 and column (6) and (7) replicate column (3) and column (5) of Panel B in Table 2, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	% PF (Scaled by Fund AUM)					% IO (Scaled by Firm Value)	
Log Social Connectedness	7.362***		7.632***	8.475***	8.210***	2.522***	3.381***
	(11.33)		(7.96)	(8.04)	(7.47)	(11.76)	(8.41)
Log Distance		4.669***	-0.244	-2.482***	-2.698***		-1.219***
		(10.47)	(-0.41)	(-3.13)	(-3.48)		(-4.15)
Firm×QTR FE	YES	YES	YES	YES	YES	YES	YES
Institution×QTR FE	YES	YES	YES	YES	YES	YES	YES
Institution×Industry FE	YES	YES	YES	YES	YES	YES	YES
Same County FE	NO	NO	NO	YES	YES	NO	YES
Same State FE	NO	NO	NO	NO	YES	NO	YES
N	2.88*108̂	2.88*108̂	2.88*108̂	2.88*108̂	2.88*108̂	2.88*108̂	2.88*108̂
R2	0.067	0.066	0.067	0.067	0.067	0.237	0.237

Table IA.2: **Using Winsorized Dependent Variables**

This table replicates the Colum (1) to (5) of Table 2A and Table 2B, respectively in the Colum (1) to (5) and the Colum (6) to (10), with dependent variables winsorized at the upper 99%.

	(1)	(2)	(3)	(4)	(5)
Panel A: Dep=% PF (Scaled by Fund AUM)					
Log Social Connectedness	1.248***		1.485***	1.860***	1.754***
	(9.39)		(6.77)	(9.69)	(8.99)
Log Distance		-0.737***	0.213		
		(-7.31)	(1.30)		
Panel B: Dep=% IO (Scaled by Firm Value)					
Log Social Connectedness	0.346***		0.514***	0.569***	0.564***
	(15.24)		(12.69)	(12.92)	(12.75)
Log Distance		-0.178***	0.151***		
		(-11.62)	(5.68)		
Firm FE	YES	YES	YES	YES	YES
Institution$\times$Industry FE	YES	YES	YES	YES	YES
Distance 500- tile FE	NO	NO	NO	YES	YES
Same State FE	NO	NO	NO	NO	YES
Same County FE	NO	NO	NO	NO	YES
N	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060

Table IA.3: Linear Probability to Holding a Socially Connected Stock

This table replicates the Column (1) to Column (5) of Panel A in Table 2 with dependent variable swithed to a dummy equal to one if the institution helds shares of the firms and zero otherwise.

	(1)	(2)	(3)	(4)	(5)
			$I(\%\ PF > 0)$		
Log Social Connectedness	0.386***		0.518***	0.580***	0.554***
	(12.29)		(9.86)	(11.84)	(11.10)
Log Distance		-0.213***	0.118***		
		(-9.01)	(3.07)		
Firm FE	YES	YES	YES	YES	YES
Institution×Industry FE	YES	YES	YES	YES	YES
Distance 500- tile FE	NO	NO	NO	YES	YES
Same State FE	NO	NO	NO	NO	YES
Same County FE	NO	NO	NO	NO	YES
N	8,694,060	8,694,060	8,694,060	8,694,060	8,694,060
R2	0.402	0.402	0.402	0.403	0.403

Table IA.4: Non-Parametric Estimation of the Effect of Social Connectedness

This table replicates column (5) of Panel A and B in Table IA.2 and column (5) of Table IA.3, respectively in columns (1) to (3), replacing the continous independent variable log SC with its quintile dummies. The quintile dummies are assigned by ranking all the inst-firm pairs as of Jun 2016 into five quintiles, with higher rank to be more socially connected.

	(1)	(2)	(3)
	Winsored at Upper 99%		
Specification	**% PF**	**% IO**	**I(% PF > 0)**
Social Connectedness = 2 (Low)	0.165	0.069*	0.083**
	(1.13)	(1.92)	(2.01)
Social Connectedness = 3	0.671***	0.260***	0.265***
	(3.51)	(5.19)	(4.87)
Social Connectedness = 4	1.125***	0.528***	0.468***
	(3.92)	(7.79)	(5.98)
Social Connectedness = 5 (High)	2.374***	0.878***	0.841***
	(5.82)	(8.76)	(7.60)
Firm FE	YES	YES	YES
Institution×Industry FE	YES	YES	YES
Distance 500- tile FE	YES	YES	YES
Same State FE	YES	YES	YES
Same County FE	YES	YES	YES
N	8,694,060	8,694,060	8,694,060
R2	0.221	0.411	0.403

Table IA.5: Heterogeneity Effect between Local and Distant Firms

This table investigates the heterogeneous effect of social connectedness on institutional investment between local and distant firms. Column (1) and (2) ((3) and (4)) replicates the column (4) and (5) of Panel A (B) in Table 2, respectively, introducing interaction terms with social connectedness. The distance dummies are assigned based on the distance between the counties where the firm and the institution are headquartered, and the dummies indicates the firms headquarterd wihtin the 100-mile radius of the institution headquarter, between 100 and 300 miles, and above 300 miles.

	(1)	(2)	(3)	(4)
	% PF (Scaled by Fund AUM)		% IO (Scaled by Firm Value)	
I(Distance$<$100MI)$\times$Log Social Connectedness	31.162***	32.846***	5.961***	6.378***
	(8.70)	(8.43)	(5.37)	(5.06)
I(100MI$<=$Distance$<=$300MI)$\times$Log Social Connectedness	10.583***	11.399***	3.508***	3.721***
	(8.68)	(8.45)	(7.00)	(6.78)
I(Distance$>$300MI)$\times$Log Social Connectedness	8.495***	8.763***	2.988***	3.058***
	(7.87)	(7.79)	(7.56)	(7.55)
Firm FE	YES	YES	YES	YES
Institution$\times$Industry FE	YES	YES	YES	YES
Distance 500- tile FE	YES	YES	YES	YES
Same State FE	NO	YES	NO	YES
Same County FE	NO	YES	NO	YES
N	8,694,060	8,694,060	8,694,060	8,694,060
R2	0.073	0.073	0.305	0.305

Table IA.6: Two-Dimensional Heterogeneity of the Effect of Social Connectedness

This table studies if the analyst coverage in Panel B of Table 4 and firm size splits in Panel A of Table 4 convoy the similar information. Below replicates the column (5) Panel B of Table 2, introducing interaction terms between social proximity to capital and 3×3 double-sorted group dummies. All institution-firm pairs are assigned into one of the 3×3 independently sorted groups on firm size and analyst coverage at the end of the quarter, with higher rank to be higher level.

Firm Size **No. of Analyst**	**Small**	**2**	**Large**
	4.217***	8.986**	12.850*
Low	(5.54)	(2.38)	(1.68)
	2,364,661	317,549	15,382
	2.590***	3.129***	1.542
2	(2.95)	(3.59)	(0.95)
	1,689,484	1,541,500	144,901
	1.759	2.347***	1.639***
High	(0.60)	(3.15)	(2.90)
	132,569	921,817	1,368,852
F-Pvalue (Null: Low=High)	0.409	0.081	0.150
F-Pvalue (Null: Low=mid=High)	0.291	0.187	0.193
F-Pvalue (Null: Jointly Low=High)		0.111	
T-Pvalue (Null: Low=High)	0.205	0.041	0.075

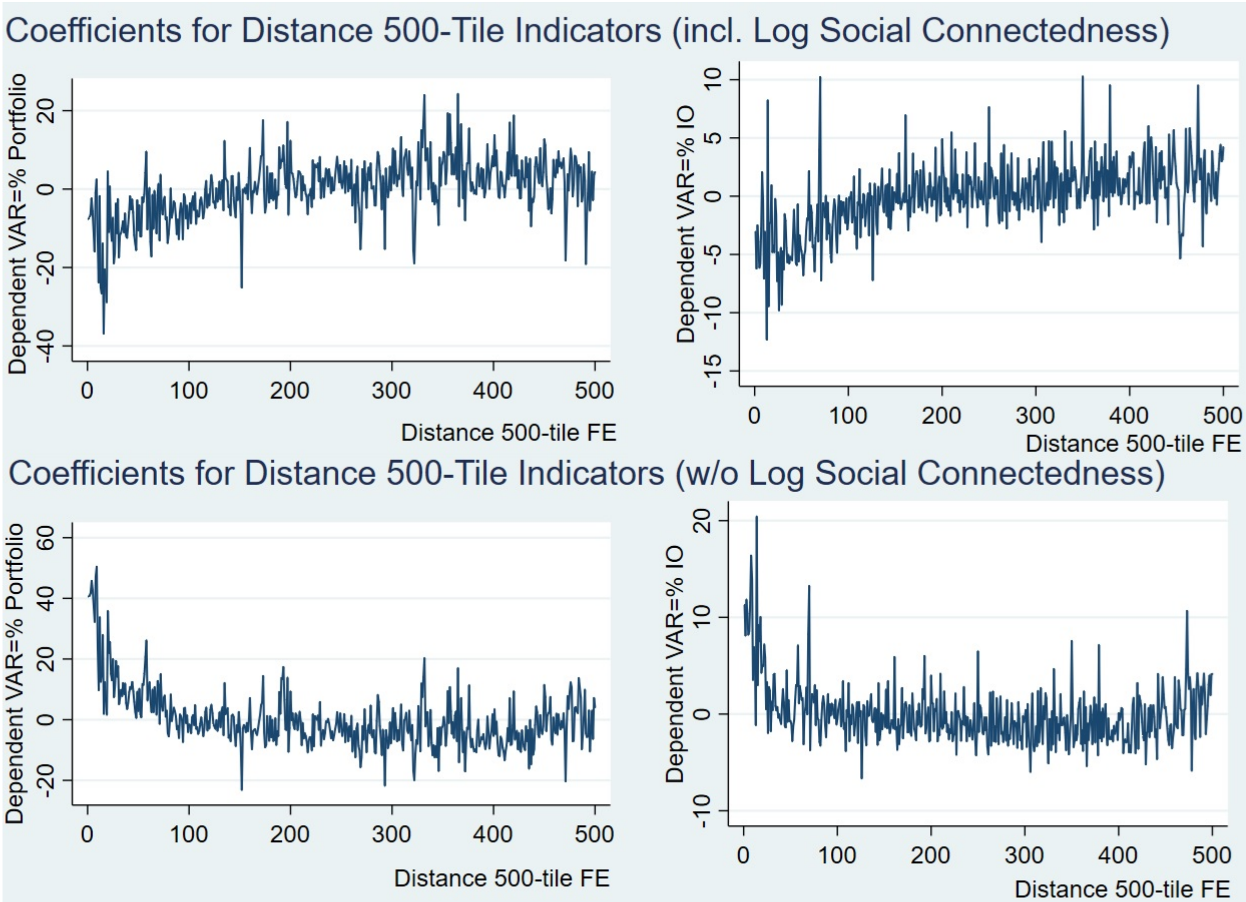


Figure IA.1: 500-Tile Indicators