

# Social Connectedness: Measurement, Determinants, Effects, and Applications

Michael Bailey, Rachel Cao, Theresa Kuchler,  
Johannes Stroebe, and Arlene Wong

Presented by Long Zhen

# Background

- **Social networks** impact many aspects of social and economic activity.
  - Migration
  - Trade
  - Job-seeking
  - Innovation
  - Preference
  - Sentiment
  - Social mobility
  - .....

# Background

- Social networks are also associated with many aspects of activity.
  - Geographic proximity
  - Historical ties
  - Political boundaries
  - .....
- Social network is such an important and influential topic
- But, the unavailability of large-scale and representative data on social networks posed a challenge for related studies.
  - Some existing studies built on microdata from Facebook.

# Motivation

- Facebook: 2.1 bil active users globally and 239 mil in North America.
  - Large enough (Duggan et al(2015): more than 58% US adults)
  - Representative: usage rates are relatively constant across groups
- → This paper introduce **a new measure of social connectedness** at the US county level.
- Contribution
  - The first work to construct social connectedness from massive social media data.
  - Empirically reveal the determinants and effects of social network.
    - Focusing on documenting and describing patterns, instead of causal analysis

# Measurement

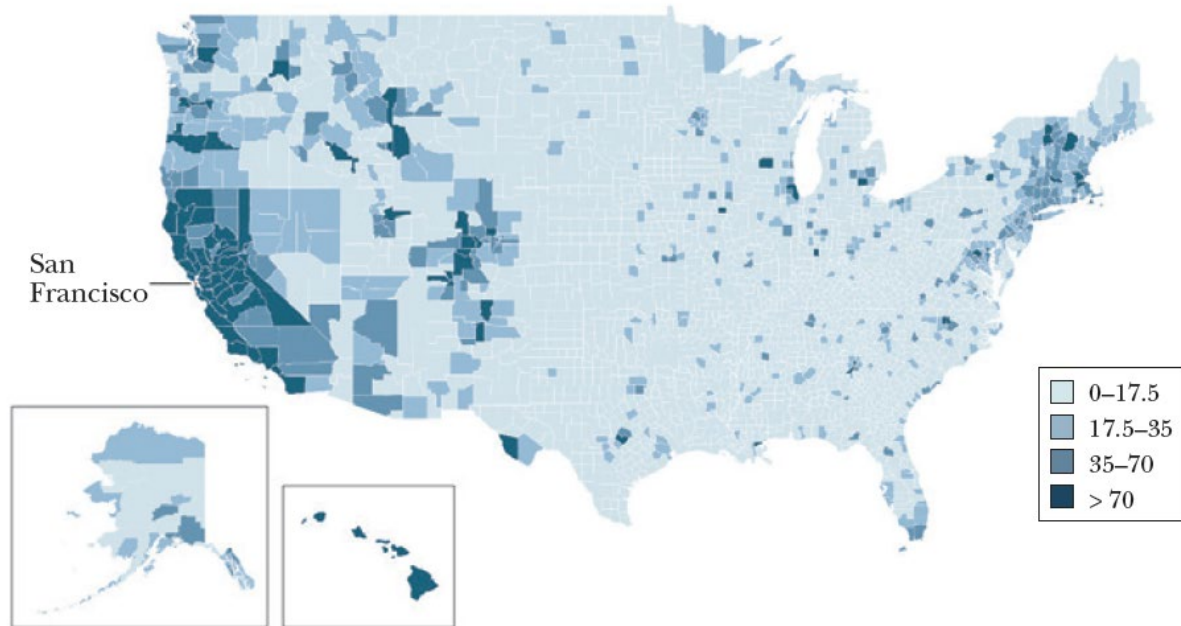
- Using aggregated information of Facebook users of 2016.4
  - (actually updated to 2020.8)
- Map users to their county/country based on information and activity
- Obtain total number of friendships links between geographies
  - Only friendship links among users who have been active over the last 30 days
  - Treat each link identically
- Get Social Connectedness Index between each county/country-foreign country – a relative probability of friendship links

$$Social\ Connectedness_{i,j} = \frac{FB\_Connections_{i,j}}{FB\_Users_i * FB\_Users_j}$$

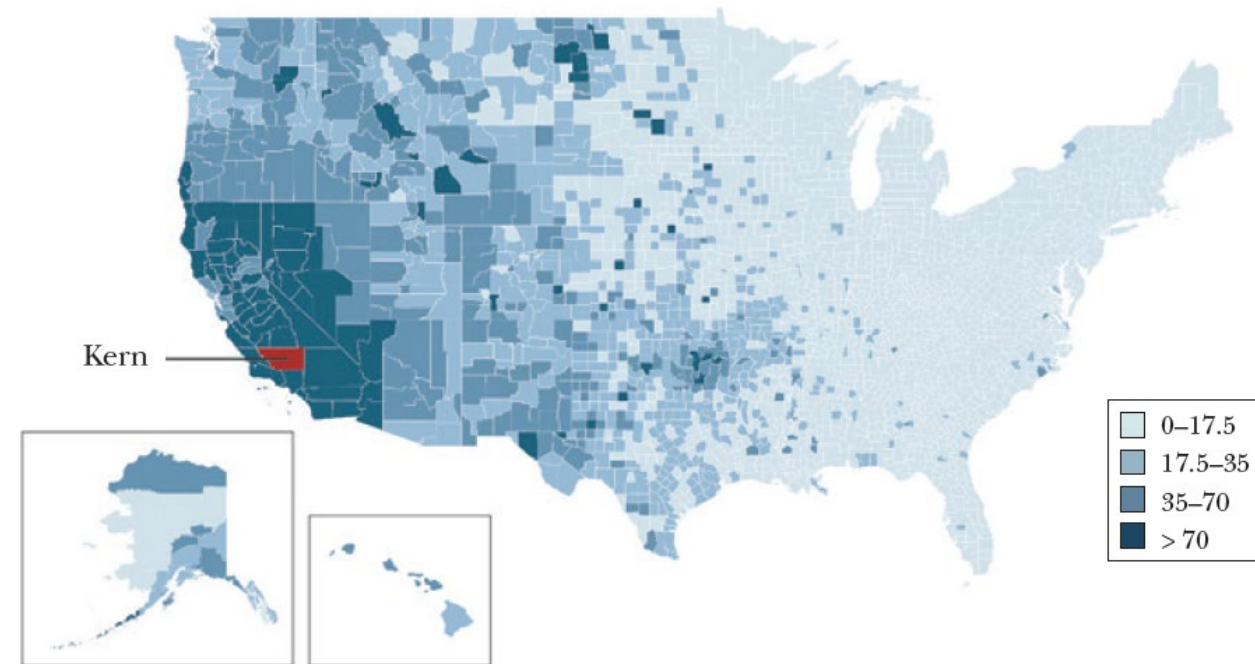
# Determinants

- Geographic distance

A: Relative Probability of Friendship Link to San Francisco County, CA



B: Relative Probability of Friendship Link to Kern County, CA



- Much more geographically concentrated

# Determinants

- Distance is obviously the primary factor

## Distance and Friendship Links: Across-County Summary Statistics for the United States

	<i>Share of friends living within:</i>			<i>Share of US population living within:</i>		
	<i>50 Miles</i>	<i>100 Miles</i>	<i>200 Miles</i>	<i>50 Miles</i>	<i>100 Miles</i>	<i>200 Miles</i>
Mean	55.4%	62.8%	70.3%	1.3%	2.8%	6.6%
P5	38.1%	46.0%	54.2%	0.1%	0.3%	1.0%
P10	42.5%	49.6%	57.1%	0.1%	0.6%	2.1%
Median	55.4%	63.9%	71.6%	0.7%	2.1%	5.8%
P90	67.4%	74.8%	81.2%	3.2%	6.2%	15.0%
P95	70.3%	76.9%	83.2%	5.4%	9.2%	15.6%

# Determinants

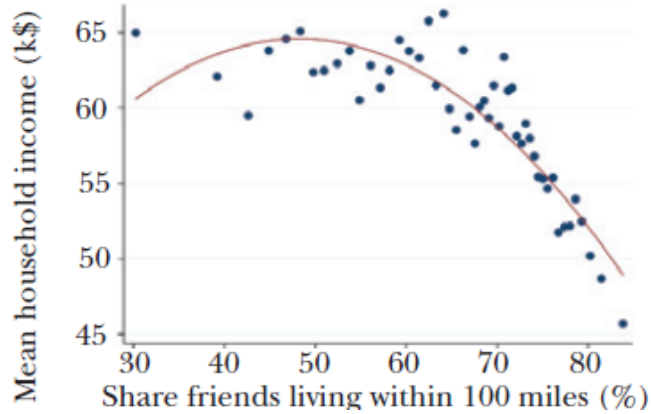
- Other factors

	Dependent Variable: Log(SCI)				
	(1)	(2)	(3)	(4)	(5)
log(Distance in Miles)	-1.483*** (0.065)	-1.287*** (0.061)	-1.160*** (0.059)	-1.988*** (0.043)	-1.214*** (0.055)
Same State		1.496*** (0.087)	1.271*** (0.083)	1.216*** (0.044)	1.496*** (0.085)
Δ Income (\$1,000)					-0.006*** (0.001)
Δ Share Population White (%)					-0.012*** (0.001)
Δ Share Population No High School (%)					-0.012*** (0.002)
Δ 2008 Obama Vote Share (%)					-0.006*** (0.001)
Δ Share Population Religious (%)					-0.002*** (0.001)

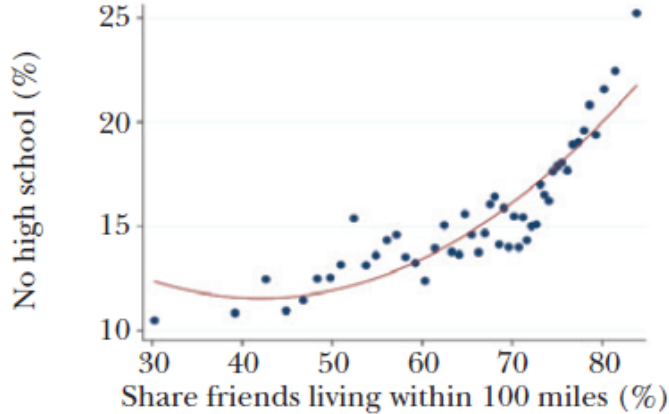


# Concentration Heterogeneity

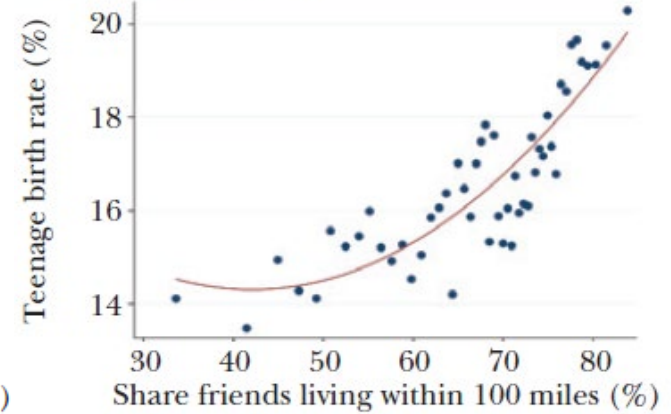
A: Average Income



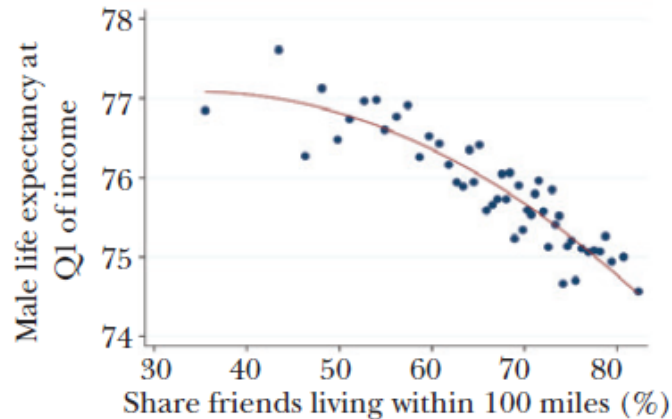
B: Percent No High School



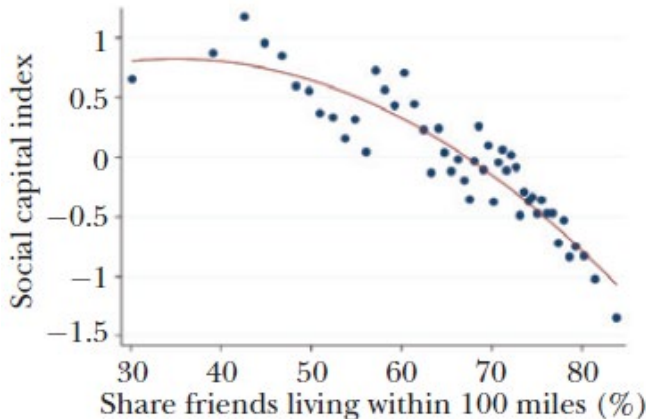
C: Teenage Birth Rate



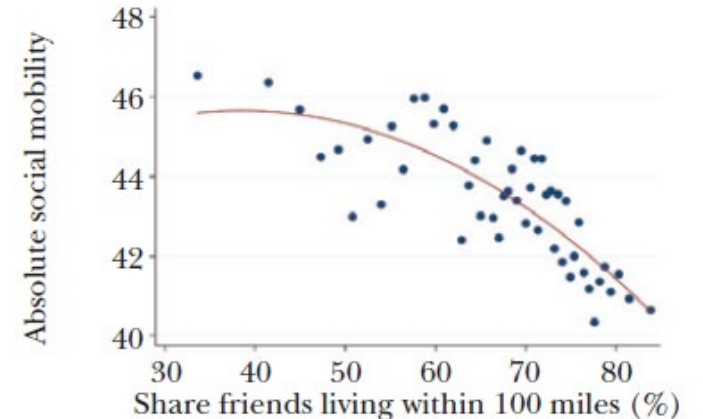
D: Life Expectancy



E: Social Capital



F: Social Mobility



# Effect

- Trade flow

	(1)	(2)	(3)	(4)
<i>Panel A: Dependent Variable: log(State-Level Trade Flows)</i>		data from the Commodity Flow Survey		
log(Distance)	-1.057*** (0.071)		-0.531*** (0.084)	-0.533*** (0.085)
log(SCI)		0.999*** (0.051)	0.643*** (0.071)	0.637*** (0.060)
State Fixed Effects	Y	Y	Y	Y
Other State Differences	N	N	N	Y
Observations	2,219	2,220	2,219	2,219
$R^2$	0.912	0.918	0.926	0.930

- SCI is strongly correlated with state–state trade flows
- controlling for SCI significantly reduces the estimated distance elasticities

# Effect

- Knowledge spillover

US Patent and Trademark Office in 2014

*Panel B: Dependent Variable: Indicator for Patent Citation*

log(Distance)	-0.048*** (0.002)		-0.011** (0.005)	-0.021** (0.009)
log(SCI)		0.063*** (0.003)	0.049*** (0.006)	0.066*** (0.012)
Technological Category + County Fixed Effects	Y	Y	Y	Y
Cited + Issued Patent Fixed Effects, Other County Differences	N	N	N	Y
Observations	2,171,754	2,171,754	2,171,754	2,168,285
$R^2$	0.056	0.059	0.059	0.101

- significant correlation between SCI and knowledge spillovers, innovation, and, ultimately, economic growth.

# Effect

- Migration

the Statistics of Income (SOI) Tax Stats Migration Data provided by the IRS

*Panel C: Dependent Variable: log(County-Level Migration)*

log(Distance)	−0.973*** (0.048)		0.023 (0.021)	0.031 (0.021)
log(SCI)		1.134*** (0.019)	1.148*** (0.024)	1.159*** (0.024)
County Fixed Effects	Y	Y	Y	Y
Other County Differences	N	N	N	Y
Observations	25,305	25,305	25,305	25,287
$R^2$	0.610	0.893	0.893	0.893

- Individuals are more likely to move to counties where they already have friends.

# Effect

## Social Connectedness, Ancestry, and International Trade

	log(SCI)			log (Exports + 1)	log (Imports + 1)
	(1)	(2)	(3)	(4)	(5)
log(Distance)	-1.159*** (0.258)	-0.690*** (0.162)	-0.493*** (0.174)	-2.092*** (0.391)	-1.627*** (0.378)
log(Ancestry in Foreign Country)		0.341*** (0.022)			
log(Born in Foreign Country)			0.367*** (0.033)		
log(SCI)				0.597*** (0.139)	0.470*** (0.103)
Fixed Effects	Y	Y	Y	Y	Y
Observations	33,146	33,146	16,527	11,015	11,014
$R^2$	0.908	0.936	0.943	0.770	0.770
Number of Countries	105	105	52	216	216

# Conclusion

- The SCI should allow researchers to overcome some of the measurement challenges that have held back empirical research on the role of social interactions in finance, economics, and the broader social sciences
- → Is there any implications for investors?

# Social Proximity to Capital: Implications for Investors and Firms

Theresa Kuchler et al. 2022 Review of Financial Studies

# Motivation

- Large regional differences in economic outcomes across US
  - Dougal et al.(2018): coastal cities have better amenities
- The role of public equity investments by institutional investors
  - Providing capital, liquidity, and corporate governance
  - Available data
  - Substantial geographic variation in the location of institutional investors
- Contribution
  - Examine how social connectedness affect economic decisions



# Research Questions

- Whether institutional investors tend to invest more in firms located in regions that the investors are socially connected to?
  - Yes
- Do firms located in regions with higher social proximity to institutional capital attract more overall institutional investment?
  - Yes
- Whether social proximity to capital affects firms' valuation?
  - Yes

# Data

- Social connectedness index

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

- Institutional holding data:
  - Thomson Reuters Institutional Holding/CRSP/SEC filing
  - Institution-firm pair holding:

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

- → 3,083 firms and 2,820 fund families

# Empirical Results

- Baseline specification

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

Social connectedness and institutional investment

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

- Log SCI ~ institutional investment
- Stronger when they are near, but even they are over 100 miles apart, the coefficient is still significantly positive.

- Heterogeneity analysis
  - Assumption: driven by an increase in investors' awareness of firms located in socially connected places
- By investor characteristics

Social connectedness and investment: by institution characteristics

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

- By firm characteristics

### Social connectedness and investment: by firm characteristics

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

- By fund manager characteristics

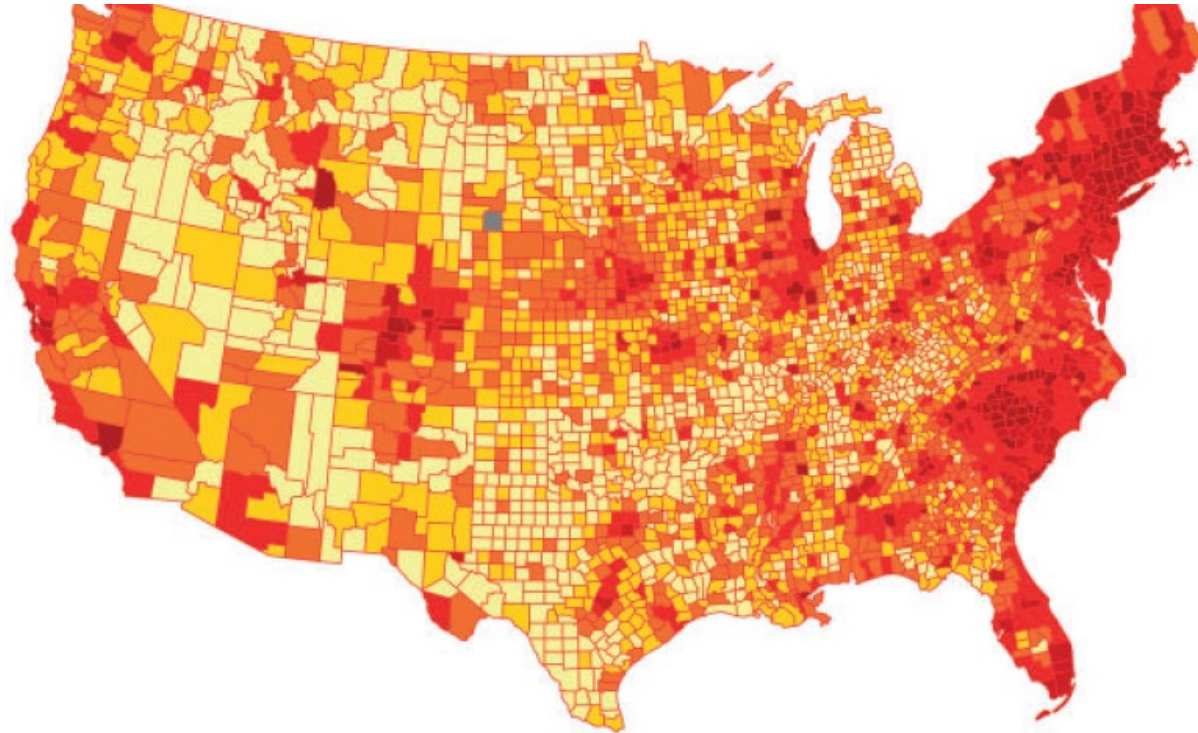
Social connectedness and mutual fund investment, with fund manager characteristics

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

# Capital market implication for firms

$$\text{Social Proximity to Capital}_{i,t} = \sum_j AUM_{j,t} \times \text{Social Connectedness}_{i,j},$$

$$\text{Physical Proximity to Capital}_{i,t} = \sum_j AUM_{j,t} / (1 + \text{Distance}_{i,j}),$$



total institutional ownership

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

Firms' social proximity to capital and institutional ownership

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



$$\begin{aligned} \text{Log Valuation}_{i,t} = & \beta \text{Log Social Proximity to Capital}_{i,t-1} + \gamma X_{i,t-1} \\ & + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \end{aligned}$$

Social proximity to capital and firm value

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

$$\begin{aligned} \text{Log Liquidity}_{i,t} = & \beta \text{Log Social Proximity to Capital}_{i,t-1} + \gamma X_{i,t-1} \\ & + \psi_t + \xi_{ind(i)} + \eta_{state(i)} + \epsilon_{i,t}, \end{aligned}$$

**Social proximity to capital and stock liquidity**

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

# Implications for institutional investors

- Investors that hold more socially connected stocks should outperform other investors that hold fewer such stocks.

Portfolio social connectedness and performance

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

**Performance of socially connected holdings**

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

- Institutional investors' investments in socially connected firms are primarily driven by awareness of these firms rather than by superior information

# Conclusion

- Institutional investors invest more in firms located in regions to which they have stronger social ties.
- Firms in regions that are socially proximate to institutional capital have higher liquidity and higher valuations.