Does Social Connectedness Affect Stock Market Participation?

Chih-Ching Hung

Baruch College, City University of New York

January 3^{rd}-5^{th}, 2021

2021 AFA PhD Poster Session

Non-Participation Puzzle

- The average equity premium each year between 1928 and 2019 is 7.8%.
- Yet, a substantial share of households of all income groups do not participate in stock market!
  Haliassos and Bertaut (1995), Campbell (2006), Guiso and Sodini (2013)
- Relevance of NPP:
  (1) Challenge traditional asset pricing model like Merton (1969).
  (2) Deteriorate income equality.
  (3) May lead to lower aggregate output.
  Bhamra and Uppal (2019)

Traditional explanations to NPP

- Preference-based:

  Barberis, Huang, and Thaler (2006): loss aversion, narrow framing.

  Garlappi, Uppal, and Wang (2007): ambiguity aversion.

- Cost-based:

  Fiscal participation costs and/or information costs.
  $\rightarrow$ cross-sectional differences.
  Vissing-Jörgensen (2003), Cole, Paulson, and Shastry (2014)

- Belief and Trust

  Guiso, Sapienza, and Zingales (2008), Gurun, Stoffman, and Yonker (2018)

IRS Statistics of Income

- Estimation of **Stock Market Participation**:

$$\frac{\#tax\ filers\ with\ dividend\ tax}{\#tax\ filers}$$

- County-level data from 2010 to 2018.

- What dividend tax captures:

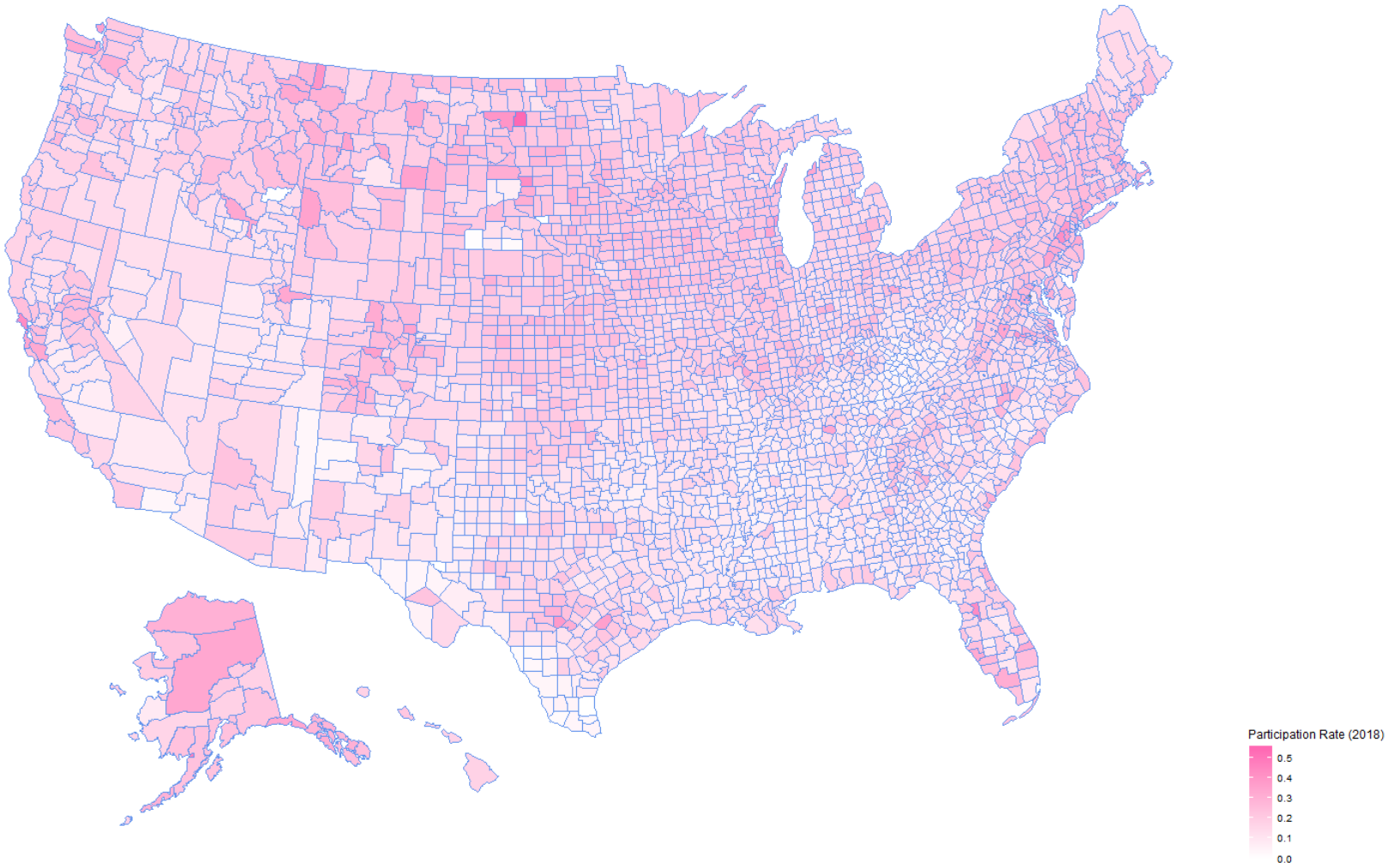
  (1) Most of the direct stock holdings.

  (2) 20% of mutual fund holdings that focus on high-payout stocks.

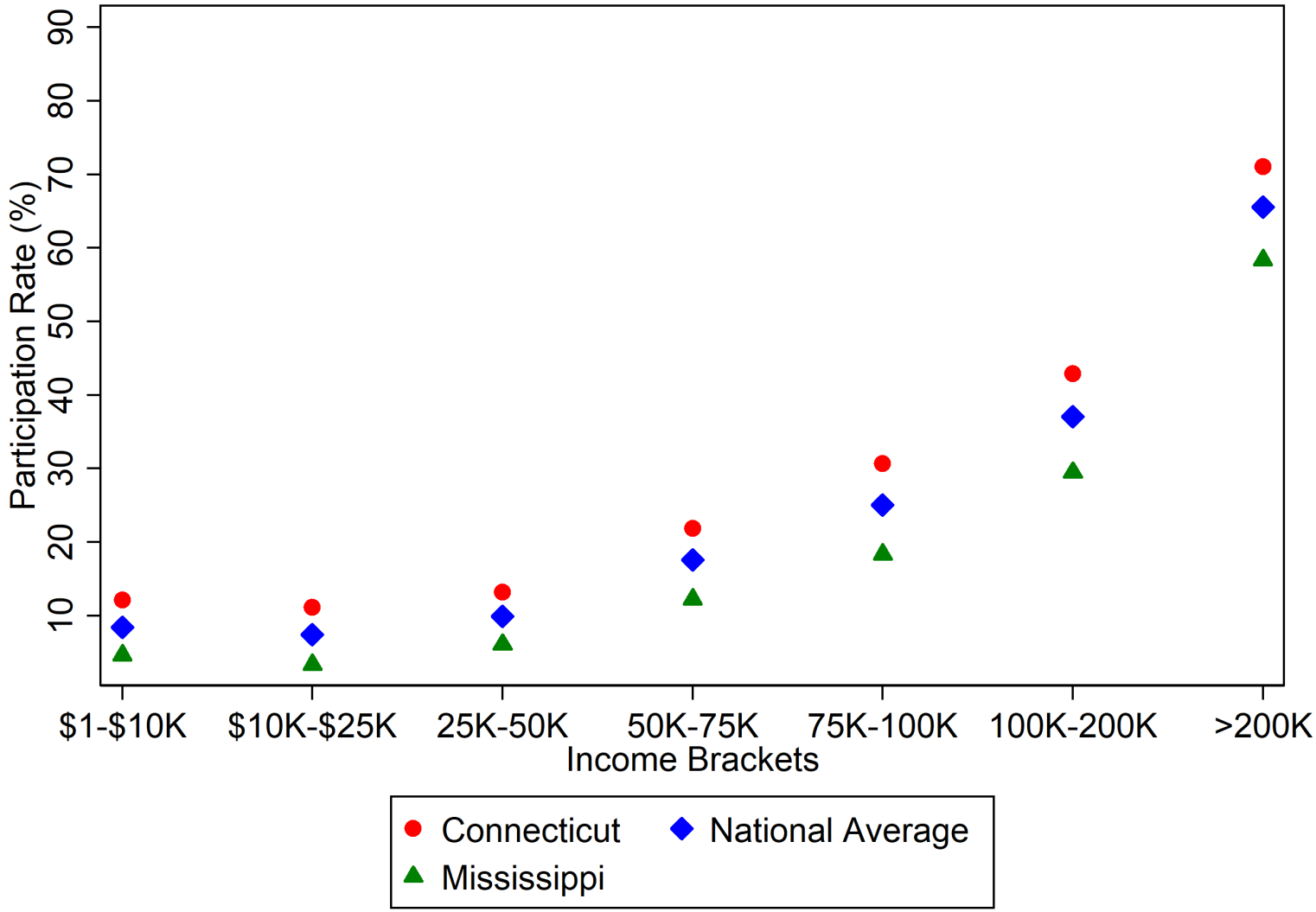
- Correlation with SCF participation rate: 96.5%

Time-Series Comparison | Income-Cross-Section Comparison

Stock market participation from IRS in 2018



Stock market participation across income in 2018



Social connectedness

- A different channel: Social Connectedness
  - We rely on our friends, parents, and opinion leaders to acquire information, form belief, and make all kinds of economic decision.
  - Participation: Hong, Kubik, and Stein (2004), Kaustia and Knüpfer (2012)
    Retirement plan: Duflo and Saez (2002), Ouimet and Tate (2020)
  - However, Angrist (2014): The perils of peer effects.
    Endogeneity problem arises when implying causality from correlation.
- Main contribution of this paper:

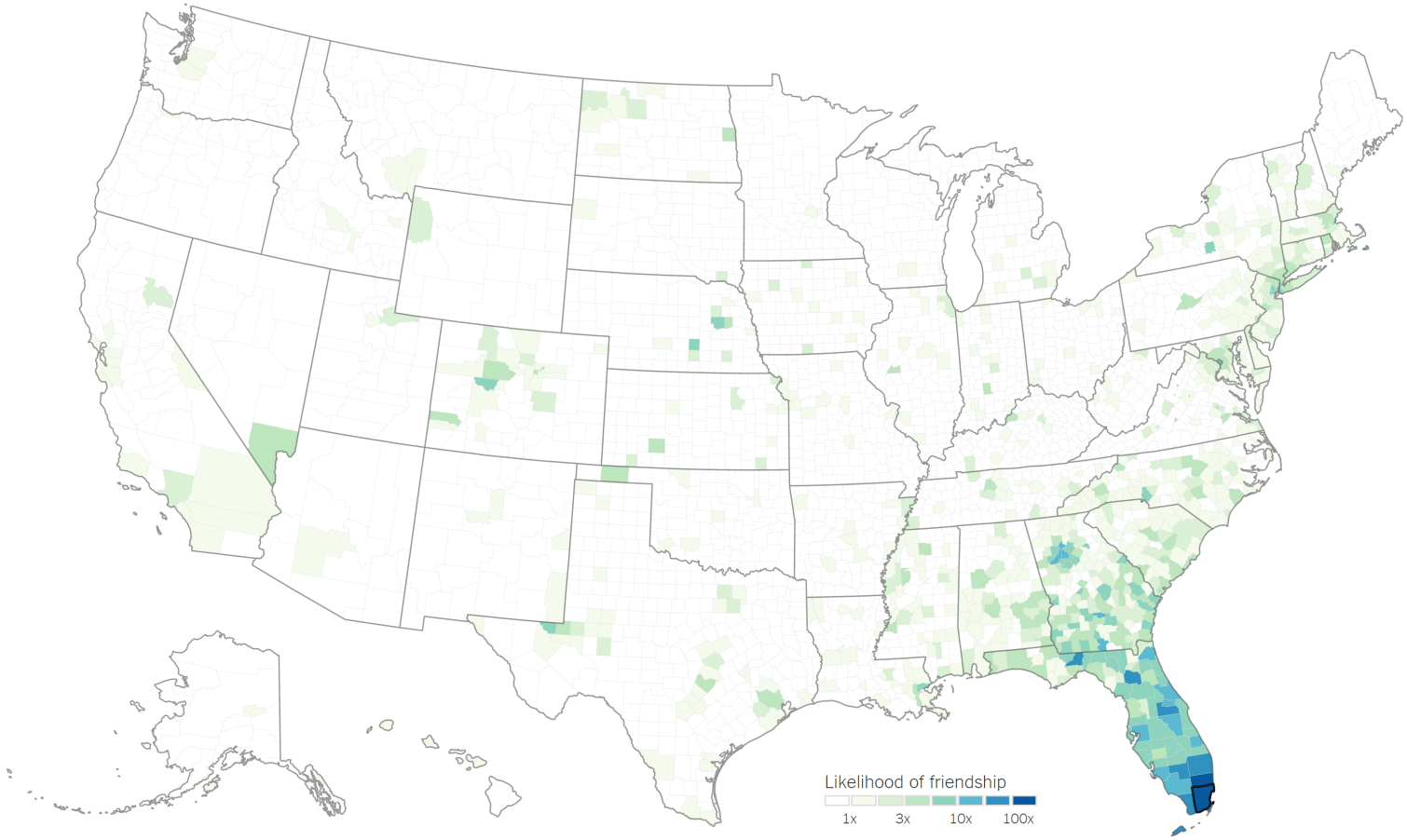
  1) Much more representative sample
  2) Identification strategy for causal inference
  3) Welfare implication

Facebook SCI

- 2016 Snapshot of Facebook user logins, large coverage of US population.
- Properties: (see Bailey, Cao, Kuchler, Stroebel, and Wong (2018))
  (1) Reflect persistent social connectedness beyond FB usages.
  (2) Spatial connectedness - cultural, industrial, immigration connections.
- Applications in Finance and other areas:
  - Proximity to capital, P2P lending expansions
  - Spread of COVID, commuting flows in urban areas, beliefs about local mortgage price

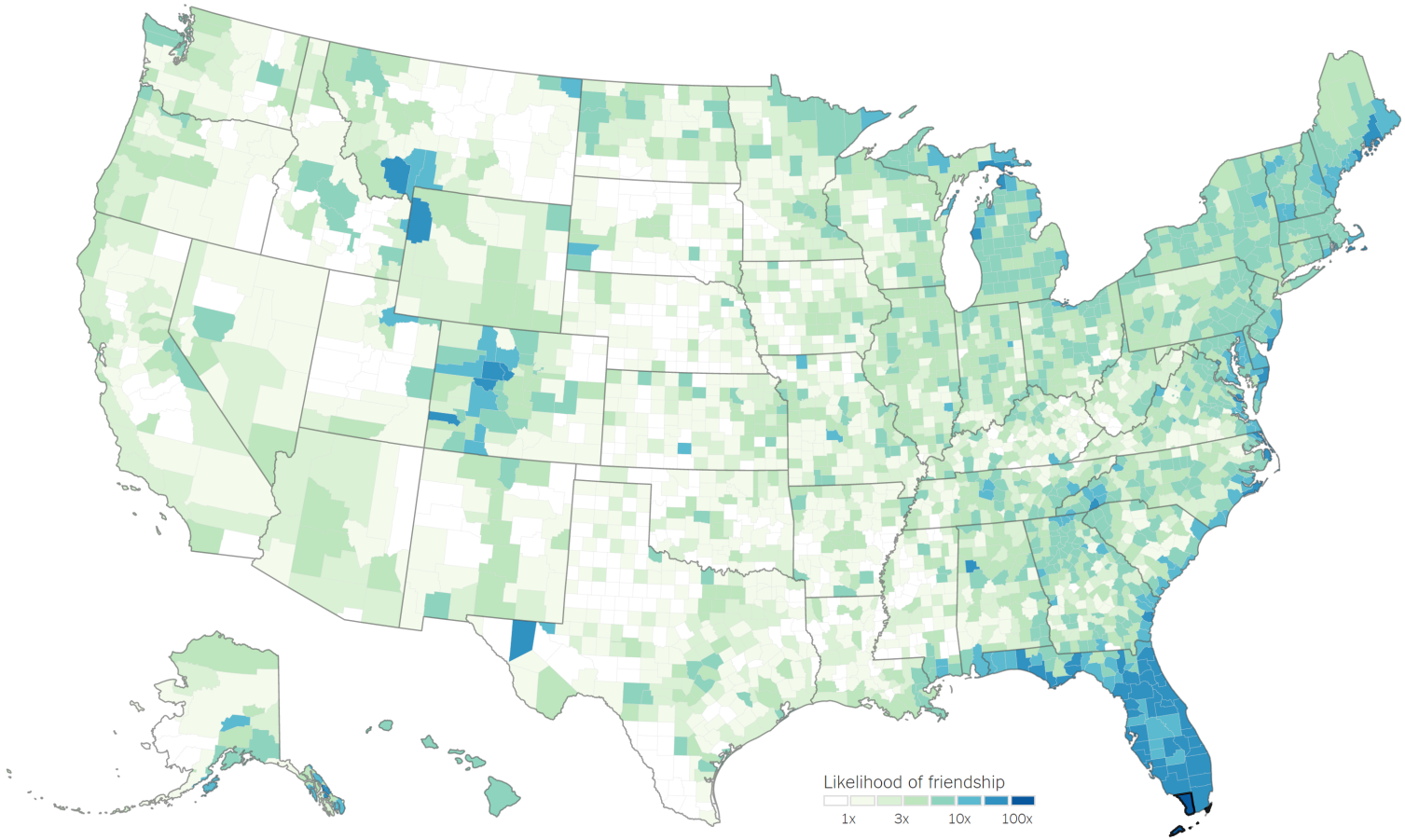
Facebook SCI: example

The relative probability that someone in any U.S. county has a Facebook friendship link to Miami-Dade County, Fla.



Facebook SCI: example

The relative probability that someone in any U.S. county has a Facebook friendship link to Monroe County, Fla.



Empirical method

- Mian and Sufi's spirit of research design.

- explanatory variable: *Friends' Participation*

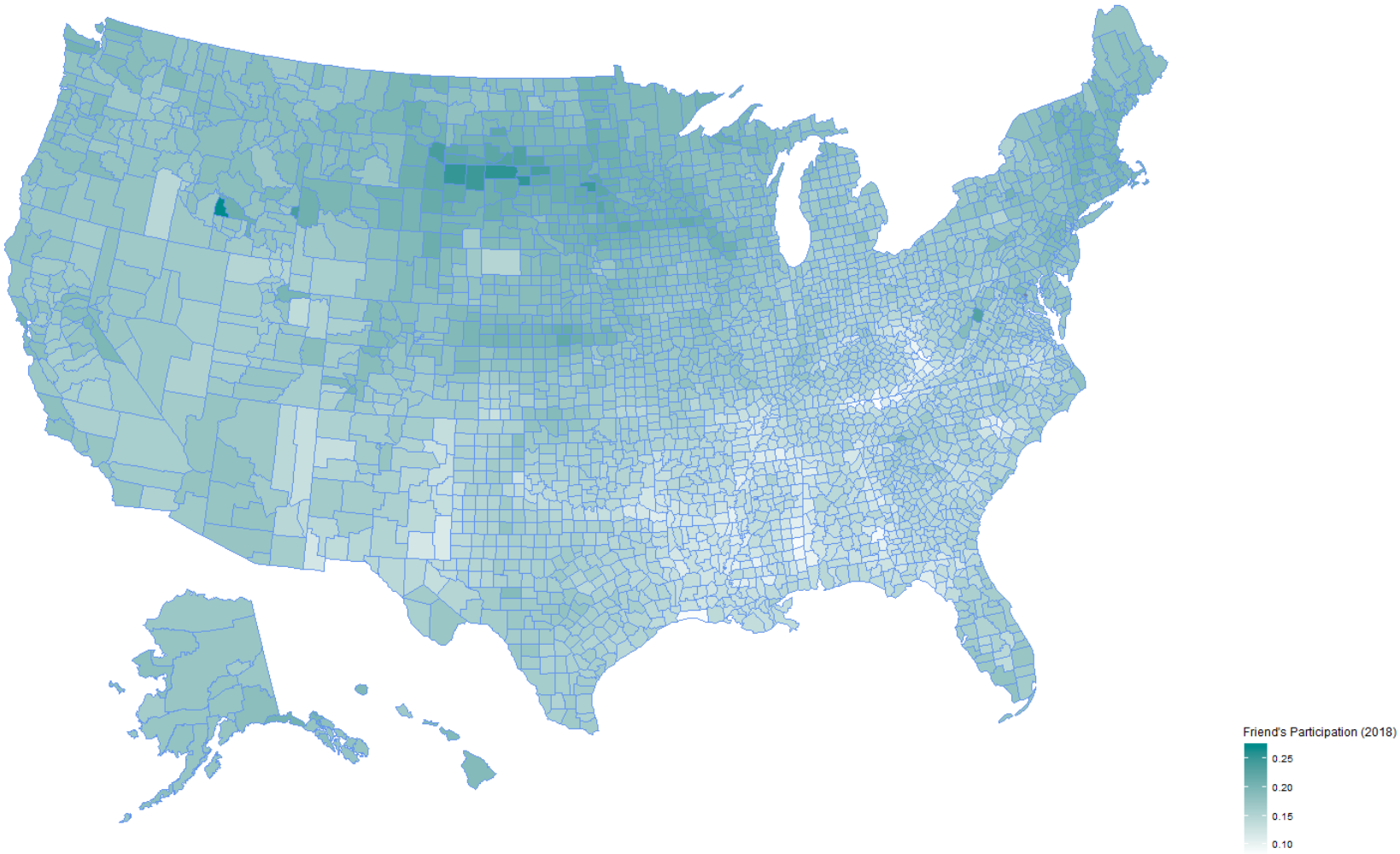
$$SCI_{i,j} = \frac{Friendships_{i,j}}{FB\ User_i\ *\ FB\ User_j}$$

$$FP_{i,t} = \frac{\sum SCI_{j \in \Omega \setminus I} * Participation_{j,t}}{\sum SCI_{j \in \Omega \setminus I}}$$

  where I is the state where county i resides.

- **Identification Strategy**:
  "**Revelation**" of financial advisory misconduct as a negative shock.

Friends' Participation in 2018



Benchmark model (1): level of participation

	(1)	(2)	(3)	(4)	(5)	(6)
			Dependent: Participation$_{i,t}$ (%)			
FP$_{i,t-1}$ (%)		1.1848***		1.0324***		0.5484***
		[30.54]		[5.81]		[4.76]
Participation$_{i,t-1}$ (%)					0.4600***	0.4297***
					[5.67]	[5.29]
Population/1,000$_{i,t}$	-0.0052***	-0.0041***	0.0083**	0.0097**	0.0057**	0.0066**
	[-4.17]	[-4.63]	[2.00]	[2.18]	[2.02]	[2.13]
Education$_{i,t}$ (%)	0.3083***	0.2181***	-0.0216***	-0.0229***	-0.0189***	-0.0198***
	[25.35]	[18.95]	[-2.87]	[-3.15]	[-3.24]	[-3.44]
Unemploy$_{i,t}$ (%)	-0.3927***	-0.1424***	0.0106	0.0071	0.0033	0.0019
	[-10.91]	[-4.00]	[1.03]	[0.75]	[0.35]	[0.21]
Income/1,000$_{i,t}$	0.1415***	0.1115***	0.0085	0.0079	0.0089*	0.0085
	[11.77]	[10.52]	[1.19]	[1.14]	[1.70]	[1.63]
Return$_{i,t-1}$ (%)	0.0001	0.0010	-0.0001	-0.0001	0.0001	0.0001
	[0.07]	[1.10]	[-1.09]	[-1.05]	[0.66]	[0.57]
Intercept	6.9311	-11.4557	16.0307	-1.0187	8.5142	-0.0463
	[10.77]	[-12.72]	[37.76]	[-0.35]	[6.07]	[-0.03]
County FE			Y	Y	Y	Y
State*Year FE			Y	Y	Y	Y
Obs	25060	25059	25060	25059	25059	25059
adj. R^2	0.479	0.612	0.971	0.972	0.977	0.977

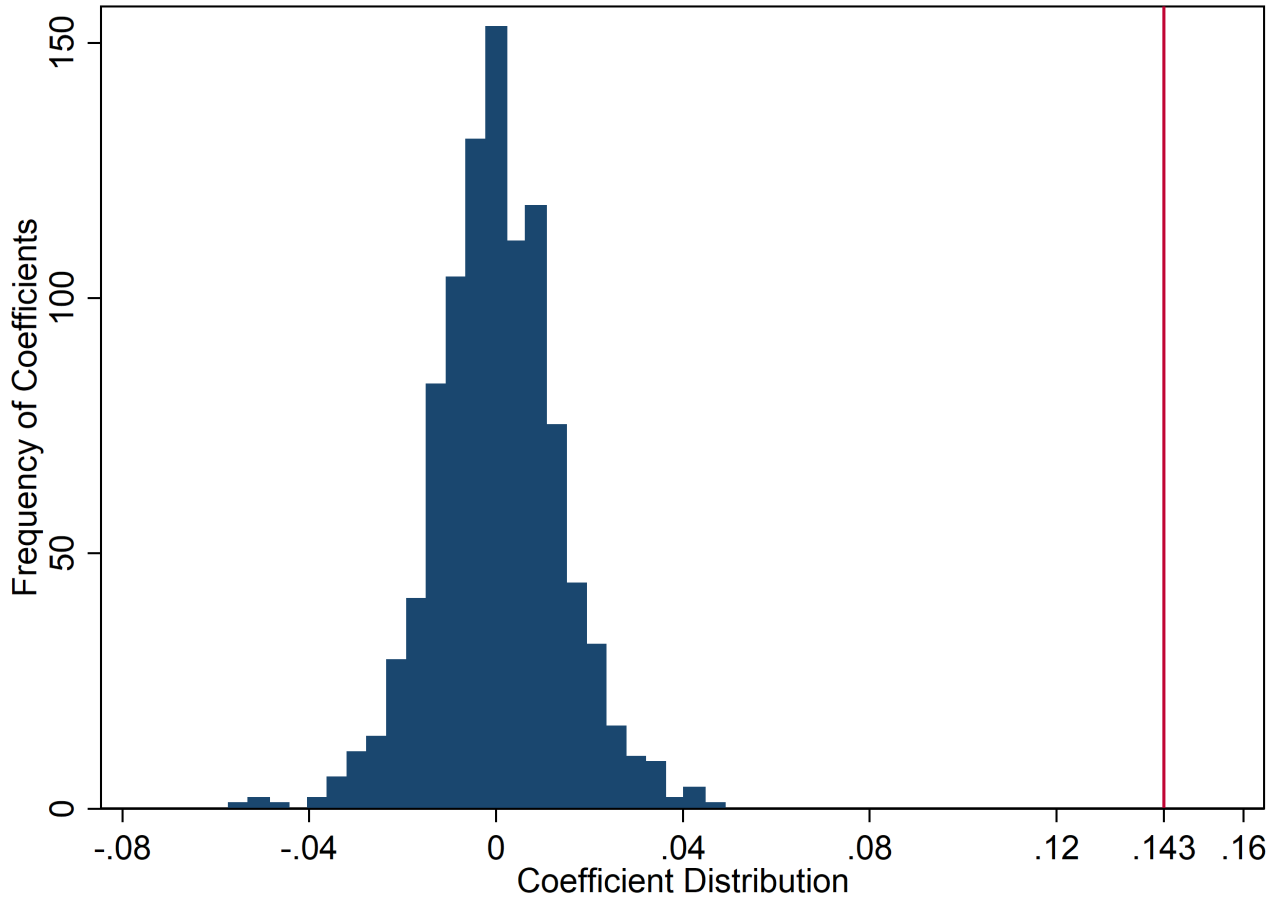
Benchmark model (2): change in participation

	(1)	(2)	(3)	(4)
	Dependent: ΔParticipation$_{i,t}$ (%)			
ΔFP$_{i,t-1}$ (%)		0.2472***		0.1417*
		[3.59]		[1.86]
ΔParticipation$_{i,t-1}$ (%)	-0.1356***	-0.1441***	-0.1942***	-0.1988***
	[-3.08]	[-3.11]	[-3.17]	[-3.11]
ΔPopulation/1,000$_{i,t}$	0.0117**	0.0121**	0.0077	0.0080
	[2.47]	[2.50]	[1.02]	[1.07]
ΔEducation$_{i,t}$ (%)	-0.0193***	-0.0190***	-0.0175***	-0.0172***
	[-3.31]	[-3.25]	[-2.79]	[-2.75]
ΔUnemploy$_{i,t}$ (%)	-0.0045	-0.0047	-0.0077	-0.0078
	[-0.47]	[-0.51]	[-0.74]	[-0.76]
ΔIncome/1,000$_{i,t}$	0.0032	0.0039	0.0015	0.0020
	[0.49]	[0.59]	[0.19]	[0.25]
Return$_{i,t-1}$ (%)	0.0002**	0.0002**	0.0002**	0.0002**
	[2.08]	[2.09]	[2.10]	[2.10]
Intercept	-0.1050	-0.0760	-0.1103	-0.0937
	[-8.78]	[-6.35]	[-8.58]	[-9.27]
County FE			Y	Y
State*Year FE	Y	Y	Y	Y
Obs	21924	21924	21924	21924
R^2	0.095	0.097	0.186	0.187
adj. R^2	0.080	0.082	0.032	0.033

- Focus on the change specification afterwards.

Placebo test

- To take care of potential global shock affecting the result, I randomly simulate 1,000 sets of artificial SCI.



Geographic distance and social distance

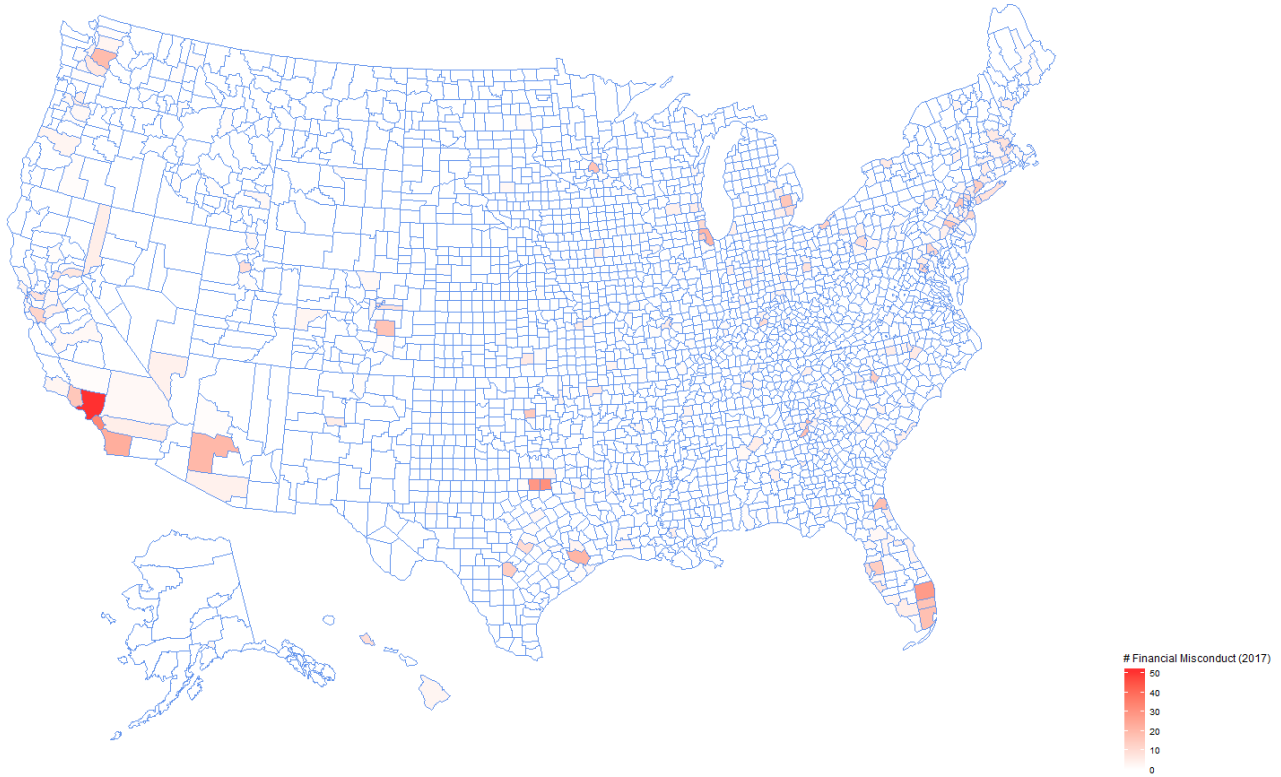
- $FP_{i,t-1}^{SameState} = \frac{\sum SCI_{j \in I} * Participation_{j,t-1}}{\sum SCI_{j \in I}}$

	(1)	(2)	(3)
	Dependent:	ΔParticipation$_{i,t}$ (%)	
ΔFP$_{i,t-1}$ (%)			0.1400*
			[1.93]
ΔFP$_{i,t-1}^{SameState}$ (%)	0.2293***	0.3262*	0.3256*
	[2.71]	[1.75]	[1.73]
Controls	Y	Y	Y
County FE		Y	Y
State*Year FE		Y	Y
Obs	21924	21924	21924
adj. R^2	0.024	0.037	0.037

Financial misconduct revelation in 2017

- Egan, Matvos, and Seru (2019): Financial Misconduct data from FINRA.

  Include: Employment Separation After Allegations, Customer Dispute (Settled), Customer Dispute (Award/Judgment), Regulatory (Final), Civil (Final), Criminal (Final Disposition)
- Use the **"revelation"** rather than **happening** of misconducts as a shock!



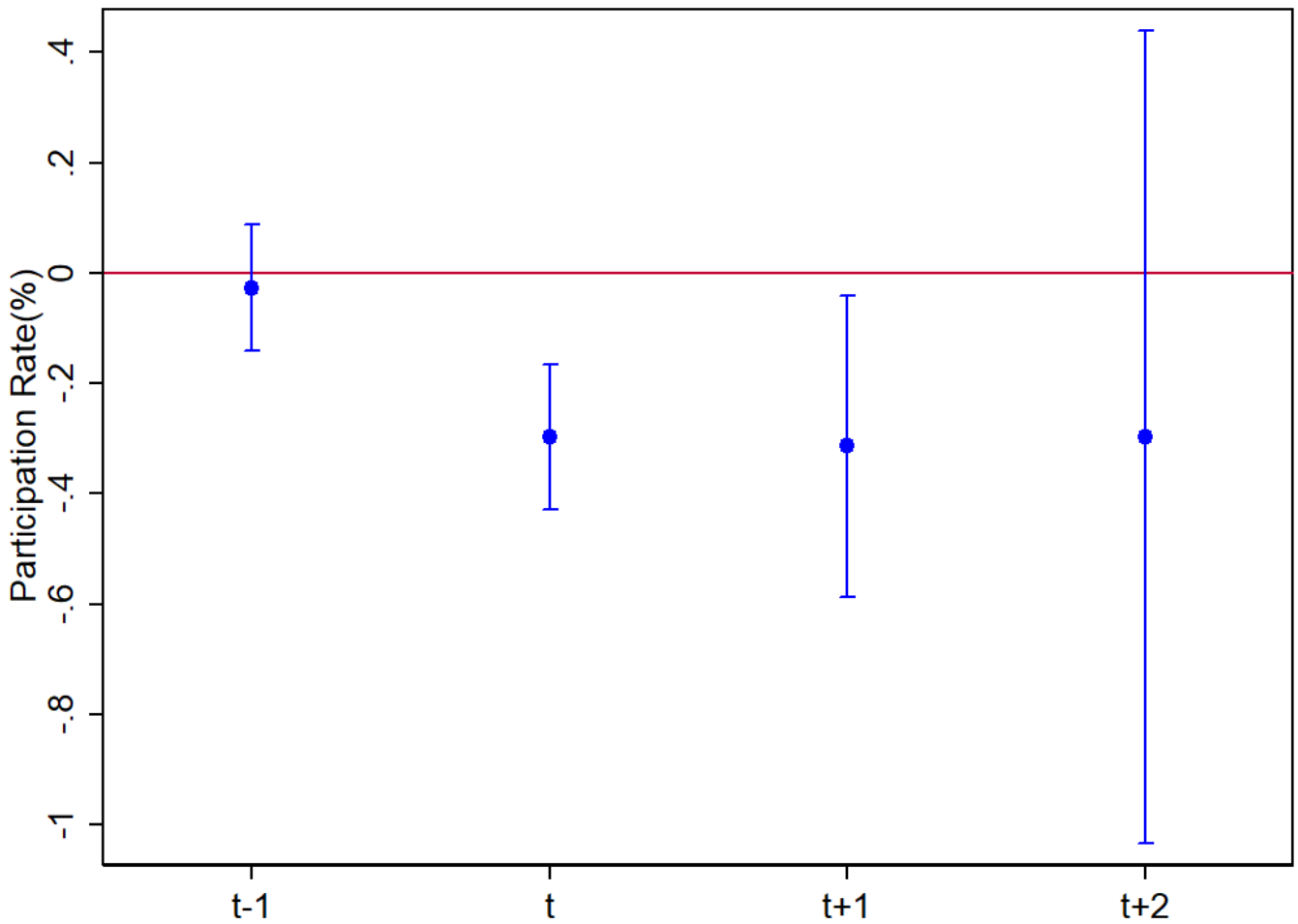
Identification: First stage

- *Prob. of Misconduct* is the number of misconduct revelation scaled by thousand population.

	(1)	(2)	(3)	(4)
		Dependent: ΔParticipation$_{i,t}$ (%)		
Prob. of Misconduct$_{i,t+1}$			-0.0094	-0.0242
			[-0.17]	[-0.37]
Prob. of Misconduct$_{i,t}$		-0.2589**	-0.2743***	-0.2948***
		[-2.54]	[-3.07]	[-3.63]
Prob. of Misconduct$_{i,t-1}$			-0.2498*	-0.3108*
			[-1.88]	[-1.91]
Prob. of Misconduct$_{i,t-2}$				-0.3003
				[-0.67]
Controls	Y	Y	Y	Y
County FE	Y	Y	Y	Y
State*Year FE	Y	Y	Y	Y
Obs	21924	18793	15657	15657
adj. R^2	0.032	0.060	0.059	0.059

- F-stats of *Prob. of Misconduct* in column (2) is 22.16.

Identification: First-stage shock



Identification: instrument variable

- $\Delta FP_{i,t}^{IV} = \frac{\sum SCI_{j \in \Omega \setminus I} * \Delta \widehat{Participation}_{j,t}}{\sum SCI_{j \in \Omega \setminus I}}$

	(1) ΔParticipation$_{i,t}$ (%)	(2) ΔParticipation$_{i,t}$ (%)
$\Delta FP_{i,t-1}$ (%)	0.1417*	
	[1.86]	
$\Delta FP_{i,t-1}^{IV}$ (%)		0.4553**
		[2.17]
Controls	Y	Y
County FE	Y	Y
State*Year FE	Y	Y
Obs	21924	18790
adj. R^2	0.033	0.042

- People react asymmetrically stronger toward negative events.

Income Cross-section of FP effect

- Estimate participation rates for low-, mid-, and high-income groups. Thresholds: 50K & 100K.

	(1)	(2)	(3)	(4)
	ΔParticipation$^{Income\ Group}_{i,t}$ (%)			
	All	Low	Mid	High
$\Delta FP_{i,t-1}$ (%)	0.1417*	0.1892**	0.1514	0.7534**
	[1.86]	[2.17]	[0.65]	[2.62]
$\Delta Par_{i,t-1}$ (%)	-0.1988***			
	[-3.11]			
$\Delta Par^{low}_{i,t-1}$ (%)		-0.2553***		
		[-3.98]		
$\Delta Par^{mid}_{i,t-1}$ (%)			-0.3514***	
			[-14.27]	
$\Delta Par^{high}_{i,t-1}$ (%)				-0.3538***
				[-21.67]
Controls	Y	Y	Y	Y
County FE	Y	Y	Y	Y
State*Year FE	Y	Y	Y	Y
Obs	21924	21924	21924	21924
adj. R^2	0.033	0.024	0.036	0.071

Metropolitan and Non-Metropolitan area (1)

- ΔFP from metropolitan or non-metropolitan counties only.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent: ΔParticipation$_{i,t}$ (%)					
$\Delta FP_{i,t-1}^{Metro}$ (%)	0.1427***		0.1120**	0.1143**		0.0997**
	[2.73]		[2.38]	[2.38]		[2.27]
$\Delta FP_{i,t-1}^{NonMetro}$ (%)		0.2047***	0.1754**		0.1052	0.0802
		[2.89]	[2.57]		[1.42]	[1.13]
Controls	Y	Y	Y	Y	Y	Y
County FE				Y	Y	Y
State*Year FE	Y	Y	Y	Y	Y	Y
Obs	21924	21924	21924	21924	21924	21924
adj. R^2	0.081	0.081	0.082	0.033	0.033	0.033

Metropolitan and Non-Metropolitan area (2)

	(1) $\Delta Par_{i,t}^{Metro}$	(2) $\Delta Par_{i,t}^{Metro}$ (%)	(3) $\Delta Par_{i,t}^{NonMetro}$	(4) $\Delta Par_{i,t}^{NonMetro}$ (%)
$\Delta FP_{i,t-1}$ (%)	0.3436***	0.3001***	0.2372***	0.1426*
	[3.65]	[3.21]	[3.31]	[1.91]
Controls	Y	Y	Y	Y
County FE		Y		Y
State*Year FE	Y	Y	Y	Y
Obs	8134	8134	13784	13784
R^2	0.356	0.490	0.082	0.165
adj. R^2	0.327	0.373	0.060	-0.002

Metropolitan and Non-Metropolitan area (3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\Delta Par_{i,t}^{Metro}$ (%)				$\Delta Par_{i,t}^{NonMetro}$ (%)			
$\Delta FP_{i,t-1}^{Metro}$ (%)	0.1952**	0.1155			0.1545***	0.1276**		
	[2.05]	[1.21]			[2.71]	[2.41]		
$\Delta FP_{i,t-1}^{NonM}$ (%)			0.2792***	0.2476**			0.1897**	0.1030
			[2.87]	[2.54]			[2.55]	[1.38]
Controls	Y	Y	Y	Y	Y	Y	Y	Y
County FE		Y		Y		Y		Y
State*Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Obs	8134	8134	8134	8134	13783	13783	13783	13783
R^2	0.358	0.488	0.360	0.490	0.081	0.165	0.081	0.164
adj. R^2	0.329	0.371	0.331	0.373	0.059	-0.003	0.059	-0.003

Stock market participation and income inequality (1)

- NPP implication - Participation is beneficial to households.
  Duflo and Saez (2002), Hong, Kubik, and Stein (2004), Ouimet and Tate (2020)

- However, behavioral biases such as overconfidence, representativeness heuristics maybe reinforced through social interaction.
  Heimer (2016), Hirshleifer (2020)

- Kogan, Papanikolaou, and Stoffman (2020):
  Benefits of technology improvement may not flow into households through stock market participation.

Stock market participation and income inequality (2)

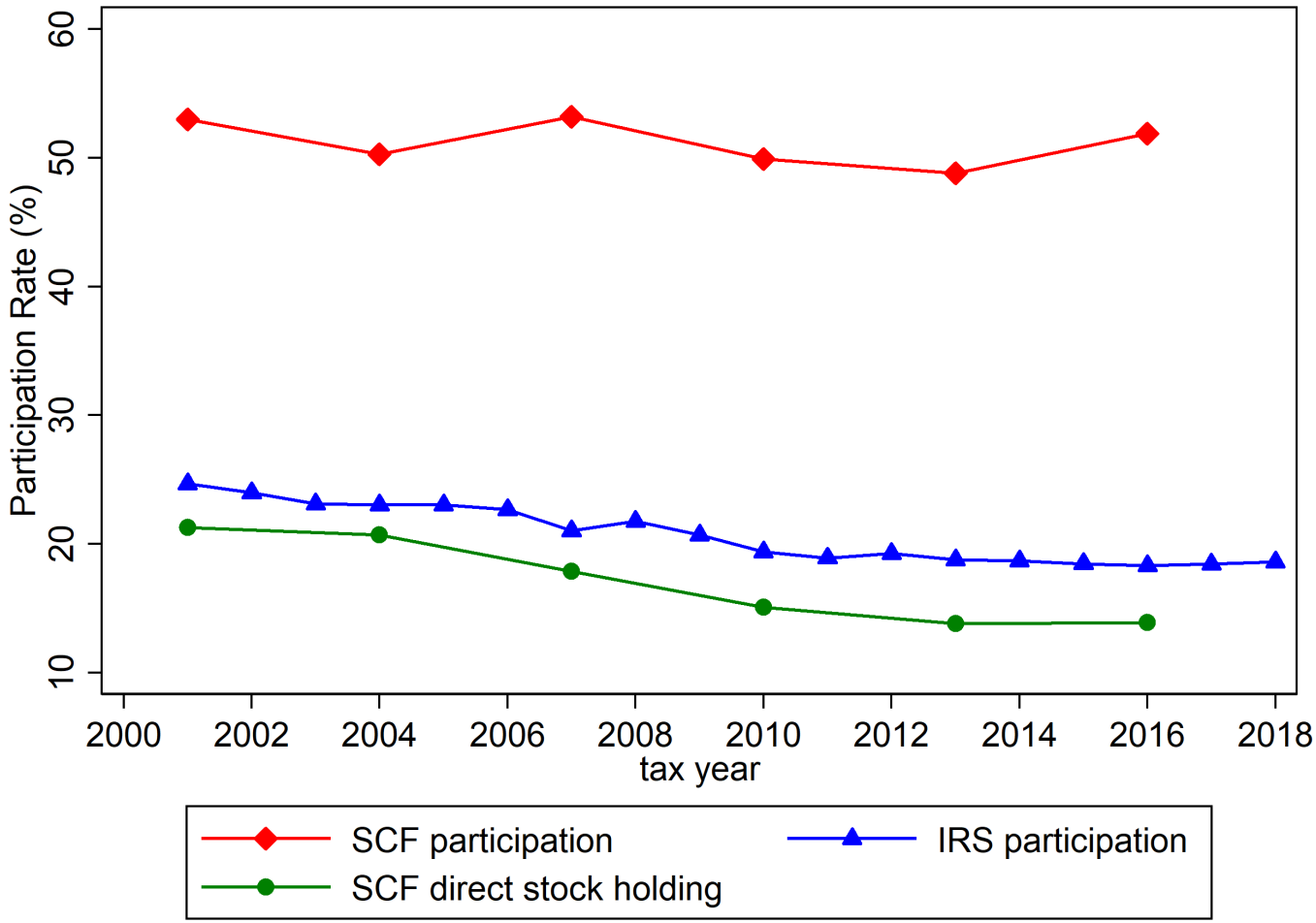
- Calculate Gini coefficients for each county each year.

	(1)	(2)	(3)
	Dependent: $\Delta \text{Gini}_{i,[t-1:t+1]}$		
	Full sample	NonMetro	Metro
$\Delta \widehat{Par^{low}_{i,t-1}}$ (%)	-0.0019	-0.0014	-0.0105***
	[-1.49]	[-1.08]	[-2.96]
$\Delta Par^{low,resi}_{i,t-1}$ (%)	0.0003	0.0003	0.0001
	[0.68]	[0.57]	[0.09]
$\Delta \widehat{Par^{mid}_{i,t-1}}$ (%)	-0.0004	-0.0003	-0.0012
	[-0.78]	[-0.53]	[-1.11]
$\Delta Par^{mid,resi}_{i,t-1}$ (%)	0.0000	0.0000	0.0002
	[0.21]	[0.13]	[0.22]
$\Delta \widehat{Par^{high}_{i,t-1}}$ (%)	0.0000	0.0000	0.0013
	[0.15]	[-0.10]	[1.47]
$\Delta Par^{high,resi}_{i,t-1}$ (%)	-0.0001	-0.0001	-0.0001
	[-0.62]	[-0.62]	[-0.36]
Controls	Y	Y	Y
County FE	Y	Y	Y
State*Year FE	Y	Y	Y
Obs.	15656	9841	5810
adj. R^2	0.129	0.135	0.146

Conclusion

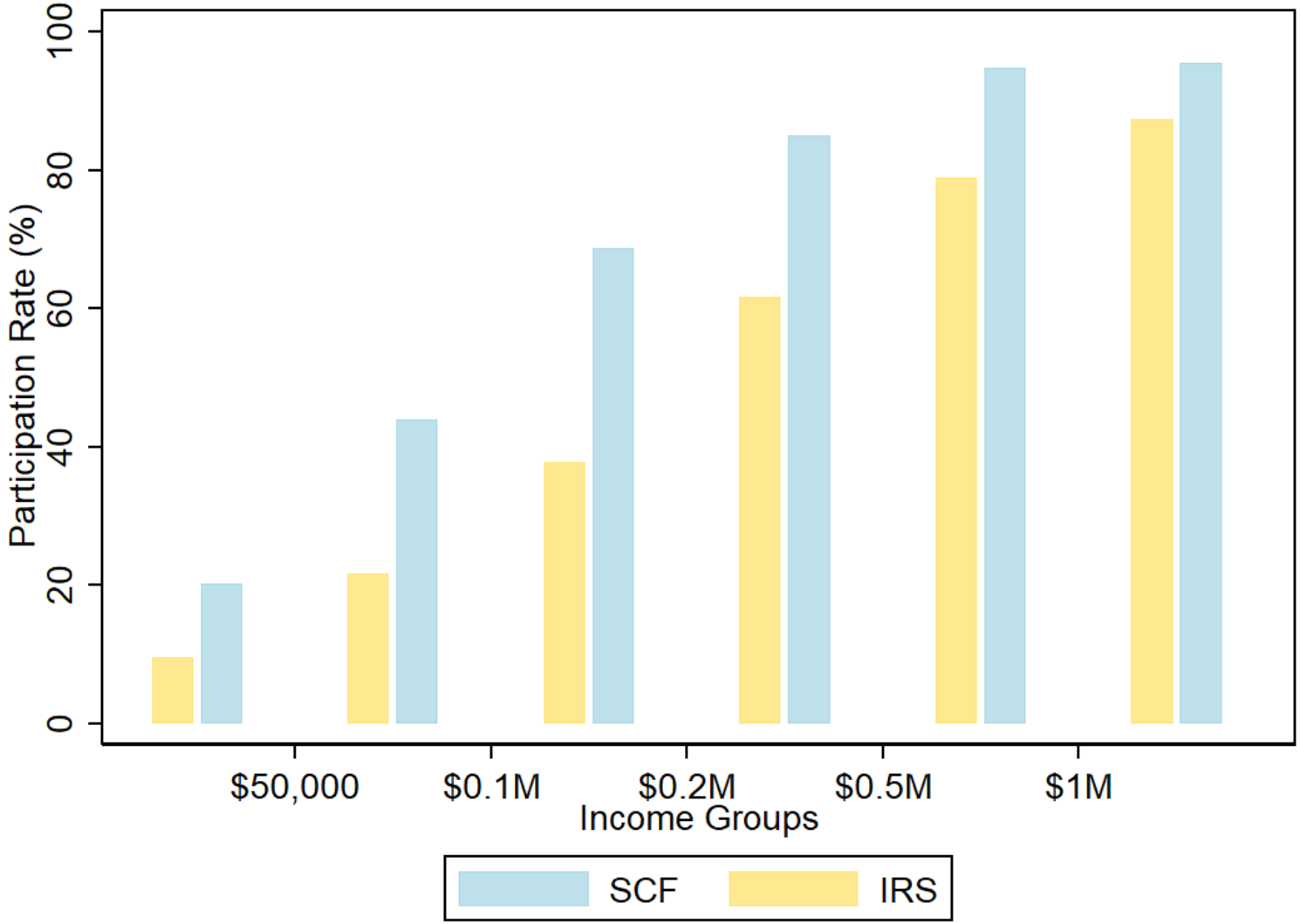
- Document significant heterogeneity in stock market participation rates that traditional explanations for NPP cannot reconcile.
- Friends' Participation positively and significantly predicts participation. 1 percentage point increase in FP leads to 14 to 25 bps increase in the local participation rate in the following year.
- Revelation of financial misconduct as an IV for identification strategy.
- Participation in low-income groups induced by Friends' Participation leads to lower income inequality in metropolitan area.

Stock market participation: IRS and SCF



back

Income cross-section of participation: IRS and SCF



back

Summary Statistics

Variable	Obs	Mean	Median	Std.	Min	Max
Participation Rate (%)	28,195	16.5	16.2	6.7	0	79.6
Participation Rate (<$50K) (%)	28,195	10.5	10.1	5.2	0	83.2
Participation Rate ($50K-$100K) (%)	28,195	21.6	21.5	7.3	0	83.3
Participation Rate (>$100K) (%)	28,195	40.0	40.6	10.9	0	92.3
Participation Rate (Metropolitan) (%)	10,458	17.4	16.9	6.5	1.5	60
Participation Rate (Non-Metropolitan) (%)	17,737	15.9	15.6	6.8	0	79.6
Friend's Participation (%)	28,195	16.5	16.5	2.5	7.8	28.1
ΔFriend's Participation (%)	25,061	-0.10	-0.09	0.29	-6.26	14.73
ΔFriend's Participation^{SameState}(%)	25,061	-0.11	-0.12	0.90	-49.59	35.51
ΔFriend's Participation^{IV}(%)	18,796	-0.12	-0.12	0.09	-3.82	1.28
ΔFriend's Participation^{Metro}(%)	25,062	-0.10	-0.12	0.37	-1.45	19.30
ΔFriend's Participation^{NonMetro}(%)	25,062	-0.10	-0.08	0.28	-6.68	9.95
Probability of Misconduct	25,063	0.009	0	0.058	0	6.816
Gini Coefficient	28,195	0.403	0.394	0.074	0.195	0.726
# of Households	28,195	37,118.6	9,857	112,421.6	22	3,306,109
Ratio higher than Bachelor Degree (%)	28,195	19.3	17.3	8.8	0	80.2
Unemployment Rate (%)	28,195	7.7	7.3	3.6	0	30.9
Median Household Income	28,195	47,079.4	45,166	12,489.4	18,972	136,268
Local Stock Return (%)	28,195	2.2	0	25.1	-662.7	782.4