Nora Marija Laurinaityte
Bank of Lithuania

Understanding Differences in Stock Market Participation: Networks Matter

2021 Canadian Stata Conference
September 23rd , 2021

Background

- Average equity risk premium from 1928 to 2016 is estimated to be around 8% p.a. in the US.
- In other developed markets the differences between the returns of risky and safe assets have also been relatively large in the last decades.
- It is puzzling that many households do not participate in the stock market (“participation puzzle”).
- Various explanations have been proposed (Mankiw and Zeldes, 1991, Gollier and Pratt, 1996, Vissing-Jorgensen, 2002, Mehra and Prescott, 2003, Paiella, 2007, Malmendier and Nagel, 2011, Christelis et al, 2013, etc.)

Background

- Chien and Morris (St Louis Fed,2017) show a large variation in stock market participation (SMP) rates across the US states.

- This heterogeneity across states cannot be explained by income (participation gap remains large for each income group), costs of living (lower participation rate in states with lower costs of living when keeping HH income constant), or tax schedules (lower participation rates in states with lower top marginal tax rate on personal dividends income and capital gains).

- Authors suggests there are regional (cultural) factors that are affecting the SMP rates (namely, influence of surrounding community).

Background

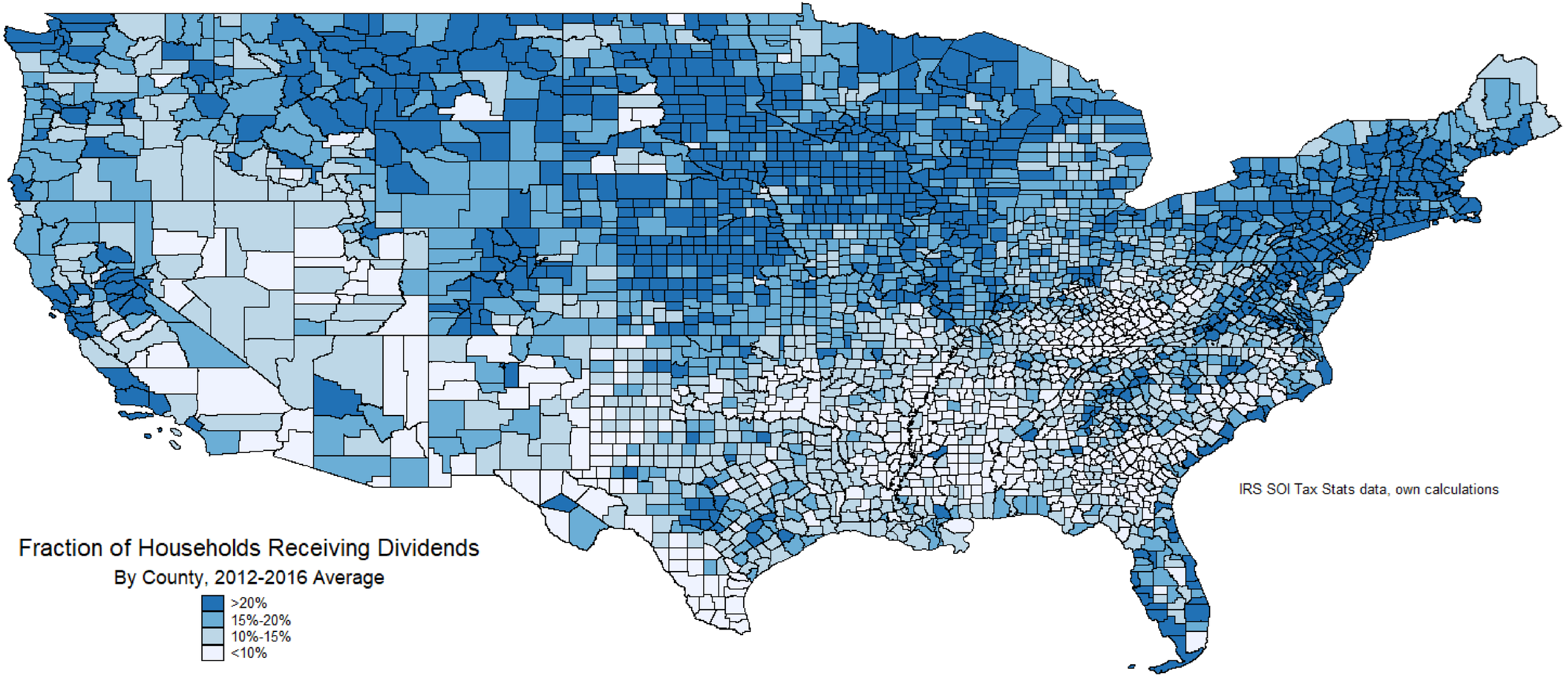
- Hong et al (2004), Ivkovic and Weisbenner (2007), Brown et al. (2008) & (2014) , Christelis et al (2010), Kaustia and Knüpfer (2012) find social households are more likely to invest due to peer effects.

- Balakina and Parakhonyak (2019) introduce an equilibrium model of the SMP rate with a social network. Model with social networks (information sharing through a network affects the decision to enter the stock market) does a good job in predicting heterogeneity of equilibrium stock market participation across the income distribution.
  - By allowing the SMP costs to vary with connectivity, can create heterogeneity in SMP levels even among agents with similar characteristics.

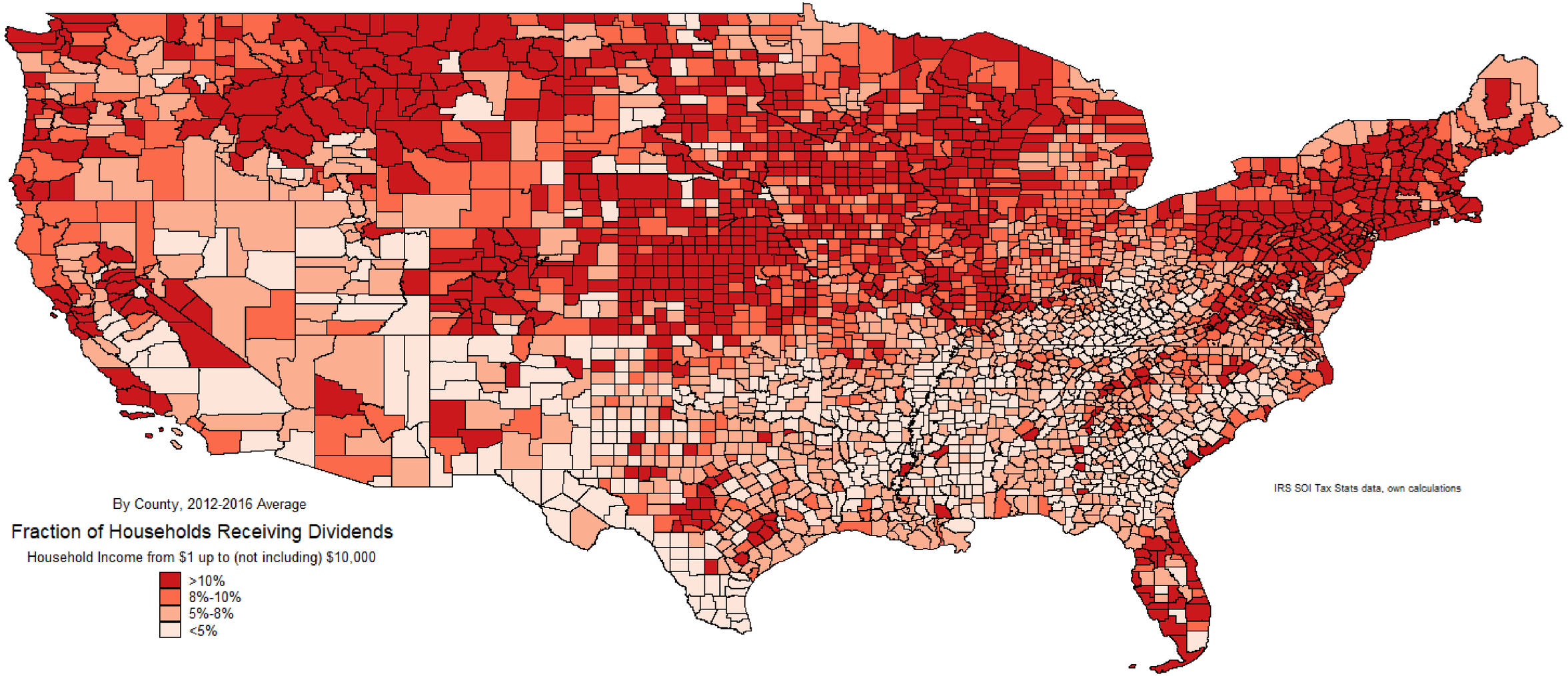
Research Idea

- Explore the variation in the SMP rates on county level across the income distribution.

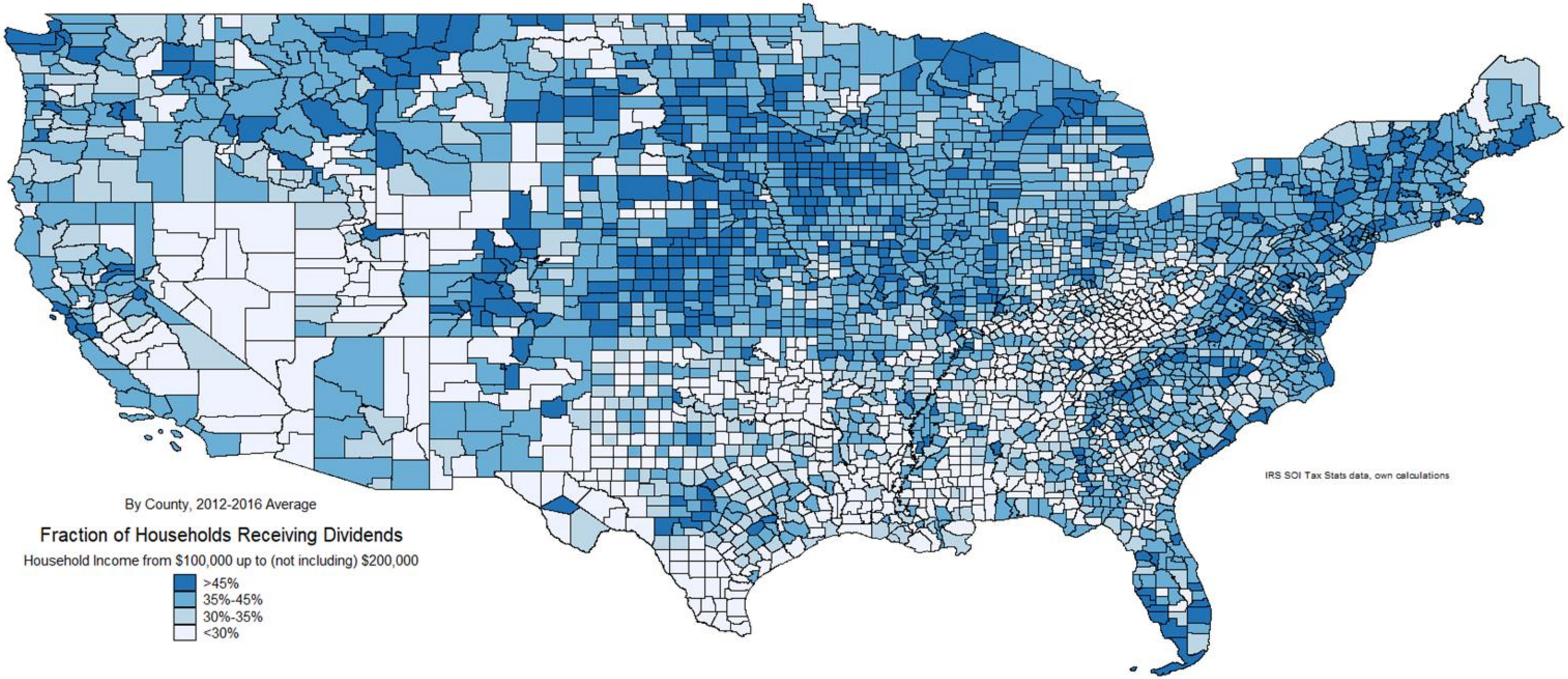
Huge Heterogeneity in SMP rates geographically.



SMP heterogeneity remains large at any HH income level.



SMP heterogeneity remains large at any HH income level.



Research Idea

- Explore the variation in the SMP rates on county level across the income distribution.
- How much of this cross-county SMP heterogeneity can be explained by traditional determinants of SMP (education, demographics, political affiliation)?
- Once traditional factors are accounted for, do features of county´s social network play a role in explaining this heterogeneity?

Results (preview)

- Traditional determinants of SMP explain the observed cross-county heterogeneity rather well, on average.
- Traditional determinants fail to explain SMP heterogeneity across the income distribution, in particular for the income rich households.
- An empirical model that accounts for county´s network, namely when the average SMP rates in the connected counties are included as determinants of county SMP, outperforms the traditional framework.
- SMP rate among households at the same income level in the county´s network is particularly important covariate of SMP for the high income level households.

Data

- **IRS' individual income and tax data for years 2012-2016**

I approximate the stock market participation rate as the (5-year average) ratio of the number of tax returns with dividend income to the number of total tax returns filed. My figures underestimate the true participation rate and should therefore be considered a lower bound. Fortunately, this shortcoming might be less of a concern when comparing participation rates across counties, as the downward bias is likely to affect counties uniformly. (as in Chien and Morris (2017) , Bäckman and Hanspal (2018))

- **Social Connectedness Index (SCI)** from Bailey et al (2017).This measure is an index based on the number of Facebook connections. Non-directional snapshot, April 2016.

- **US Census Bureau American Community Survey (ACS)** for gender, age, race, education composition

- Voting information (2016 Presidential election) from Townhall.com

How much of the cross-county SMP heterogeneity can be explained by traditional determinants of SMP?

County-level SMP Covariates

VARIABLES	(1) SMP	(2) SMP	(3) SMP	(4) SMP	(5) SMP
Log Median Income		0.132***	0.0995***	0.0475***	0.0415***
		(0.00669)	(0.0154)	(0.0109)	(0.0124)
Average Household Size			-0.0596***	-0.0224***	-0.0222***
			(0.00638)	(0.00554)	(0.00496)
Fraction Without a College Degree				-0.322***	-0.321***
				(0.0191)	(0.0196)
Unemployed				-0.133***	-0.114***
				(0.0437)	(0.0422)
Fraction Voting in 2016					0.0424***
					(0.0146)
Fraction Voters Voting Democrat 2016					-0.0227**
					(0.00942)
State Dummies	YES	YES	YES	YES	YES
Observations	3,101	3,101	3,099	3,099	3,099
R-squared	0.448	0.629	0.781	0.837	0.839

Robust standard errors clustered at State level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

SMP Covariates Over the Income Distribution

VARIABLES	(1) SMP2	(2) SMP3	(3) SMP4	(4) SMP5	(5) SMP6	(6) SMP7
Log Median Income	0.0156	0.0250***	-0.0118	-0.0142	-0.0263	-0.0302
	(0.00989)	(0.00710)	(0.0118)	(0.0163)	(0.0213)	(0.0318)
Average Household Size	0.00315	-0.0158***	-0.0277***	-0.0325***	-0.0672***	-0.0615***
	(0.00879)	(0.00449)	(0.00694)	(0.00854)	(0.00799)	(0.0162)
Fraction Without a College Degree	-0.170***	-0.116***	-0.165***	-0.275***	-0.223***	-0.422***
	(0.0234)	(0.0189)	(0.0235)	(0.0337)	(0.0345)	(0.0609)
Unemployment Rate	-0.119**	-0.105***	-0.0774**	-0.102	-0.259***	-0.455***
	(0.0583)	(0.0337)	(0.0382)	(0.0775)	(0.0888)	(0.137)
Fraction Voting in 2016	0.0545***	0.0540***	0.0604**	0.108***	-0.0561	-0.00732
	(0.0181)	(0.0148)	(0.0257)	(0.0289)	(0.0472)	(0.0520)
Fraction Voters Voting Democrat 2016	-0.0314***	-0.00771	-0.0299***	-0.0277	0.0332	-0.00496
	(0.0102)	(0.00967)	(0.0109)	(0.0166)	(0.0221)	(0.0299)
State Dummies	YES	YES	YES	YES	YES	YES
Observations	3,105	3,105	3,106	3,105	3,100	3,096
R-squared	0.564	0.719	0.698	0.667	0.597	0.530

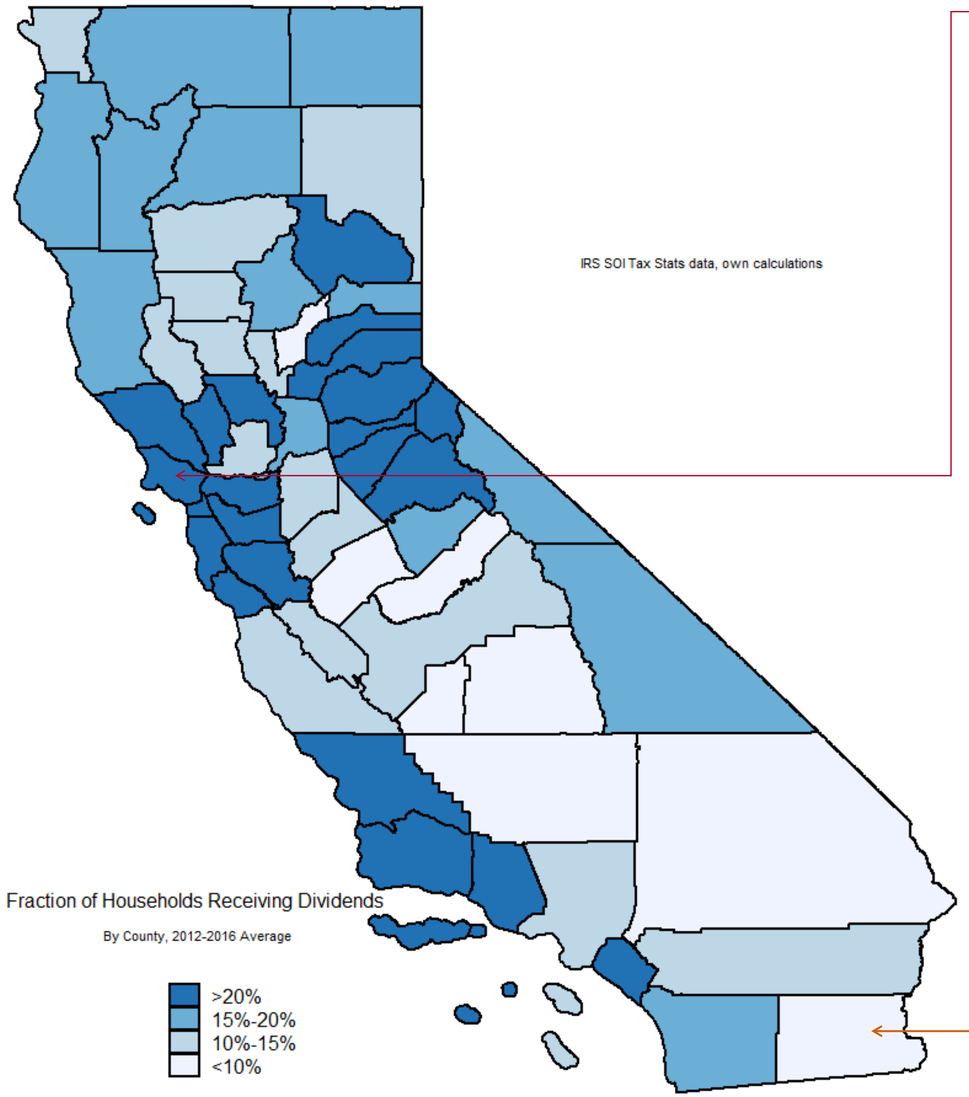
Robust standard errors clustered at State level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

- Traditional determinants do pretty well in explaining the average cross-county SMP heterogeneity.

- Traditional determinants do less well in explaining the cross-county heterogeneity in SMP rates among households in the same income group, in particular among the income-rich households (with income between $100,000 up to, but not including, $200,000) and the most income-poor households (with income between $1 and up to, but not including, $10,000)

Do SMP rates in county´s social network play a role in explaining cross-county SMP heterogeneity?

Intuition. Take California, for example….



Marin County

Population 189,717 (104,846 HHs)

Median HH income (2017) $112,735

41% of households receive dividends

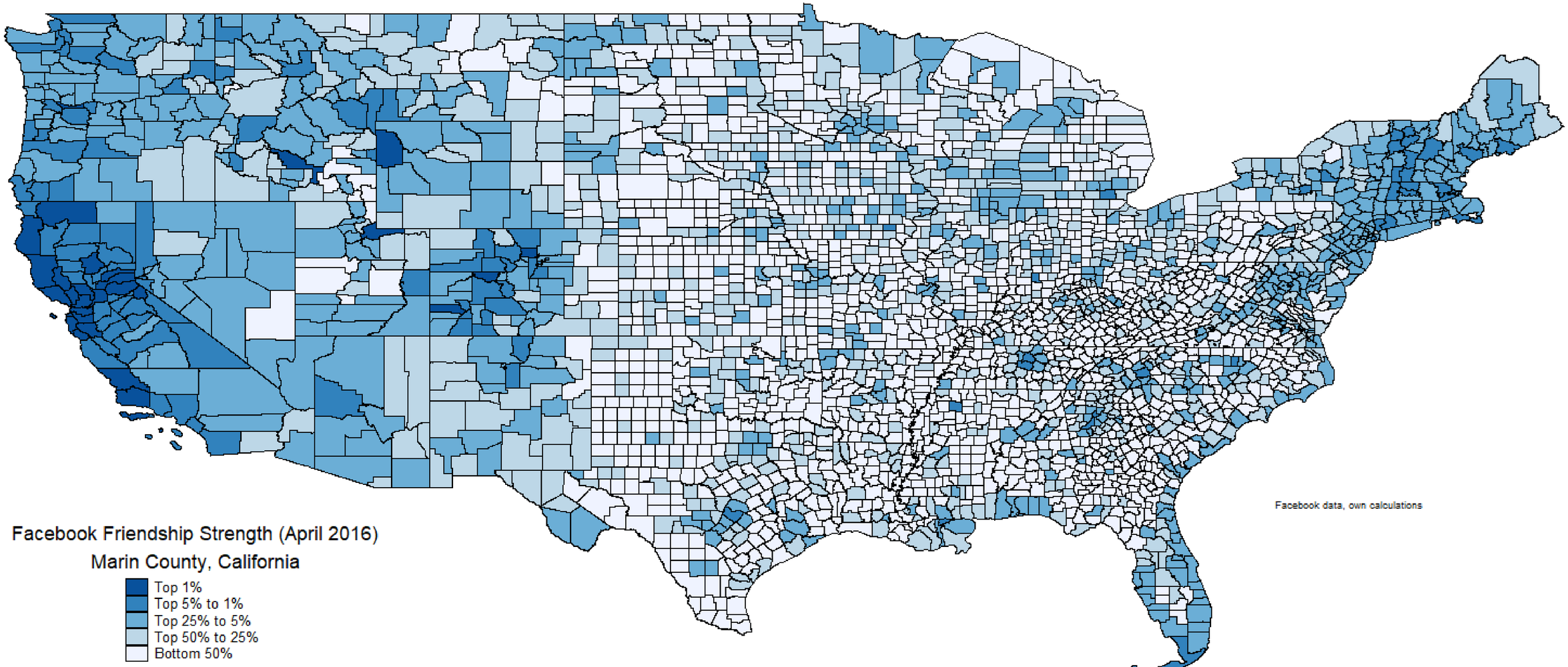
Imperial County

Population 107,679 (45,198 HHs)

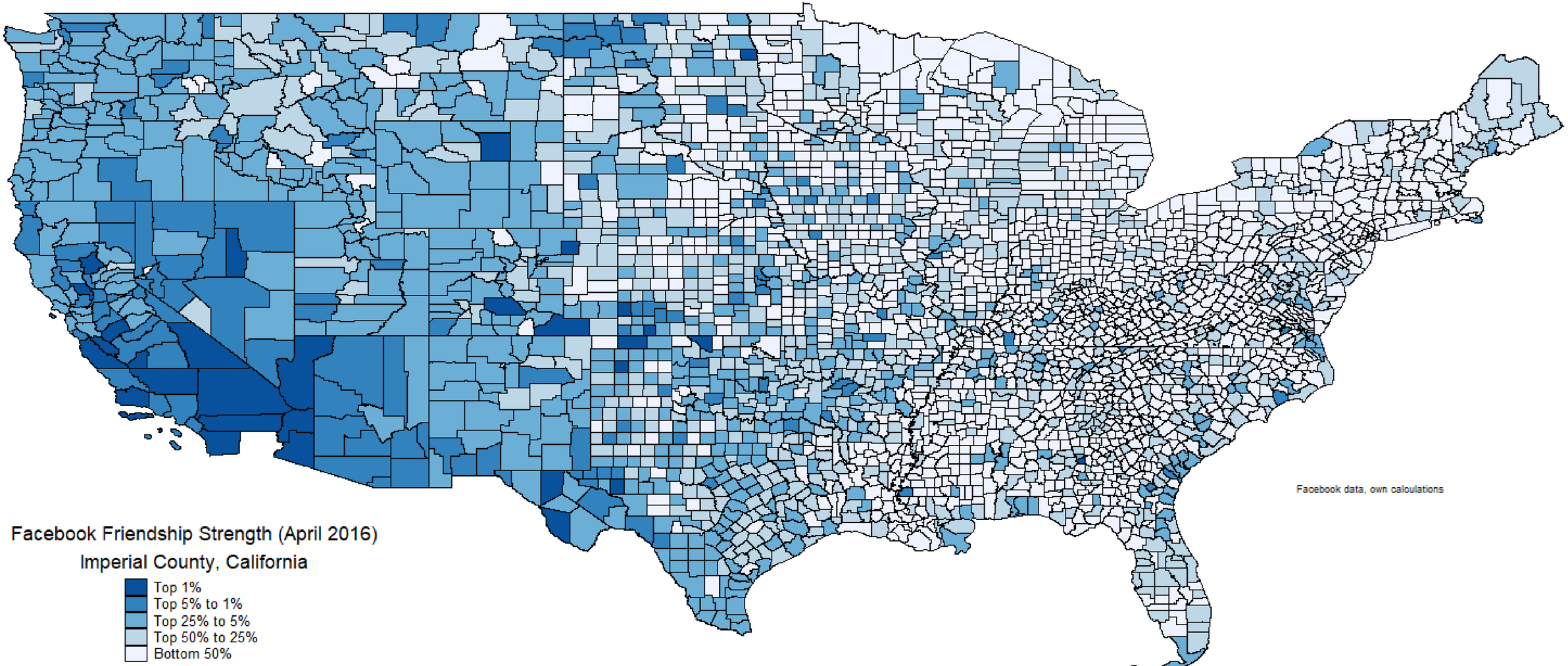
Median HH income (2017) $43,413

4% of households receive dividends

Where Do the Marin County Facebook Friends Live?



Where Do the Imperial County Facebook Friends Live?



Do Their Facebook Friends Own Stocks?

Marin County SMP is 41%. Imperial County SMP is 4%.

Take their „top 1%“ friend counties (by relative probability of friendship on Facebook).

- **Marin County**: friend county average SMP 24%. In the out-of-state friend counties, on average, 28% of their population hold stocks.
- **Imperial County**: friend county average SMP 13%. In the out-of-state friend counties, on average, less than 13% of their population hold stocks.

County SMP and the SMP in County Network

VARIABLES	(1) SMP	(2) SMP	(3) SMP	(4) SMP
Average SMP of 30 „best friend“ counties		0.603***		0.497***
		(0.0623)		(0.0839)
Average SMP of 310 „best friend“ counties			0.928***	
			(0.124)	
Average SMP of „best friend“ counties from out-of-state				0.374***
				(0.112)
Log Median Income	0.0506***	0.0345***	0.0483***	0.0372***
	(0.00650)	(0.00503)	(0.00583)	(0.00472)
Average Household Size	-0.0514***	-0.0428***	-0.0484***	-0.0434***
	(0.00535)	(0.00565)	(0.00459)	(0.00524)
Fraction Without a College Degree	-0.271***	-0.271***	-0.239***	-0.252***
	(0.0173)	(0.0145)	(0.0170)	(0.0130)
Unemployment Rate	-0.0491	-0.00424	-0.0135	0.00872
	(0.0438)	(0.0374)	(0.0401)	(0.0359)
Fraction Voting in 2016	0.126***	0.113***	0.115***	0.108***
	(0.0203)	(0.0190)	(0.0192)	(0.0188)
Fraction Voters Voting Democrat 2016	-0.0249**	-0.0439***	-0.0443***	-0.0522***
	(0.0110)	(0.00849)	(0.00876)	(0.00886)
State Dummies	YES	YES	YES	YES
Observations	3,101	3,101	3,101	3,101
R-squared	0.806	0.829	0.822	0.832

Robust standard errors clustered at State level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

County SMP and the SMP in County Network

VARIABLES	(1) SMP2	(2) SMP3	(3) SMP4	(4) SMP5	(5) SMP6	(6) SMP7
Average SMP2 of 30 „best friend“ counties	0.609***					
	(0.183)					
Average SMP2 of „best friend“ counties from out-of-state	0.327**					
	(0.126)					
Average SMP3 of 30 „best friend“ counties		0.518***				
		(0.0724)				
Average SMP3 of „best friend“ counties from out-of-state		0.390***				
		(0.109)				
Average SMP4 of 30 „best friend“ counties			0.622***			
			(0.0674)			
Average SMP4 of „best friend“ counties from out-of-state			0.317***			
			(0.111)			
Average SMP5 of 30 „best friend“ counties				0.598***		
				(0.0970)		
Average SMP5 of „best friend“ counties from out-of-state				0.482***		
				(0.107)		
Average SMP6 of 30 „best friend“ counties					0.798***	
					(0.0687)	
Average SMP6 of „best friend“ counties from out-of-state					0.319***	
					(0.108)	
Average SMP7 of 30 „best friend“ counties						0.883***
						(0.0630)
Average SMP7 of „best friend“ counties from out-of-state						0.228*
						(0.129)
Log Median Income	0.00404	0.0165***	-0.00865	-0.0281***	-0.0166	-0.0651***
	(0.00766)	(0.00411)	(0.00688)	(0.00863)	(0.0105)	(0.0180)
Average Household Size	-0.0178***	-0.0325***	-0.0391***	-0.0535***	-0.0711***	-0.0716***
	(0.00631)	(0.00420)	(0.00455)	(0.00628)	(0.00606)	(0.0112)
Fraction Without a College Degree	-0.113***	-0.0510***	-0.139***	-0.232***	-0.266***	-0.390***
	(0.0132)	(0.0130)	(0.0170)	(0.0195)	(0.0197)	(0.0326)
State Dummies	YES	YES	YES	YES	YES	YES
Observations	3,098	3,099	3,100	3,099	3,093	3,089
R-squared	0.565	0.708	0.694	0.669	0.615	0.572

Robust standard errors clustered at State level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Accounting for Financial Misconduct, Self-Employment, and Working in Finance

VARIABLES	(1) SMP	(2) SMP2	(3) SMP3	(4) SMP4	(5) SMP5	(6) SMP6	(7) SMP7
Average SMP of 30 „best friend“ counties	0.159**						
	(0.0753)						
Average SMP2 of 30 „best friend“ counties		0.318***					
		(0.110)					
Average SMP3 of 30 „best friend“ counties			0.168*				
			(0.0853)				
Average SMP4 of 30 „best friend“ counties				0.162*			
				(0.0856)			
Average SMP5 of 30 „best friend“ counties					0.241**		
					(0.115)		
Average SMP6 of 30 „best friend“ counties						0.287***	
						(0.0805)	
Average SMP7 of 30 „best friend“ counties							0.404***
							(0.118)
Financial Adviser Misconduct Rate as of 2015	0.145***	0.120***	0.0686**	0.105***	0.185***	0.213***	0.234**
	(0.0539)	(0.0348)	(0.0323)	(0.0377)	(0.0649)	(0.0785)	(0.0888)
Log Median Income	0.0942***	0.0496***	0.0485***	0.0223*	-0.00944	-0.0283	-0.0615**
	(0.0187)	(0.0135)	(0.0123)	(0.0125)	(0.0205)	(0.0248)	(0.0290)
Fraction Self Employed	-0.445	-0.515	-1.112*	-1.150*	-0.975	-0.254	0.117
	(0.716)	(0.599)	(0.650)	(0.591)	(0.857)	(1.055)	(1.272)
Fraction in Finance as a Profession	-0.399	-0.00749	-0.0209	-0.453	0.237	-1.485	1.150
	(2.023)	(1.243)	(0.885)	(1.046)	(1.447)	(2.098)	(1.894)
Average Household Size	-0.0688***	-0.0323***	-0.0427***	-0.0520***	-0.0679***	-0.0817***	-0.0927***
	(0.0151)	(0.0101)	(0.00949)	(0.00996)	(0.0152)	(0.0183)	(0.0234)
Fraction Without a College Degree	-0.253***	-0.113***	-0.0292	-0.0915***	-0.192***	-0.286***	-0.410***
	(0.0386)	(0.0275)	(0.0215)	(0.0258)	(0.0389)	(0.0445)	(0.0524)
State Dummies	YES	YES	YES	YES	YES	YES	YES
Observations	471	471	471	471	471	471	471
R-squared	0.880	0.814	0.855	0.840	0.819	0.782	0.744

Robust standard errors clustered at State level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Conclusion

- There is substantial geographic heterogeneity in SMP rates across the US.
- Traditional SMP determinants do well in explaining the average cross-county SMP heterogeneity, but less well in explaining the cross-county heterogeneity in SMP rates among households in the same income group, especially for income rich households.
- County residents´ connections to others via a social network are important in explaining cross-county SMP heterogeneity given the HH income level. An empirical model that accounts for the average SMP rates in the county´s network in addition to traditional SMP determinants improves in terms of explanatory power.

Conclusion

- This finding highlights the importance of social connectivity in household´s investment decisions, and sheds light on why some wealthy households do not hold stocks.

- Findings suggest, that non-participating wealthy households observe low SMP for similar households in their social network.

- However, whether they then do not enter equity markets themselves because of desire to conform to what they perceive as social norm, or because they fail to benefit from some informational spill-overs compared to households experiencing high SMP rates in their network, cannot be established in the present setting.

Thank you for your attention!

Social Network Data

- **"Social Connectedness Index" (SCI)** is based on the number of friendship links on Facebook.

- For each county pair, this provides a relative measure of the total number of Facebook friendship links between individuals located in the two counties as of April 2016.

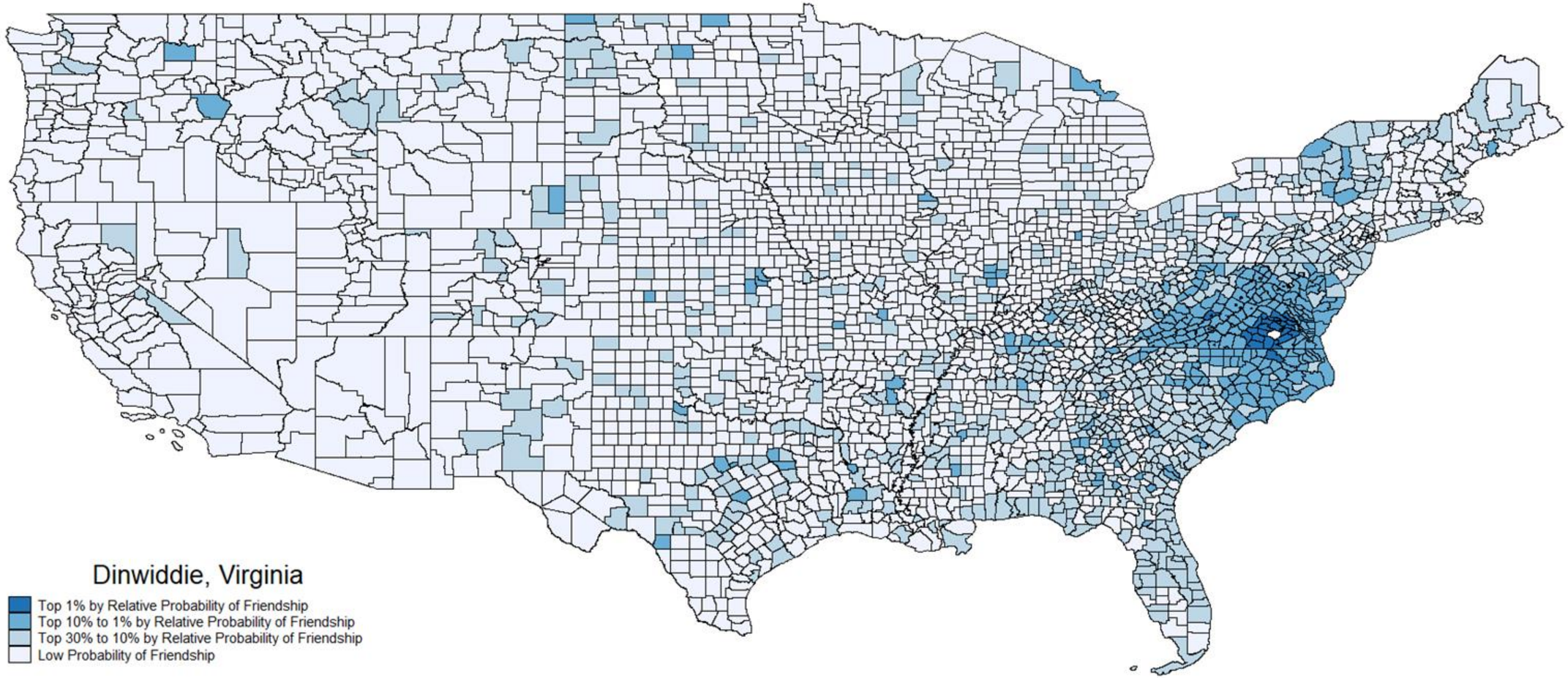
- The data is normalized to have a maximum value of 1,000,000, which is assigned to Los Angeles to Los Angeles connections, the county pair with the largest number of friendship links. The normalized data is then rounded to the nearest 0.0025.

- For the county-county pairs, the relative probability of friendship is calculated as

  rel_prob_friend = 10^12 * sci / (pop_i * pop_j)

- The scaling by 10^12 is introduced to minimize the number of decimal places. Since absolute values of rel_prob_friendship are not meaningful, this re-scaling does not affect the interpretation.

Where do the Facebook Friends live?



Where do the Facebook Friends live?