

# ソーシャル・キャピタル、その1： 測定および世代間所得移動との関連

Chetty, Raj et al. 2022. “Social Capital I: Measurement and Associations with Economic Mobility.” *Nature*. <http://dx.doi.org/10.1038/s41586-022-04996-4>

麦山 亮太 Ryota Mugiyama

学習院大学法学部政治学科

[ryota.mugiyama@gakushuin.ac.jp](mailto:ryota.mugiyama@gakushuin.ac.jp)

# 背景

ソーシャル・キャピタル（個人の社会的ネットワークやコミュニティの強さ）は、教育から健康に至るまで様々な成果を決定する可能性があるとされてきた。

しかし、どのような種類のソーシャル・キャピタルがこれらのアウトカムに重要であるかを理解する試みは、ソーシャルネットワークデータの不足によって妨げられてきた。

# 目的

Facebookの210億の友人関係のデータを用いて、以下の3つの観点からCountyレベルのソーシャル・キャピタルの多寡を測定する

1. 異なる種類の人びととのつながり Cross-type connectedness (Loury, Lin, Bourdiew)
2. ネットワーク凝集性 Network cohesiveness (Coleman)
3. 市民的参加 Civic engagement (Putnam)

作成したソーシャル・キャピタルの指標を用いて、どのソーシャル・キャピタルが世代間の所得上昇移動を促すのかを明らかにする。

# データ

**Facebook**ユーザーの全データ（Facebook社から提供された匿名化されていない全情報が入ったデータ）

うち、アメリカに居住する25–44歳の（Facebookの使用率は84%）、過去30日以内にログインした、100人以上の友人がいる、ZIP-code（日本でいう郵便番号）が欠損でない者を対象

個人レベルのデータからCountyレベルに集計したデータを構築して、Countyごとのソーシャル・キャピタル指標と上昇移動率との関連を分析する

# ソーシャル・キャピタルの測定

# 1. 経済的つながり指標 Economic Connectedness

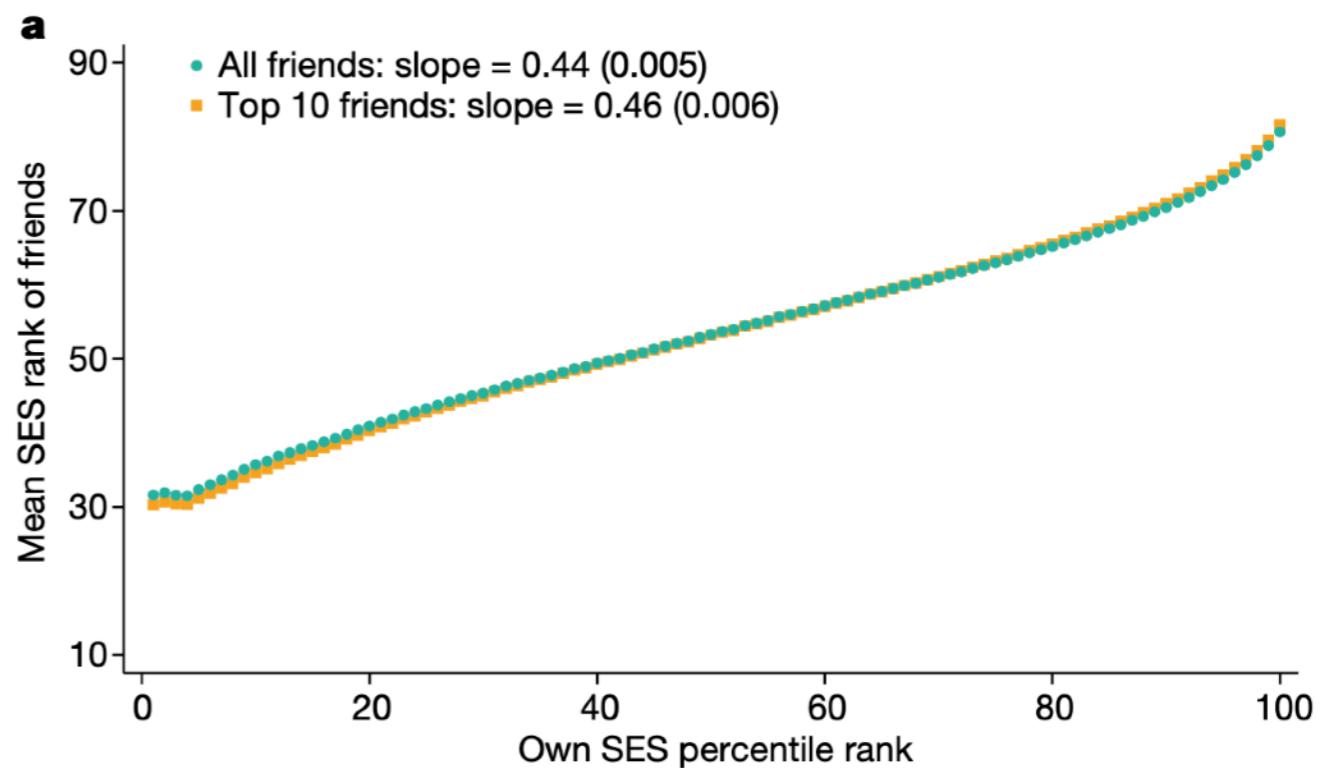
あるCountyに住むLow-SESの各個人について、その友人に占めるHigh-SESの個人の割合を計算する。この値をCountyごとに平均し、0.5で割った指標を経済的つながり指標とする。最小値は0で、1であれば友人の半分はHigh-SESであることを表す。

County	Individual	Individual's SES	Friend	Friends' SES	EC to high SES
1	1	Low	1	Low	<b>0.25</b>
1	1	Low	2	High	
1	1	Low	3	Low	
1	1	Low	4	Low	
1	2	High	1	High	<b>0.5</b>
1	2	High	2	High	
1	2	High	3	Low	
1	2	High	4	Low	
1	3	Low	1	Low	<b>0</b>
1	3	Low	2	Low	

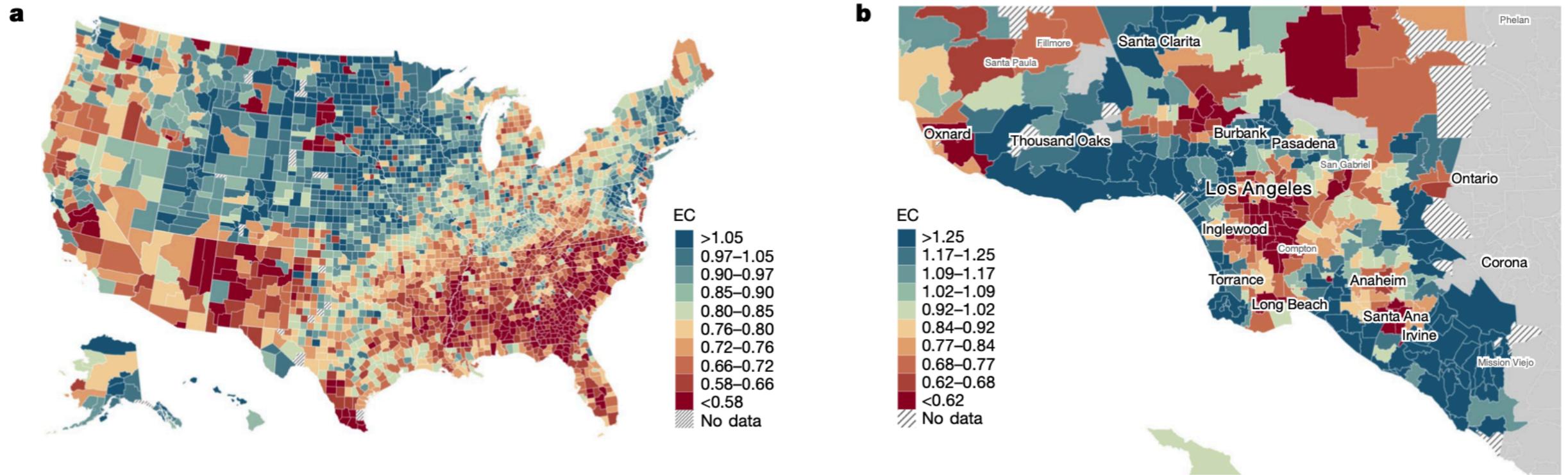
Low: SESが中央値以下、High: SESが中央値以上。

# SESの測定

1. 各個人に付随するFacebookの位置情報（またはZIPコード）を国勢調査のブロックにマッチングさせて、当該ブロックにおける世帯年収の中央値を外挿
2. 割り当てられた世帯年収の値をFacebookから利用できるすべての情報で回帰して（機械学習で最も予測力の高いモデルを作る）、予測値を割り当てる。
3. 同年齢内ごとに100分位のランクを作成して各個人をランクごとに割り当てる。これをSESの指標とする
4. 個人の友人、親（後述）についても同じようにしてSESの指標を作成する



# County別にみた経済的つながり指標の分布



上記は現在25–44歳の対象者から作成した結果であるが、子ども時代のSESを用いて作成したCounty別変数との相関係数は0.61と高く、時系列的に安定している言語的つながり、年齢的つながりに関する変数も作成

**H1.** 経済的つながりにめぐまれているほど、上昇移動が起こりやすい

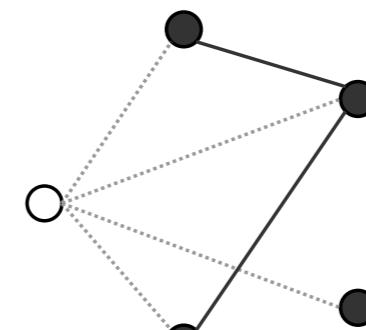
## 2. ネットワーク凝集性 Network cohesiveness

個人レベルおよびCountyレベルの3つの方法で測定する

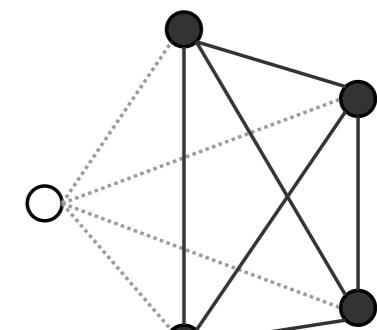
1. 密度 **Clustering** : 個人のもつネットワークの密度 (0-1)

2. サポート比率 **Support ratio** : 個人のもつつながりのうち、共通の友人（トライアド）を有する割合 (0-1)

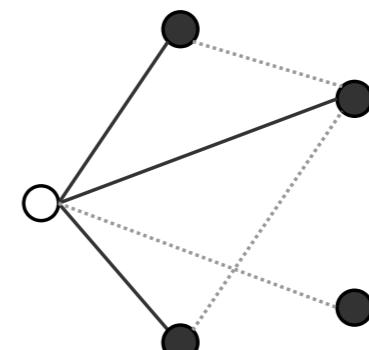
3. Countyの分断度 **Spectral homophily** :  
Countyのホールネットワークがどの程度  
クラスタに分かれているか (0-1)



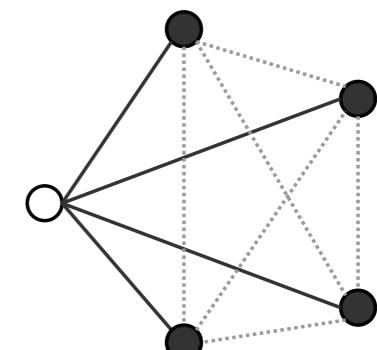
$$\text{密度} = 2/6 = 0.333$$



$$\text{密度} = 6/6 = 1$$



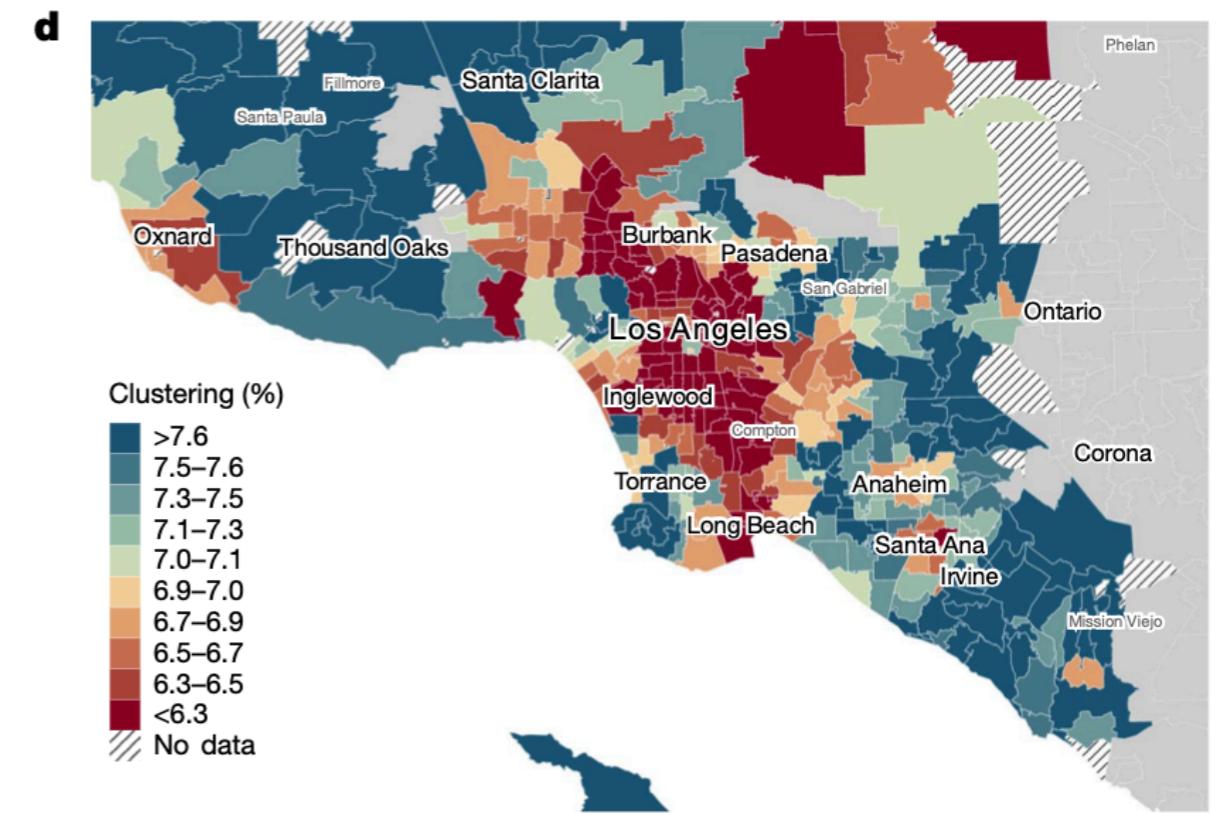
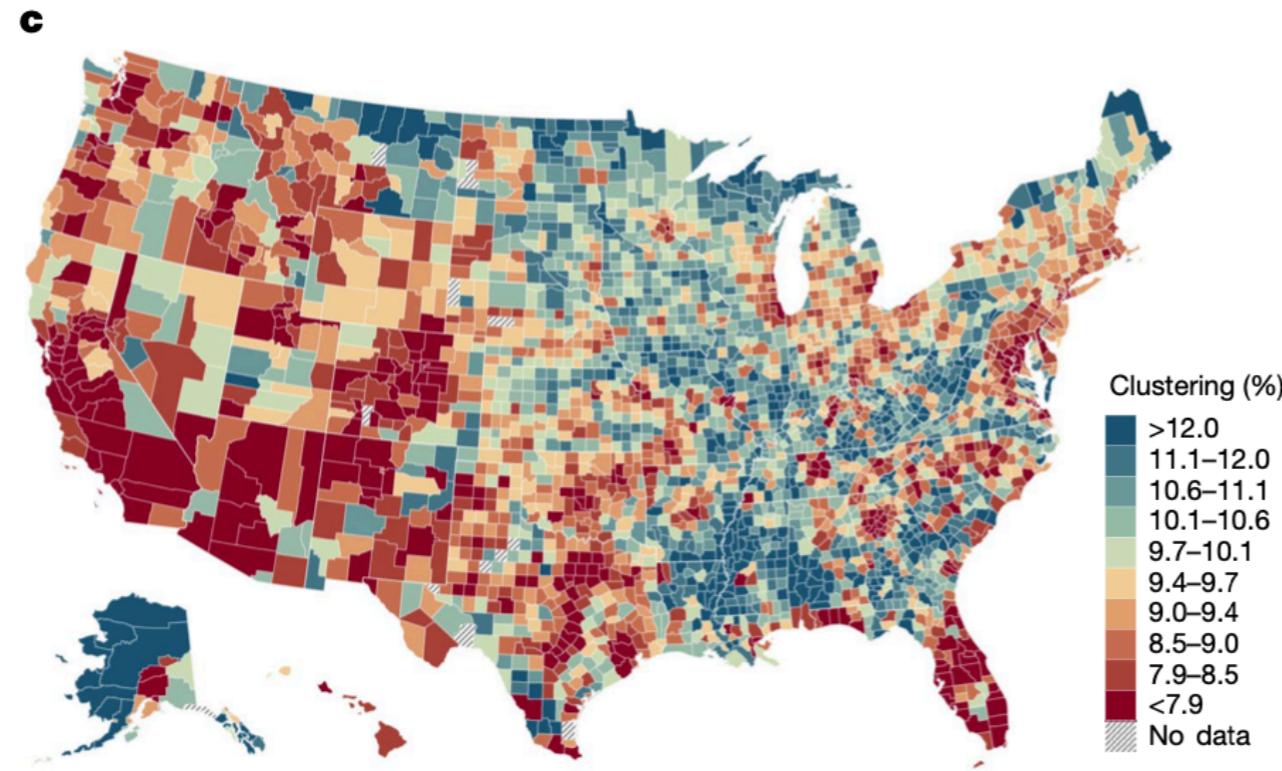
$$\text{サポート比率} = 3/4 = 0.75$$



$$\text{サポート比率} = 4/4 = 1$$

# County別にみたネットワーク凝集性（密度）の分布

地区によってネットワーク凝集性の程度にはばらつきがある



H2. ネットワーク凝集性が高いほど、（社会規範が遵守されやすいと考えられるため）上昇移動が起こりやすい

### 3. 市民的参加 Civic engagement

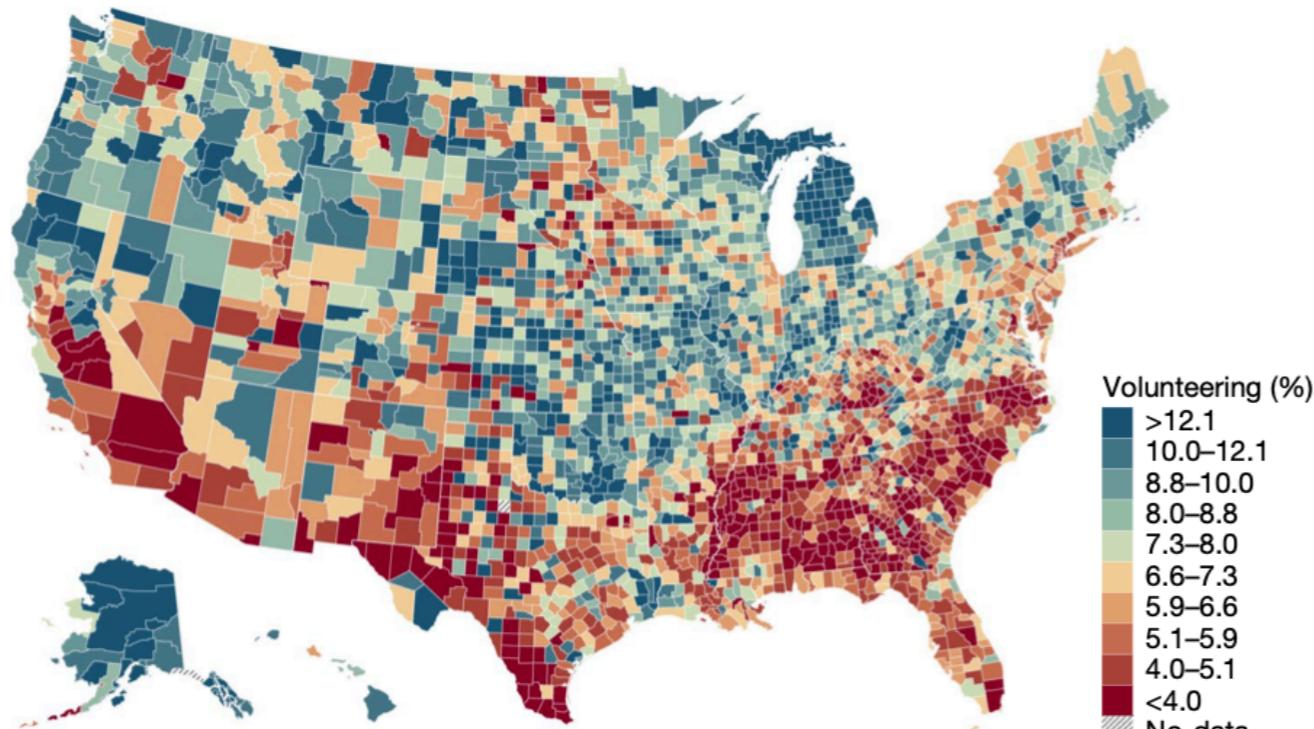
Countyレベルおよび集合レベルの3つの方法で測定する

1. **Countyにおけるボランティア団体加入率**：個人が加入する団体の名前からボランティア団体に加入しているかを特定し、Countyごとに割合をとる (0-1)
2. **市民団体**：Facebookのアカウントのうち“public goods”に分類されるアカウントのzip codeをもとに、CountyのFacebookユーザー1000人に対していくつの中間団体があるかを計算 (0-)
3. **Penn-State index**：先行研究が作成したCountyごとの市民的参加の指標

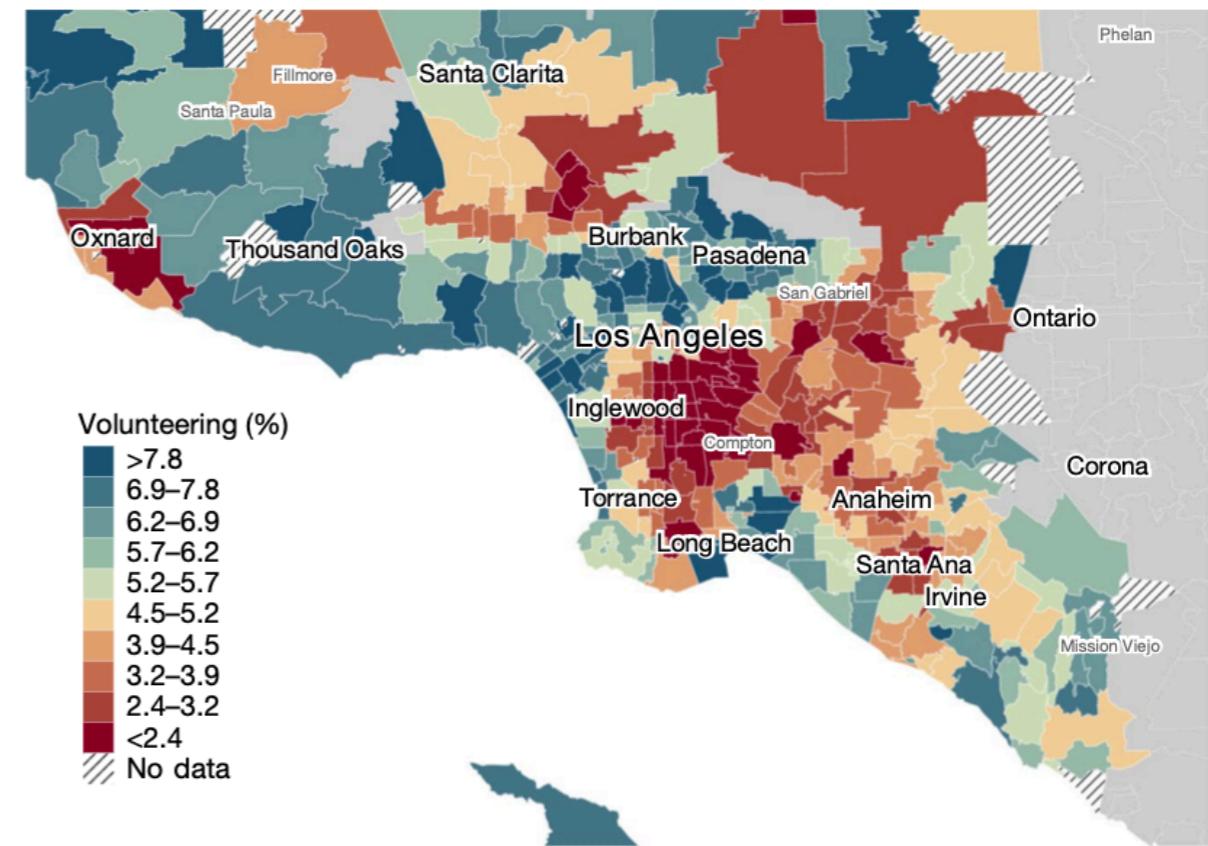
# County別にみた市民的参加（ボランティア）の分布

地区によって市民的参加にはばらつきがある

e



f



H3. 市民的参加が多いほど、（社会規範が遵守されやすいと考えられるため）上昇移動が起こりやすい

# 各指標の相関係数

各指標の相関は顕著に高いわけではなく、それぞれが独立の側面を測定している。

コミュニティにおけるソーシャル・キャピタルを測定する際には異なる側面を分け

**Table 1 | Correlation matrix for social capital measures across counties**

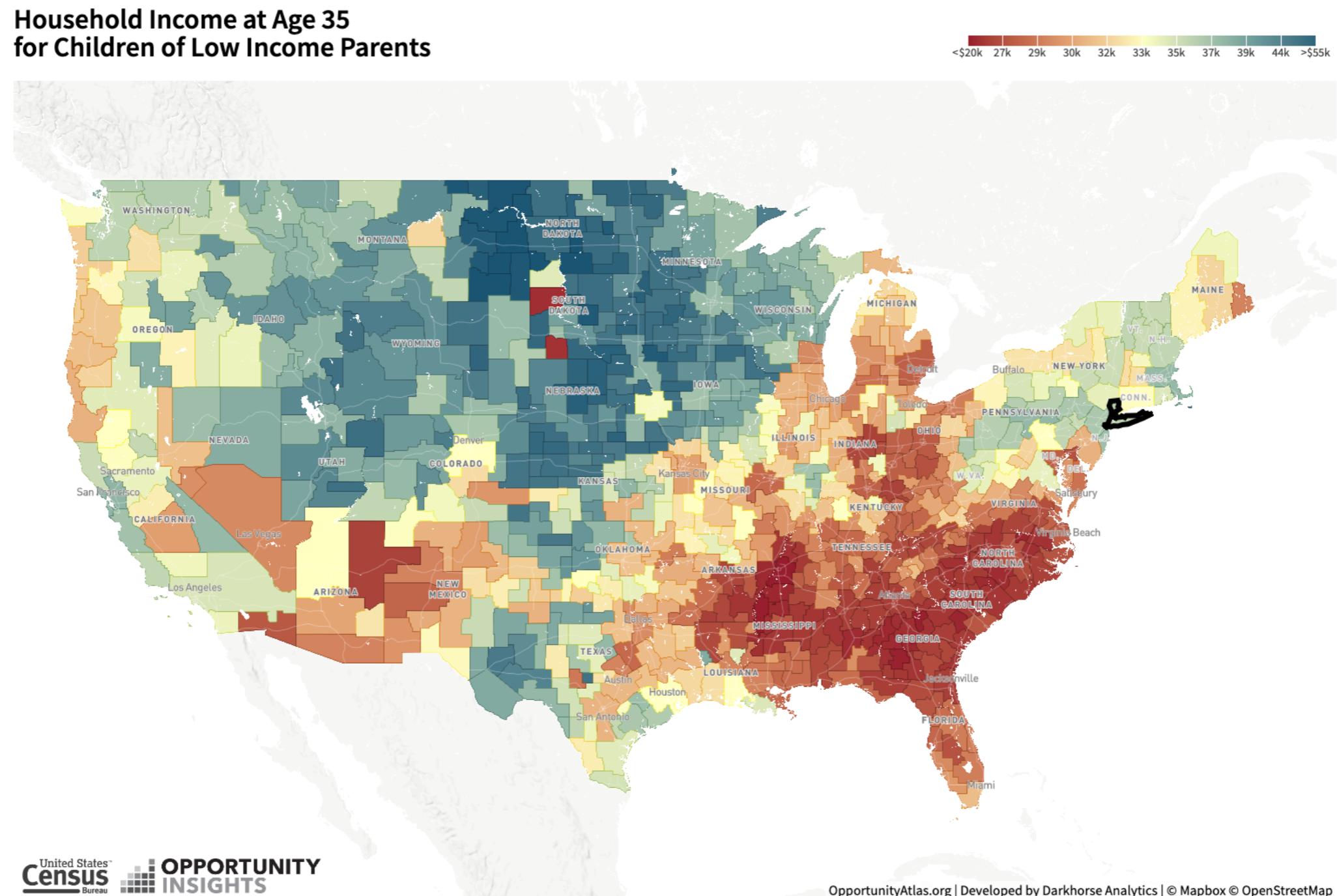
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Economic connectedness	1.00	-	-	-	-	-	-	-	-
(2) Language connectedness	0.10	1.00	-	-	-	-	-	-	-
(3) Age connectedness	-0.45	0.17	1.00	-	-	-	-	-	-
(4) Clustering	0.01	0.38	0.51	1.00	-	-	-	-	-
(5) Support ratio	-0.25	0.30	0.50	0.64	1.00	-	-	-	-
(6) Spectral homophily	-0.09	-0.37	-0.49	-0.61	-0.51	1.00	-	-	-
(7) Penn State index	0.31	0.08	-0.04	0.39	0.28	-0.25	1.00	-	-
(8) Civic organizations	0.27	0.16	0.05	0.37	0.23	-0.33	0.67	1.00	-
(9) Volunteering rate	0.46	0.28	-0.04	0.30	0.23	-0.35	0.44	0.46	1.00

This table reports county-level pairwise correlations of the primary social capital measures that we analyse, weighted by the number of children with below-median parental income in each county as calculated in the Opportunity Atlas<sup>72</sup> using Census data. Economic connectedness is twice the share of above-median-SES friends among below-median-SES people. Language connectedness is the share of friends who set their Facebook language to English among users who do not set their language to English, divided by the national share of users who set their language to English. Age connectedness is the share of friends who are aged 35–44 years among users who are aged 25–34 years, divided by the national share of users aged 35–44 years. Clustering is the share of an individual's friend pairs who are also friends with each other, averaged over all individuals in the county. Support ratio is the share of friendships between people in the county with at least one other mutual friend in the county. Spectral homophily is the second largest eigenvalue of the row-stochasticized network adjacency matrix, a measure of the extent to which the county-level friendship network is fragmented into separate groups. The Penn State index<sup>63</sup> is an index of participation in civic organizations and other measures of civic engagement. Civic organizations is the number of civic organizations with Facebook pages per 1,000 Facebook users in the county. Volunteering rate is the percentage of Facebook users in the county who are members of volunteering or activism groups. See Supplementary Table 1 for an expanded version of this correlation table that includes all social capital measures that we construct. Further details on all the social capital measures are provided in the Methods ('Variable definitions').

# 世代間所得上昇移動との関係

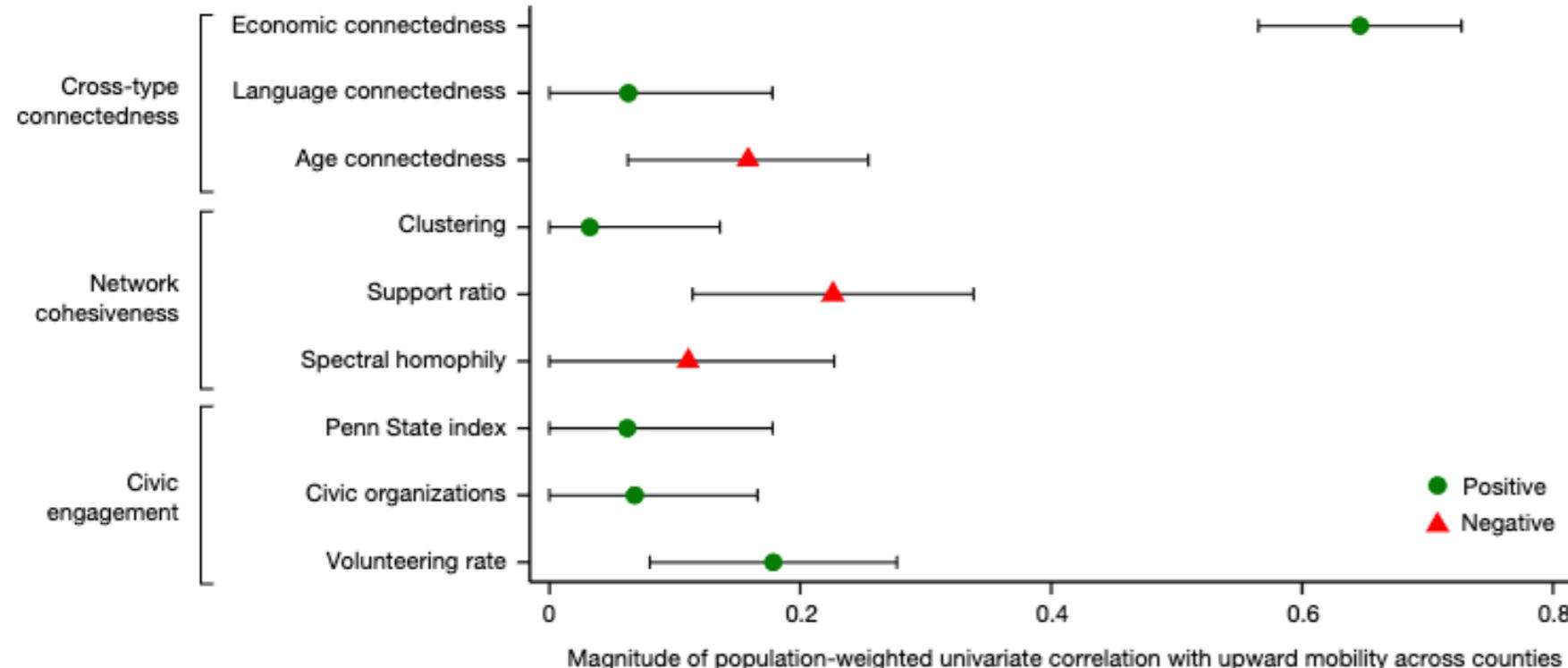
# County別の所得上昇移動との関連

従属変数：County別・1978–83年生の親所得下位25%出身者の35歳時の平均世帯所得を個人レベルの世帯所得の分布100分位のランクに変換したもの



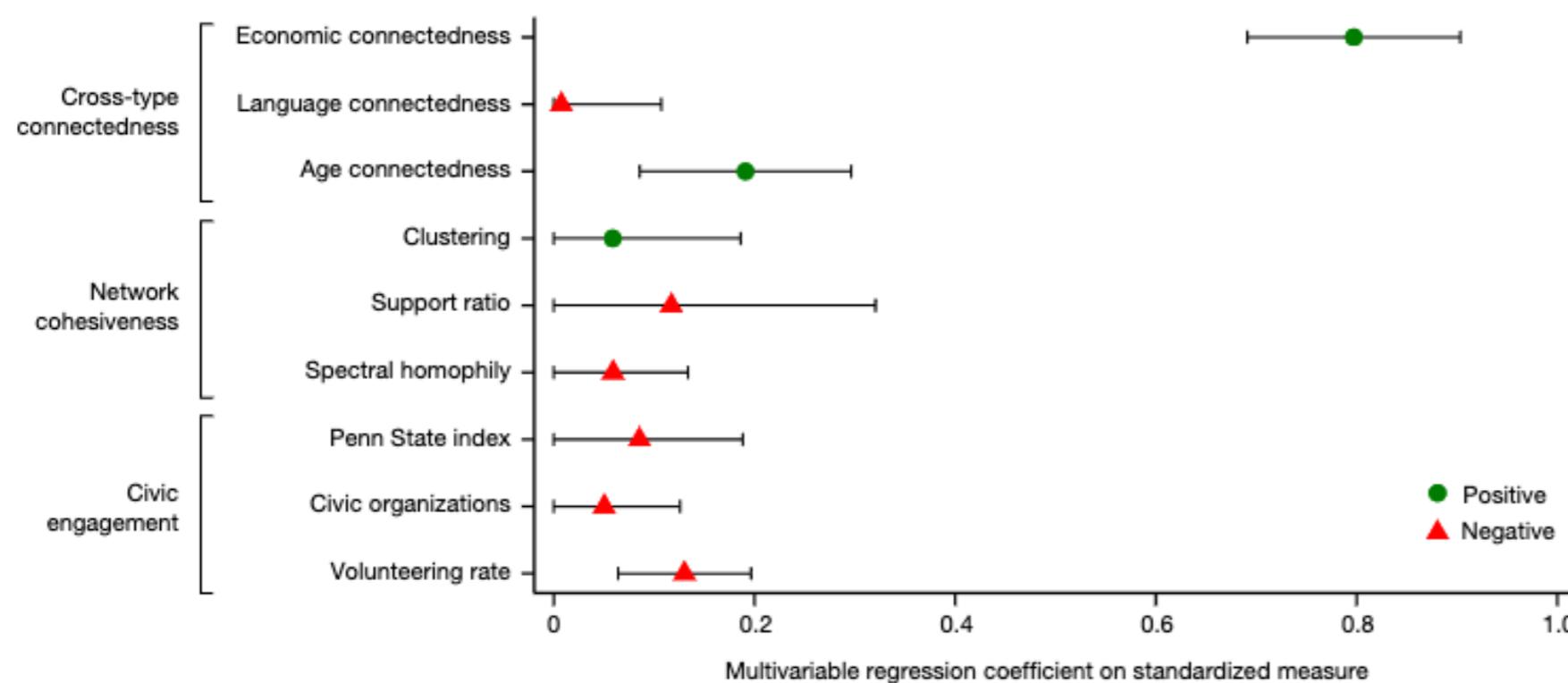
# 多くの指標のうち、経済的つながりが上昇移動率と正に相関

a



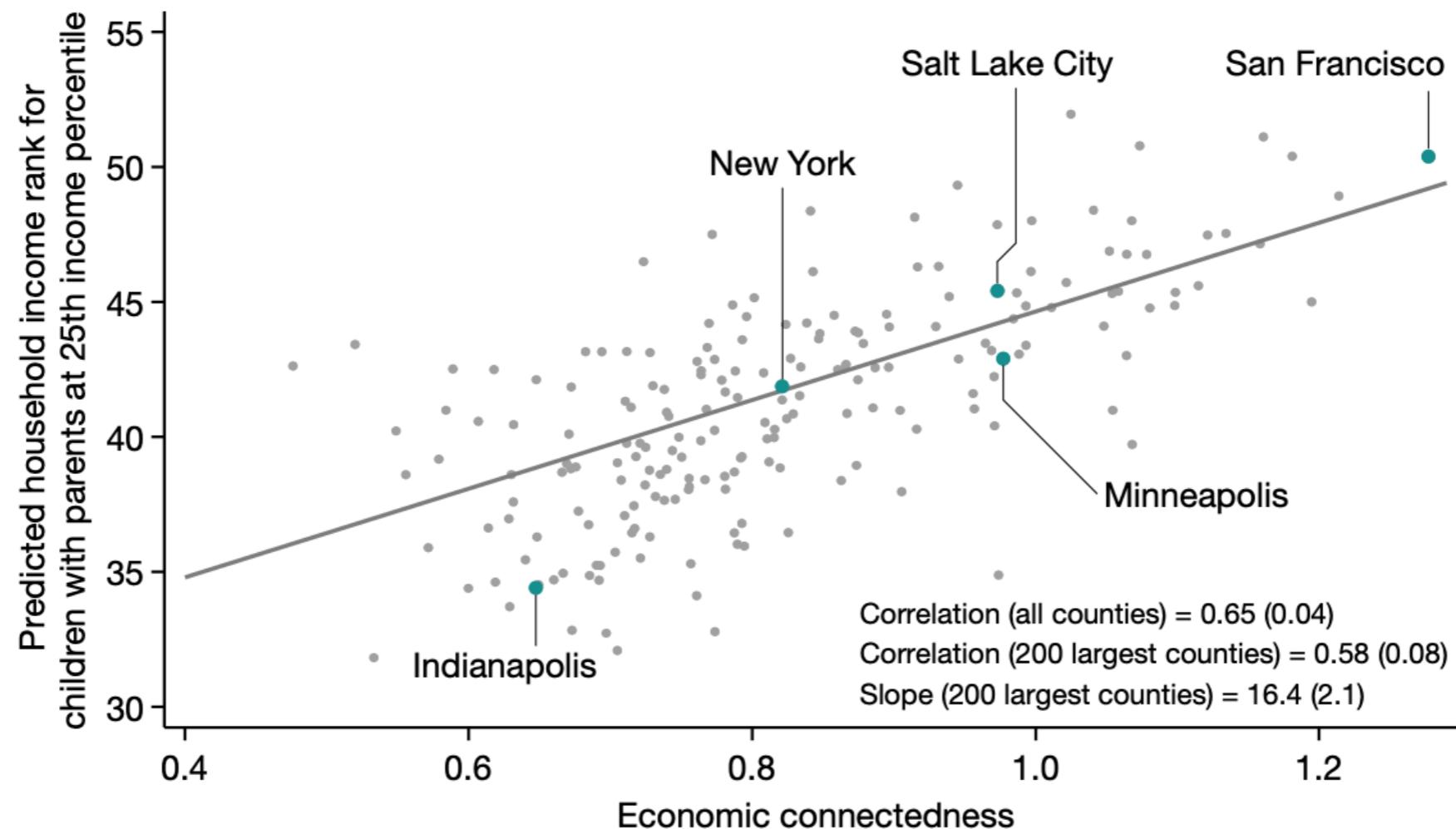
a. county-levelの指標と上昇移動率との2変量相関係数（標準化偏回帰係数）

b



b. county-levelの指標すべてを投入した回帰分析の標準化偏回帰係数

# 経済的つながりと所得移動の散布図（大規模地区）



なお、Countyごとにデータを分割してZIP-code単位で分析しても経済的つながりは依然として上昇移動と正の関連がある

# 経済的つながりと上昇移動の関連は因果関係か？

因果関係以外の3つのあり得る説明を検証

1. **逆因果Reverse causality**：上昇移動の多い地区では、低所得層は高所得層を選んでつながりやすいのではないか
2. **セレクション**：上昇移動の多い地区に住む者は、たとえ同じように低所得だったとしても上昇移動しやすい特徴を持つ
3. **他の似たような変数の効果**：経済的つながりの多い地区にはその他の望ましい特性が備わっており、経済的つながりそのものの効果ではない

# 1. 逆因果の検証

子ども時代のSES = 親のSESにもとづいて経済的つながり指標を作成し、子ども時代の経済的つながりの多寡と上昇移動に与える効果を測定することで、逆因果の問題に対処する

**対象者の親**：対象者のFacebookの登録情報や投稿などから特定（31%は特定できた）。

**対象者の子ども時代の友人**：対象者とFacebookでつながっていて、同じ高校出身で、かつ対象者の年齢±3歳の範囲の者。

**対象者の子ども時代の友人の親**：対象者の親と同様。

この指標を用いても依然として上昇移動率との相関は強い（0.44 / 0.62\*）

\* 2022年のInstagramデータを使った指標の場合。手続きは読み飛ばした

## 2. セレクションの検証

Countyの人種の割合を条件づけて分析（白人割合が高い地区、黒人割合が高い地区）しても、結果は同じ

Countyを移動した個人のデータをもとにして求めたCountyが上昇移動に与える効果（係数）を従属変数として分析 (Chetty & Hendren, 2018) → 正の相関あり

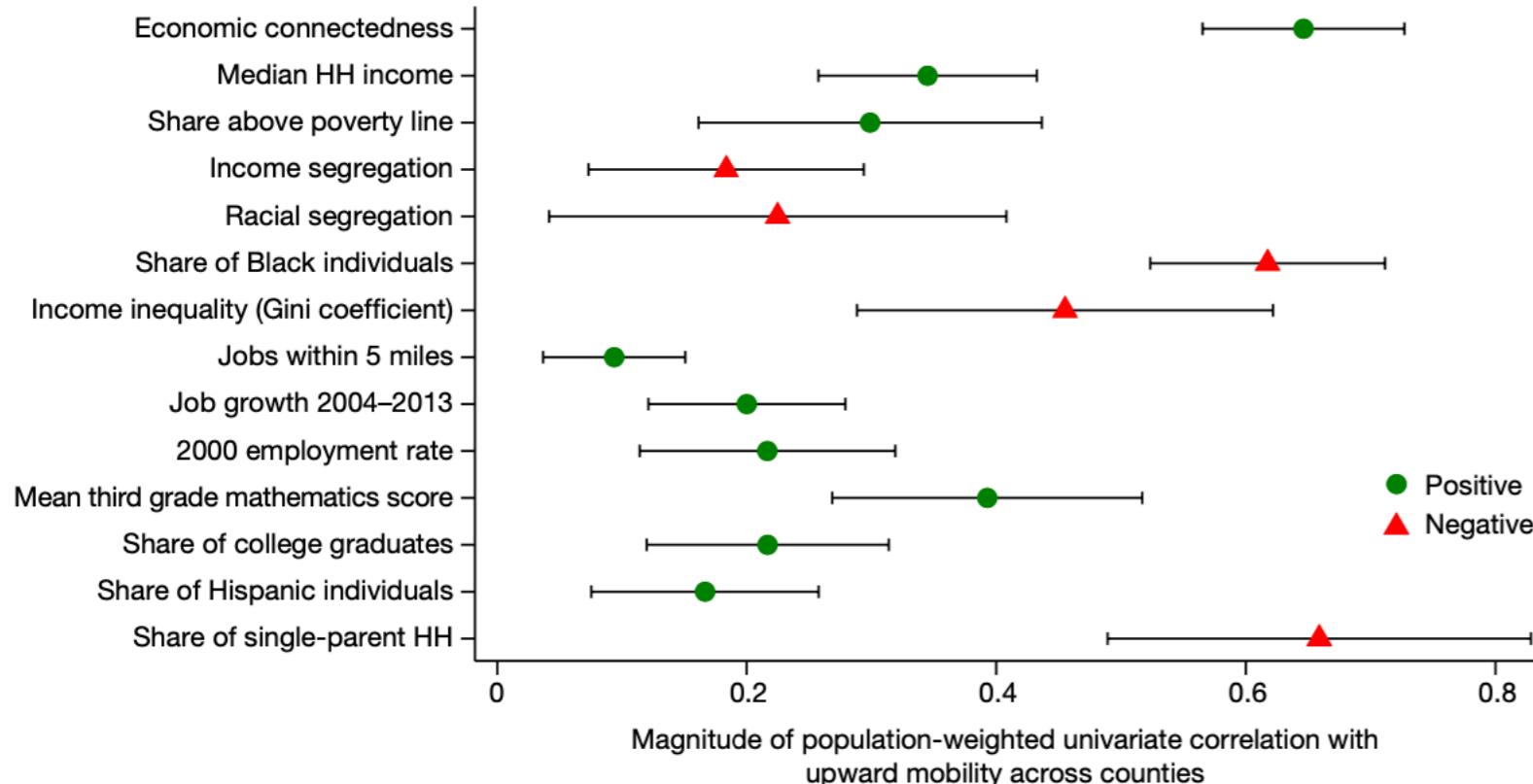
個人	0歳	1歳	2歳	…	17歳	18歳
1	A	A	B		B	B
2	A	A	A		A	B
3	B	C	C		A	A

$$Y_i = \beta_A X_{iA} + \beta_B X_{iB} + \beta_C X_{iC} + \dots$$

→ County Aに一年住むことが上昇移動に与える効果

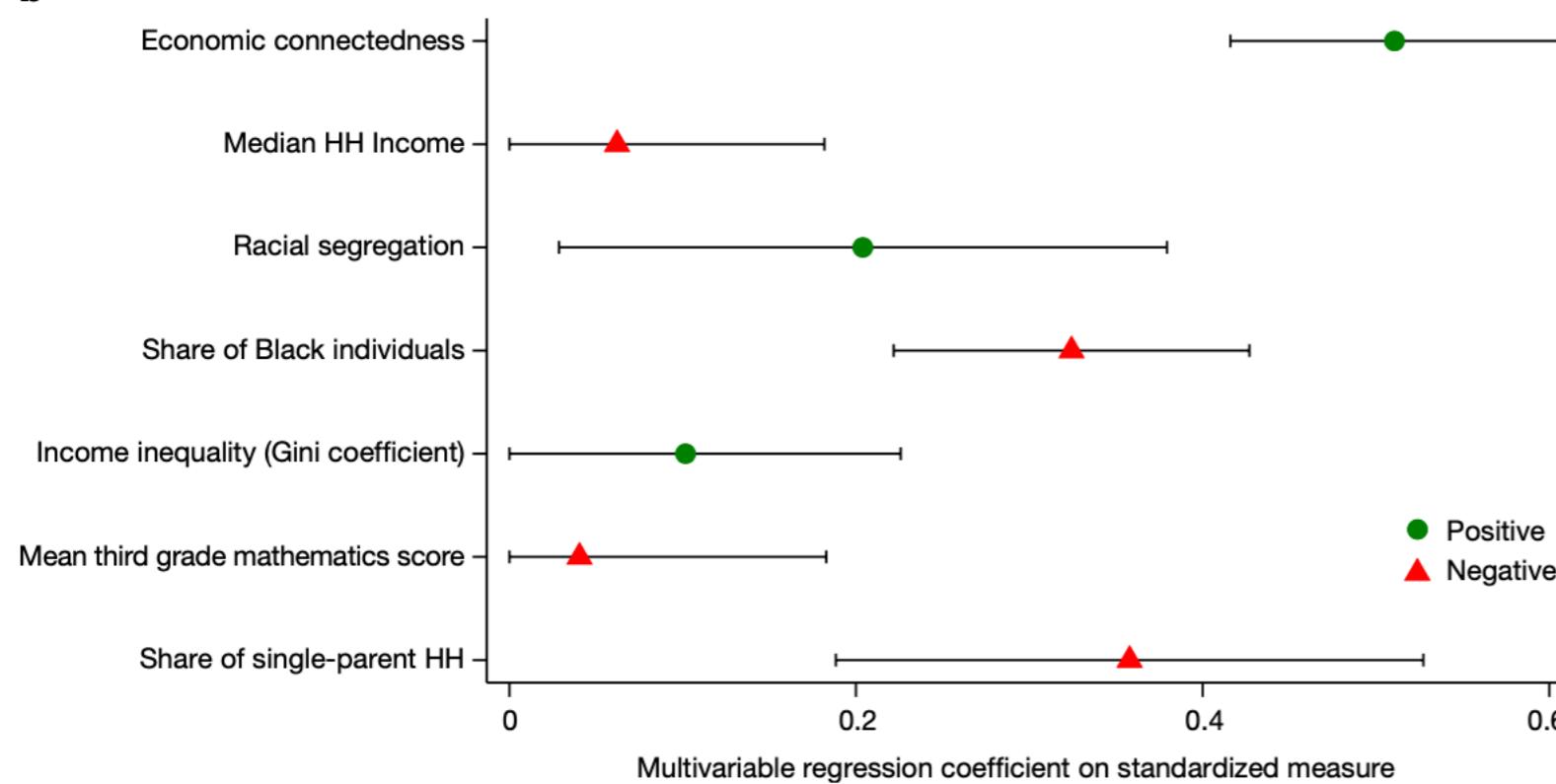
### 3. 他の似たような変数の効果の検証

a



a. county-levelの指標と上昇移動率との2変量相関係数（標準化偏回帰係数）

b



b. county-levelの指標のうち単変量相関の大きい7つの指標を投入した回帰分析の標準化偏回帰係数

# 経済的つながりによる媒介効果

**Table 2 | Associations between upward income mobility, EC and other neighbourhood characteristics**

EC versus median income and poverty rates								
Dependent variable	Upward income mobility							
	Counties				ZIP codes			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Median income	0.345*** (0.045)	-0.006 (0.069)	-	-	0.574*** (0.023)	0.209*** (0.029)	-	-
Poverty rate	- -	- (0.070)	-0.299*** (0.069)	0.142** (0.043)	- -	- (0.038)	-0.543*** (0.052)	-0.195*** (0.054)
Economic connectedness	- -	0.649*** (0.058)	- -	0.732*** (0.043)	- -	0.548*** (0.038)	- -	0.568*** (0.033)
Observations	2,984	2,984	2,984	2,984	24,165	24,165	24,165	24,165
R <sup>2</sup>	0.119	0.418	0.089	0.430	0.330	0.496	0.295	0.496
EC versus segregation and inequality								
Dependent variable	Upward income mobility							
	(1)	(2)	(3)	(4)	(5)	(6)		
Income segregation	-0.173*** (0.053)	-0.071 (0.054)	-	-	-	-	-	-
Racial segregation	- -	- (0.088)	-0.212** (0.086)	-0.027 (0.054)	-	-	-	-
Income inequality (Gini coefficient)	- -	- (0.084)	- (0.084)	- (0.091)	-0.449*** (0.084)	- (0.091)	-0.103 (0.091)	
Economic connectedness	- -	0.601*** (0.044)	- -	0.604*** (0.054)	- (0.054)	- (0.063)	0.577*** (0.063)	
Observations	1,820	1,820	1,821	1,821	2,741	2,741		
R <sup>2</sup>	0.034	0.413	0.051	0.408	0.207	0.424		
EC versus share of Black residents								
Dependent variable	Upward income mobility for Black individuals				Upward income mobility for white individuals			
	Counties		ZIP codes		Counties		ZIP codes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of Black individuals	-0.158** (0.068)	0.078 (0.076)	-0.204*** (0.057)	-0.014 (0.071)	-0.128** (0.057)	0.151** (0.067)	-0.250*** (0.018)	0.035* (0.018)
Economic connectedness	- -	0.502*** (0.095)	- -	0.468*** (0.083)	- -	0.582*** (0.051)	- -	0.631*** (0.027)
Observations	1,885	1,885	11,147	11,147	2,982	2,982	24,020	24,020
R <sup>2</sup>	0.025	0.222	0.042	0.224	0.016	0.277	0.063	0.380

# 今後の研究課題

1. 今回作成したさまざまなソーシャル・キャピタル指標が所得移動以外のアウトカムに与える効果の分析
2. アメリカ以外でのネットワークデータを活用したソーシャル・キャピタル指標の作成
3. 経済的つながりを増やすような介入を行えば世代間所得移動は増加するのかの検証。そのためには経済的つながりがどのようなメカニズムで生まれるのかを理解する必要がある。次の論文\*でこれを分析する。

\*Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt. 2022. "Social Capital II: Determinants of Economic Connectedness." *Nature* 1–13.

# 感想

- データがすごすぎて同じことは絶対にできないが、それでも完璧ではないデータから工夫によってある程度良質な変数を作っている（SESの作成など）、効果的な図の提示の仕方など、学べるところも多い。
- ソーシャル・キャピタルは個人レベルの資源だと位置づけている研究者もいるが（Colemanなど）、この論文では一貫して地域レベルの変数として位置づけているのが特徴的（「SESの高い人とのつながりを持つ人ほど上昇移動しやすい」という主張ではない）。
- 著者らは経済学者だが、社会学の論文も多く引用している。この研究に限らず、真に革新的な研究は学問領域を飛び越えて重要な問題に挑戦しているように思う。
- Natureは当分出てこなさだとしても、最近は社会学のトップ中のトップの研究者はPNASなど元々自然科学中心のジャーナルに論文を掲載することが増えている印象。このあたりのジャーナルはチェックしておくとよいかもしれない。