Supplementary Information

Social Capital Around the World: Measurement and Correlates

Drew Johnston Theresa Kuchler Ayush Kumar
Johannes Stroebel Manas Kulkarni Mike Bailey

Contents

1 Data for Outcome Analysis

1.1 Country-level measures of intergenerational income mobility

Intergenerational mobility refers to the extent to which measures of a person's socioeconomic status (such as income or education) are independent of those of their parents. For example, a common way to measure intergenerational income mobility is through the *intergenerational income elasticity (IGE)*, which quantifies the relationship between parents' and children's incomes via the following regression:

$$y_{i,t} = \alpha + \beta y_{i,t-1}^{parents} + \varepsilon_{i,t}.$$

Here, $y_{i,t}$ is the log income of individual i in generation t, and $y_{i,t-1}^{parents}$ is the log of individual i's parents' income in generation $t-1$. The coefficient β captures the IGE, and intergenerational income mobility is thus given by $1-\beta$. If $\beta = 0$, parents' income does not predict children's income and there is high mobility; if $\beta = 1$, children's income rises proportionally with parents' income and there is low mobility.

The World Bank's Global Database on Intergenerational Mobility (GDIM) provides estimates of intergenerational income elasticity for 87 countries, covering approximately 84 percent of the world's population (Munoz and van der Weide 2025). The database reports IGE estimates for father-son pairs, constructed from at least 156 household surveys using the Two-Sample Two-Stage Least Squares (TSTSLS) method. This approach enables estimation of IGE in the absence of long-running panel data by predicting parental income from retrospective information on education and, where available, occupation. As a result, the approach broadens global coverage compared to what would otherwise be possible.

1.2 Country-level measures of intergenerational educational mobility

The GDIM also provides estimates of intergenerational educational mobility for 153 economies, covering about 97 percent of the world's population (Van der Weide et al. 2024). The data provides cohort-based estimates for individuals born between 1940 and 1990, with cohorts defined by birth decade, and includes mobility estimates across various definitions of parental education (mother, father, average, or maximum) and child outcomes (sons, daughters, or all individuals). We use the estimates where parental education is represented by the maximum level attained among the two parents and child education is measured among all respondents. Specifically, we use the following variables from the GDIM:

- COR: Correlation coefficient between children's and parents' years of schooling.
- BETA: Beta coeffficient from regressing children's on parents' years of schooling.
- MU050: Expected child educational rank from a person born in the bottom half.
- BHQ4: Probability that a child from the bottom half reaches the top quartile.

In the main analysis, we use (1 − COR) as the measure of intergenerational educational mobility. We focus on the 1980s cohort, and only use estimates derived from retrospective surveys, that is, surveys where respondents report both their own and their parents' educational attainments. This approach ensures comparability across countries and avoids reliance on co-resident data.[1]

[1] In surveys where children are not asked about their parents' education level, parent and child education can only be matched for children still sharing a household with their parents, which is a selected sample.

To assess the robustness of our findings to these choices, Table SI-12 also presents results using various alternative measures of educational intergenerational mobility from the GDIM: MU050, BHQ4, and (1 − BETA).[2] We also explore estimates for earlier cohorts (1960s and 1970s). In addition, we conduct robustness checks that incorporate data from co-resident surveys, which are used in the GDIM when retrospective data are unavailable, and which thus allow for an expansion of the sample. Across these robustness checks, we find a consistently strong negative relationship between EC stratification and intergenerational educational mobility.

1.3 Subnational measures of upward income mobility

Our analysis of the within-country relationships between economic connectedness and upward income mobility draws on a variety of existing estimates of such mobility. The typical measure of upward mobility in the literature captures the average adult income rank for a child born to parents at the 25th income percentile in their birth cohort, but exact specifications vary. We summarize the data sources in Table SI-1, and discuss them briefly here.

1.3.1 Australia

Deutscher and Mazumder (2020) measure upward income mobility in Australia using Australian federal income tax returns from 1991 to 2015 covering 1.1 million children born between 1978 and 1982. Children are linked to parents using dependent status from tax returns. Measures are decomposed at the Australian Statistical Area 4 (SA4) geographic level (N = 87). Children are assigned to the associated SA4 region using the first geographic location associated with their primary parent. The primary income measure is individual total pretax income. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.2 Brazil

Britto et al. (2024) measure intergenerational mobility in Brazil at the level of the municipality (N = 5,570). The authors combine administrative records and household surveys covering 1.3 million children born between 1988 and 1990. The authors predict informal income using a machine learning model trained on household surveys. Capital income and other formal non-labor income is imputed using the same methodology. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.3 Canada

Connolly et al. (2019) provide estimates of intergenerational mobility for Canadian Census Divisions (N = 288) using income tax records to identify child-parent pairs. The analytic sample consists of children born between 1972 and 1985. Children's adult incomes were measured in 1991, 1996, or 2001. Children born to parents with less than 500 US dollars of income are dropped from the sample. The final sample consists of 557 thousand children with both parent and child income measured pre-tax but post-transfer. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

[2] When multiple parents report the same years of schooling, tie-breaking is necessary to assign ranks. To account for this ambiguity, Van der Weide et al. (2024) assign upper and lower bounds for rank-based mobility measures. Following their work, we use the midpoint of these bounds as our point estimate for intergenerational elasticity in education.

1.3.4 Denmark

Cholli et al. (2024) estimate upward income mobility for Danish parishes (N = 1,949) using 560 thousand parent-child links derived from administrative records in Denmark. Observations with missing data before age 17, low child or parent income, or children who are first- or second-generation immigrants are excluded. Measures are constructed using gross income (before taxes and transfers). The top 0.5% of the child and parent income distributions are winsorized. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.5 Ecuador

Del Pozo and Moreno (2024) combined social security records, census data, and labor force surveys to build measures of upward mobility for cantons in Ecuador (N = 221). Similar to Britto et al. (2024), machine learning techniques are used to estimate informal labor income. The parent-linked analytic sample contains 15 million children, representing 70% of the total registered population. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.6 England

The Behavioral Insights Team built measures of social mobility for local authorities (N = 325) by linking English longitudinal education outcomes (LEO) data with adult earnings records (Zaman et al. 2024). Parental income during childhood is proxied for by eligibility for free school meals at age 16. The children linked in the LEO dataset represent around 14% of individuals born between 1986 and 1992 from England's lowest-income families. Our analyses use the expected child income rank for children eligible for free school meals as the preferred measure of upward mobility.

1.3.7 France

Kenedi and Sirugue (2023) use administrative data on French departments (N = 96) to estimate intergenerational mobility for children born in the 1970s. The main data source is the permanent demographic sample (EDP) which combines several administrative records on individuals born in the first four days of October. This data is combined with tax records, civil register data, census data, and the French all-employee panel. This dataset does not include parental income directly, so the authors use a two-stage least-squares approach to predict parents' incomes from a subsample of parents with income records available. This process mirrors machine learning approaches used to predict informal income and formal capital income in developing nations. Child income is computed directly from administrative records pre-tax but post-transfer. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.8 Italy

Acciari et al. (2022) use 1.7 million parent-child pairs constructed from tax records across Italian NUTS3 regions (N = 110) to document intergenerational mobility across Italy for children born between 1979 and 1983. The main sample uses gross income (before taxes and transfers) for both parent and child income

measures. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.9 Netherlands

Lam et al. (2025) use linked administrative data covering the full Dutch population (4.6 million parent-child pairs) to construct measures of intergenerational mobility for Dutch Gemeenten (N = 342). Outcomes are observed from birth through age 35 and span health, education, labor market activity, social insurance, housing, and fertility. Parental characteristics include income, wealth, and education, derived from tax and register data. Child income is measured at age 35 using gross income (before taxes and transfers). Absolute upward mobility is measured as the expected adult income rank of children born to parents at the 25th percentile of the national income distribution.

1.3.10 New Zealand

Crichton and Jenkins (2026) use birth records and administrative tax data from New Zealand's Integrated Data Infrastructure to construct estimates of intergenerational income mobility for Territorial Authority Local Boards (N = 85) in New Zealand. The analysis sample includes around 230 thousand children, and location data is weighted by proportion of time spent in a given location by a child from age 10 to 17. Parent income is averaged across tax years the child turned 15-19, while child income is measured in the tax year the child turned 28, 29, or 30 depending on the birth cohort. Absolute mobility is defined as the expected income rank of children whose parents were at the 25th percentile of the national income distribution.

1.3.11 Spain

Soria Espin and Medina (2025) use administrative tax records from Spanish postal code areas (N = 3,930) for children born in the 1980s and 1990s to measure absolute upward mobility as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution. Parental households reporting negative or low income while holding significant real estate were removed to address misreporting concerns. The bottom 5% of the child income distribution was removed for similar reasons. Parental income is measured using total gross household income. Child income is measured using individual gross income. The main sample includes over 1.5 million parent-child pairs.

1.3.12 Sweden

Nybom and Stuhler (2025) use Swedish administrative registers covering the universe of residents aged 16-64 to measure intergenerational mobility across municipalities (N = 290). Children are linked to their biological parents using multigenerational registers, and the analysis focuses on cohorts born between 1981 and 1989. Families are assigned to municipalities based on the child's place of residence at age 16. Parent and child labor earnings come from employer-reported tax records covering 1968-2020. Educational attainment is measured using Statistics Sweden's classification of schooling categories, recoded into years of education. Absolute upward mobility is defined as the expected adult earnings rank of children whose fathers were at the 25th percentile of the national earnings distribution.

1.3.13 Switzerland

Chuard and Schmiedgen-Grassi (2020) combine individual-demographic information with official survey records and social security earnings records from Swiss mobility regions (N = 106). The main analytic sample contains native children born between 1967 and 1984. Child and parent incomes are measured pre-tax and post-transfer. The main sample covers just over 900 thousand individuals. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.3.14 United States

Chetty, Friedman, et al. (2026) construct The Opportunity Atlas using linked U.S. Census and Internal Revenue Service (IRS) tax returns, providing data on upward mobility for U.S. ZCTAs (N = 26,962). The main sample consists of all children born between 1978 and 1983. Parental and child income are measured as gross pre-tax income. Absolute upward mobility is measured as the expected adult income rank for children born to parents at the 25th percentile of the national income distribution.

1.4 Subnational measures of educational mobility

We also analyze subnational measures of educational mobility for many developing nations. We focus on measures of education based on an individual's relative educational status within their country.

1.4.1 Africa

Alesina et al. (2021) build measures of educational intergenerational mobility for various subnational regions (N = 2,846) across 27 countries in Africa. The sample consists of data from 69 national censuses covering 93 million individuals. After applying sample restrictions there are 25.8 million individuals linked to parents with educational records. Individuals are labeled literate if they completed primary education and illiterate if they did not. The measure of upward mobility used in our analyses is the probability of a child achieving literacy if they were born to illiterate parents. The regional mobility estimates analyzed in the paper are residualized on country fixed effects.

1.4.2 Germany

Dodin et al. (2024) use German Microcensus data on education and parent income to construct measures of educational intergenerational mobility for German local labor market regions (N = 258). Given the tracking system in German secondary education, the authors use attainment of the A-Level degree—necessary for university entry—as their measure of opportunity. Parent income is measured using the household's equivalized monthly net income (excluding children's income), converted into percentile ranks within the national income distribution. The regional dataset is built using census survey waves from 2011 to 2018 and consists of a sample of 231 thousand children. The measure of educational mobility used in our analyses is the share of children whose parents are in the first or second quintile of the national income distribution who obtain or are on track to obtain an A-level-equivalent degree. The regional mobility estimates analyzed in the paper are residualized on state fixed effects, to account for cross-state differences in educational systems and therefore the outcome variable.

1.4.3 India

Measures of upward educational mobility for subdistricts in India (N = 5,720) were developed by the co-author team of Sam Asher, Paul Novosad, and Charlie Rafkin and are based on data and methodology from Asher et al. (2024). The mobility estimates disaggregated by subdistrict were obtained through correspondence with the co-author team, but are not available to the public.

Asher et al. (2024) use parent-child linked data from the India Human Development Survey (IHDS), a survey covering 41,554 households. Educational mobility is estimated for birth cohorts between 1950 and 1989. The authors estimated bounds for children born to parents in the bottom-half of the parental educational distribution for the 1960-1969 birth cohorts in India. Education is coded into seven categories and rank-transformed for parent- and child-cohorts. The measure of mobility used in our analyses is the expected child education rank for a child born to parents in the bottom-half of the educational attainment distribution. Because education is a coarser measure of SES than income, the bottom-half mobility measure is more difficult to estimate accurately. Instead of estimating a conditional expectation function with a linear regression, the authors bound their bottom-half mobility measure with statistical theory. Our analyses use the midpoint of these bounds as our measure of upward mobility.

1.4.4 Latin America & the Caribbean

Muñoz Saavedra (2024) uses data from 91 censuses in 24 countries across Latin America and the Caribbean to estimate upward educational mobility in subnational regions (N = 6,684), measured as the likelihood of a child finishing primary school if they were born to parents who did not finish primary school. The regional mobility estimates analyzed in the paper are residualized on country fixed effects.

1.5 Socio-political outcomes

In the main analysis of the relationship between EC stratification and socio-political outcomes, we use data from the Joint European Values Study/World Values Survey, which provides comprehensive, cross-national measures of social attitudes, political preferences, and cultural norms (Haerpfer et al. 2020). To complement the main analysis, we also draw on additional survey data from the Gallup World Poll, the Ipsos Populism Report, and the International Social Survey Programme. We next describe these data sets, and the specific questions explored in our work.

1.5.1 Joint European Values Study/World Values Survey (EVS/WVS)

The Joint EVS/WVS 2017-2022 dataset is a comprehensive, cross-national social survey resource, offering harmonized data from 92 countries and territories gathered between 2017 and 2022 (EVS/WVS 2022). This dataset unifies the European Values Study (EVS) and the World Values Survey (WVS). The joint data incorporates 156,658 respondents (59,438 from EVS and 97,220 from WVS) from 92 countries and territories (36 from EVS and 66 from WVS, with 10 countries having participated in both). The data includes responses to 231 harmonized survey questions. Respondents were generally interviewed face-to-face, though mixed methods (such as online or phone interviews) were used in certain contexts.

Table SI-2 describes the questions from the survey that we explore in our work.[3] For each variable, the table includes the title by which the question is referred to in our figures, the variable code in the

[3] All outcome variables are sourced from the Joint EVS/WVS 2017-2022 dataset, with the exception of the question labeled *Importance of Luck for Success* (see Table SI-2). This question was only included in the World Values Survey Wave 7.

joint dataset, the exact wording of the survey question, the response scale, and whether our analysis reversed the scale for interpretability. Because the Joint EVS/WVS dataset harmonizes variables from both the EVS and WVS, the wording of the survey questions is largely consistent across the two surveys; consequently, only the WVS wording is displayed in the table. In some cases, the original survey coding assigned lower values to responses representing higher levels of a given attitude; these scales were flipped so that higher values consistently indicate a more positive or higher-level response.

To generate country-level measures, the mean of individual responses to each question within each country was calculated. In addition, to explore potential heterogeneity by economic status, we used respondents' self-reported income deciles to calculate separate averages for the top- and bottom-halves of the income distribution for each question.

1.5.2 Gallup World Poll

The Gallup World Poll covers over 140 countries representing more than 95% of the world's population (Gallup 2024). The poll surveys at least 1,000 adults per country (with larger samples in certain large countries), employing nationally representative sampling either face-to-face or by telephone. The target population is all civilians aged 15 and older, with questionnaire translations into all major languages spoken in each country. Table SI-3 summarizes the questions used from the Gallup survey. For each variable, the table reports the title by which it is referenced in the figures, the year in which the question was asked in the survey, the exact question wording from the survey, and the response scale.

1.5.3 Ipsos Populism Report

The Ipsos Populism Report 2025 is a global survey that explores attitudes toward populism, societal trust, leadership, economic discontent, and nativism (Young 2025). The survey covered 31 countries and involved a sample of 23,228 adults aged 16–74. It was fielded between February 21st and March 7th, 2025. Sample sizes were approximately 1,000 per country. In countries such as Australia, Canada, France, Germany, Great Britain, Italy, Japan, Spain, and the United States, the samples are broadly representative of the adult population under age 75. In other countries, including Brazil, Chile, Indonesia, Malaysia, South Africa, Thailand, and Türkiye, the samples tend to be more urban, educated, and affluent than the population at large. To enhance comparability, all country samples are weighted by Ipsos to align with the demographic profile of the adult population based on the most recent census data.

Table SI-4 summarizes the questions used from the Ipsos survey. For each variable, the table reports the title by which it is referenced in the figures, along with the exact question wording. All questions were asked on a three-point scale ("agree," "neither agree nor disagree," and "disagree"). The country-level measure used in this analysis is the share of respondents selecting "agree."

1.5.4 International Social Survey Programme

The International Social Survey Programme (ISSP) conducts coordinated surveys on social science topics, enabling cross-country and longitudinal comparisons of attitudes, values, and behaviors (ISSP Research Group 2019; ISSP Research Group 2022). The outcomes used in our analyses were from two modules of the survey: the ISSP 2017 "Social Networks and Social Resources" module (ZA6980) and the ISSP 2019 "Social Inequality V" module (ZA7600). The 2017 module collected data from 32 countries, with national samples of about 1,000-1,700 adults. The 2019 module collected data from 34 countries,

with national samples of similar sizes. To generate country-level measures, the mean of individual responses to each question in each country was calculated.

Table SI-5 summarizes the questions used from the ISSP surveys. For each variable, the table reports the title by which it is referenced in the figures, the variable code in the survey dataset, the exact survey question wording, the response scale, and whether the scale was reversed for interpretability.

1.6 Inequality

We use the Gini index as our preferred measure of inequality. The Gini index captures the degree of inequality within a distribution, typically income or consumption. The index ranges from 0, which captures perfect equality (all individuals have the same income or consumption) to 1, which captures perfect inequality (one individual holds all income or consumption). It is derived from the Lorenz curve, which plots cumulative income shares against population percentiles, with the Gini index corresponding to the normalized area between the Lorenz curve and the line of perfect equality. Since no single source provides comprehensive Gini estimates for all countries, we compile data from multiple sources, documented below in the order of preference.

To construct decade-level measures of income inequality, we harmonize Gini coefficients from three major sources—the World Bank, the Standardized World Income Inequality Database (SWIID), and the All the Ginis dataset—and apply a consistent priority ordering when multiple estimates are available. For each country and decade, we compute the mean Gini from each source, then select the first non-missing value following a fixed sequence. For the 1990s, values are drawn in order from the World Bank, SWIID, and All the Ginis; if all are missing, the estimate is used from the World Bank in the 2000s. The same hierarchy applies to earlier decades: the 1980s pull first from the World Bank and SWIID, then All the Ginis, or otherwise use the 1990s value, while the 1970s use the World Bank or SWIID and fall back to the 1980s. Contemporary inequality measures follow a similar logic but focus on the most recent available data: we select the best Gini in order from the World Bank (2020s, then 2010s), SWIID (2020s, then 2010s), and finally the latest available SWIID observation. All Gini values are on a 0–1 scale.

In columns 1-3 of Panel A of Table 1 and Supplementary Table SI-10, inequality is measured using Gini coefficients lagged by 20 years from the present day to reflect the level of inequality during respondents' childhoods. In columns 4-6 of Panel A of Table 1 and Supplementary Table SI-10, analyzing the educational mobility of the 1980s cohort, we use Gini values from the 1990s, corresponding to a 10-year lag relative to birth. In Panels B and C of Table 1 and Supplementary Table SI-10, as well as in Supplementary Tables SI-11 and SI-13, which examine present-day sociopolitical outcomes, we use the most recently available Gini estimates.

1.6.1 World Bank Gini index

The World Bank's Poverty and Inequality Platform compiles nationally representative household survey data from government statistical agencies and, for high-income countries, from the Luxembourg Income Study database (World Bank 2025).

The World Bank calculates the Gini index based on either income or consumption data, depending on which is most consistently available for each country; whenever possible, consumption is preferred as it provides a more stable welfare measure, especially in developing economies. The World Bank

harmonizes the data to improve cross-country comparability. For high-income economies, they apply estimation methods aligned with those used for developing countries.

The World Bank Gini data span the period from the 1960s through the 2020s. For each country, we calculate an average Gini coefficient by decade to provide a smoothed measure of inequality.

1.6.2 Standardized World Income Inequality Database

In cases where Gini figures from the World Bank are missing, we use data from the Standardized World Income Inequality Database (SWIID), which aims to provide internationally comparable measures of income inequality and redistribution (Solt 2020).

The SWIID harmonizes data from the OECD Income Distribution Database; the Socio-Economic Database for Latin America and the Caribbean generated by CEDLAS and the World Bank; Eurostat; the World Bank's PovcalNet; the UN Economic Commission for Latin America and the Caribbean; national statistical offices around the world; and academic studies, using the Luxembourg Income Study as a benchmark. The SWIID provides comparable Gini indices of disposable and market income inequality for 199 countries from 1960 onward, along with measures of absolute and relative redistribution.

In our analyses, we use the SWIID's estimated Gini index for household disposable income, which captures inequality in incomes after taxes and government transfers.

1.6.3 All the Ginis

In the few cases where Gini estimates are missing from both the World Bank and SWIID data, we use data from the All the Ginis database (Milanovic 2019).

1.7 Varieties of Democracy (V-Dem)

The Varieties of Democracy (V-Dem) dataset measures democracy and governance worldwide (Coppedge et al. 2025; Maerz et al. 2025). It covers over 200 countries from 1789 to the present, and contains indicators and indices on electoral, liberal, participatory, deliberative, and egalitarian principles of democracy. V-Dem data are compiled annually by more than 4,000 expert country coders. Each variable is typically coded by five or more independent country experts for every country-year, with statistical models used to resolve coder inconsistencies and account for potential biases.

We use certain country-level indices from the V-Dem as controls in our analyses of the relationship between EC stratification and socio-political outcomes. Table SI-6 summarizes the indices used from the V-Dem dataset. For each index, the table reports the corresponding variable code in the dataset and a short description of what the index measures. The indices are derived from detailed descriptions that guide the expert coders in rating each country-year. Importantly, the indices themselves are aggregates, built from a range of more specific measures within the V-Dem dataset. For each country we compute the average of its V-Dem scores over the years 2018–2024, and use these averages in our regressions. The overall Democracy Index employed in our analysis is constructed as the simple arithmetic mean of V-Dem's five core democracy indices.

1.8 Population Measures

For some of our subnational analyses, we weight our measures by population. Unfortunately, high-quality population measures from administrative sources are only available for some high-income coun-

tries. For our analyses, we construct population measures by aggregating gridded 1km-by-1km cells from LandScan (Lebakula et al. 2024) to every geographic aggregation we need. Cells that cross administrative boundaries are split based on the proportion of land-area which falls within a given boundary. The resulting population estimates align closely with administrative data where available (see Supplementary Fig. SI-6).

Table SI-1: Sources of Subnational Intergenerational Mobility Estimates

Geography	Granularity	Regions	Citation
Income Mobility			
Australia	SA4 Regions	87	Deutscher and Mazumder (2020)
Brazil	Municipalities	5,570	Britto et al. (2024)
Canada	Census Divisions	288	Connolly et al. (2019)
Denmark	Parishes	1,949	Cholli et al. (2024)
Ecuador	Cantons	221	Del Pozo and Moreno (2024)
England	Local Authorities	325	Zaman et al. (2024)
France	Departments	96	Kenedi and Sirugue (2023)
Italy	NUTS3 Regions	110	Acciari et al. (2022)
Netherlands	Gemeenten	342	Lam et al. (2025)
New Zealand	Territorial Authority Local Boards	85	Crichton and Jenkins (2026)
Spain	Postal Codes	3,930	Soria Espin and Medina (2025)
Sweden	Municipalities	290	Nybom and Stuhler (2025)
Switzerland	Mobility Regions	106	Chuard and Schmiedgen-Grassi (2020)
United States	ZCTAs	26,962	Chetty, Friedman, et al. (2026)
Education Mobility			
Africa	Varied (27 Countries)	2,846	Alesina et al. (2021)
Germany	Local Labor Markets	258	Dodin et al. (2024)
India	Subdistricts	5,720	Asher et al. (2024)
Latin America	Varied (20 Countries)	6,684	Muñoz Saavedra (2024)

Note: This table lists the data sources for measures of intergenerational mobility at the subnational level. For each dataset, the table reports the granularity, the number of units available, and the original source.

Table SI-2: Joint European Values Study/World Values Survey Outcomes

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
Trust in Others				
Trust Most People	a165	Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?	2	✓
Trust Your Neighbor-hood	g007_18_b	I would like to ask you how much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all? People in your neighborhood.	4	✓
Trust Strangers	g007_34_b	I would like to ask you how much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all? People you meet for the first time.	4	✓
Trust Other Nationalities	g007_36_b	I would like to ask you how much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all? People of another nationality.	4	✓
Trust in the Electoral Process				
Votes are Counted Fairly	e265_01	In your view, how often do the following things occur in this country's elections? Votes are counted fairly.	4	✓
Election Officials are Fair	e265_06	In your view, how often do the following things occur in this country's elections? Election officials are fair.	4	✓

Continued on next page

Table SI-2: Joint European Values Study/World Values Survey Outcomes (Continued)

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
Opposition is Allowed to Run	e265_02	In your view, how often do the following things occur in this country's elections? Opposition candidates are prevented from running.	4	×
Voters are not Bribed	e265_04	In your view, how often do the following things occur in this country's elections? Voters are bribed.	4	×
Confidence in Institutions				
Confidence in Courts	e069_17	Please look at this card and tell me, for each item listed, how much confidence you have in them, is it a great deal, quite a lot, not very much or none at all? The justice system.	4	✓
Confidence in the Police	e069_06	Please look at this card and tell me, for each item listed, how much confidence you have in them, is it a great deal, quite a lot, not very much or none at all? The police.	4	✓
Confidence in Parliament	e069_07	Please look at this card and tell me, for each item listed, how much confidence you have in them, is it a great deal, quite a lot, not very much or none at all? Parliament.	4	✓
Confidence in Civil Service	e069_08	Please look at this card and tell me, for each item listed, how much confidence you have in them, is it a great deal, quite a lot, not very much or none at all? Civil service.	4	✓
Civic Misbehavior				
Welfare Fraud is Justifiable	f114a	Please tell me for each of the following whether you think it can always be justified, never be justified, or something in between, using this card. Claiming state benefits which you are not entitled to.	10	×

Continued on next page

Table SI-2: Joint European Values Study/World Values Survey Outcomes (Continued)

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
Cheating on Taxes is Justifiable	f116	Please tell me for each of the following whether you think it can always be justified, never be justified, or something in between, using this card. Cheating on tax if you have the chance.	10	×
Fare Evasion is Justifiable	f115	Please tell me for each of the following whether you think it can always be justified, never be justified, or something in between, using this card. Avoiding a fare on public transport.	10	×
Democracy and Populism				
Importance of Democracy	e235	How important is it for you to live in a country that is governed democratically? On this scale where 1 means it is "not at all important" and 10 means "absolutely important" what position would you choose?	10	×
Support for Strong Leadership	e114	I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having a strong leader who does not have to bother with parliament and elections.	4	✓
Prioritize Natives for Jobs	c002_01	For each of the following statements I read out, can you tell me how strongly you agree or disagree with each. Do you strongly agree, agree, disagree, or strongly disagree? When jobs are scarce, employers should give priority to [NATIONALITY] people over immigrants.	5	✓
Income Redistribution				
Incomes Should be Equalized	e035	On this card you see a number of opposite views on various issues. How would you place your views on this scale? 1 - Incomes should be made more equal; 10 - We need larger income differences as incentives.	10	✓

Continued on next page

Table SI-2: Joint European Values Study/World Values Survey Outcomes (Continued)

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
Happiness				
Feeling of Happiness	a008	Taking all things together, would you say you are: 1 - Very happy; 2 - Quite happy; 3 - Not very happy; 4 - Not at all happy.	4	✓
Overall Life Satisfaction	a170	All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole?	10	×
Feeling of Control over Life	a173	Some people feel they have completely free choice and control over their lives, and other people feel that what they do has no real effect on what happens to them. Please use the scale to indicate how much freedom of choice and control you feel you have over the way your life turns out? 1 - None at all; 10 - A great deal.	10	×
World Values Survey only				
Importance of Luck for Success	q110	How would you place your views on this scale? 1 - In the long run, hard work usually brings a better life; 10 - Hard work doesn't generally bring success—it's more a matter of luck and connections	10	×

Note: This table summarizes the Joint European Values Study/World Values Survey questions analyzed in this study. For each variable, it reports the figure label, dataset code, WVS question wording, response scale, and whether the scale was reversed for interpretability. Because the Joint EVS/WVS dataset harmonizes the two surveys, wording is largely consistent and only the WVS version is shown. Scales were reversed where necessary so that higher values uniformly indicate more positive or higher-level responses.

Table SI-3: Gallup World Poll Outcomes

Outcome	Survey Year	Survey Question	Points in Response
Trust in neighbours	2020	How much do you trust each of the following? Do you trust them a lot, some, not much, or not at all? If you don't know, please just say so. How about ... the people in your neighborhood.	4
Neighbours care about your wellbeing	2021	How much do you think most of your neighbors care about you and your wellbeing? A lot, somewhat or not at all?	3
Confidence in judicial system	2024	In this country, do you have confidence in each of the following, or not? How about the judicial system and courts?	2
Confidence in local police	2024	In the city or area where you live, do you have confidence in the local police force, or not?	2
Confidence in military	2024	In this country, do you have confidence in each of the following, or not? How about the military?	2
Trust in journalists	2020	How much do you trust each of the following? Do you trust them a lot, some, not much, or not at all? If you don't know, please just say so. How about ... journalists in this country.	4
Perceived freedom of media	2024	Do the media in this country have a lot of freedom, or not?	2
Confidence in honesty of elections	2024	In this country, do you have confidence in each of the following, or not? How about honesty of elections?	2
Corruption in government is widespread	2024	Is corruption widespread throughout the government in this country, or not?	2

Continued on next page

Table SI-3: Gallup World Poll Outcomes (Continued)

Outcome	Survey Year	Survey Question	Points in Response
Trust in government	2020	How much do you trust each of the following? Do you trust them a lot, some, not much, or not at all? If you don't know, please just say so. How about ... the national government in this country.	4
Government cares about your wellbeing	2021	How much do you think the government of [COUNTRY] cares about you and your wellbeing? A lot, somewhat or not at all?	3
Satisfied with opportunities to make friends	2024	In the city or area where you live, are you satisfied or dissatisfied with ... the opportunities to meet people and make friends.	2
Satisfied with freedom to make life decisions	2024	In this country, are you satisfied or dissatisfied with your freedom to choose what you do with your life?	2

Note: This table summarizes the Gallup survey questions used in the analysis. For each variable, it reports the label by which the question is referred to in the figures as well as the exact question wording from the Gallup World Poll survey. The table also reports the survey year from which the question was taken and the number of points in the scale of responses for the question.

Table SI-4: Ipsos Populism Report Outcomes

Outcome	Survey Question
Need strong leader against rich & powerful	[COUNTRY] needs a strong leader to take the country back from the rich and powerful.
Need strong leader willing to break rules	To fix [COUNTRY], we need a strong leader willing to break the rules.
Politicians don't care about people like me	Traditional parties and politicians don't care about people like me.
Elites don't care about hardworking people	The political and economic elite don't care about hardworking people.
Main divide in society is citizens vs elite	The main divide in our society is between ordinary citizens and the political and economic elite.
Experts don't understand people like me	Experts in this country don't understand the lives of people like me.
Media prioritizes profit over truth	The mainstream media is more interested in making money than telling the truth.
Economy is rigged to advantage the rich	[COUNTRY] economy is rigged to advantage the rich and powerful.
Society is broken	Society is broken.
Country is in decline	[COUNTRY] is in decline.

Note: This table summarizes the Ipsos survey questions used in the analysis. For each variable, it reports the label by which the question is referred to in the figures as well as the exact question wording from the Ipsos survey. All questions use a three-point scale, and country-level measures reflect the share of respondents who "agree."

Table SI-5: International Social Survey Programme Outcomes

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
ISSP 2017 - Social Networks and Social Resources				
People try to take advantage of you	v34	How often do you think that people would try to take advantage of you if they got the chance, and how often would they try to be fair?	4	✓
Richer people should help less rich friends	v40	To what extent do you agree or disagree with the following statements? People who are better off should help friends who are less well off.	5	✓
Care for self & family before helping others	v39	To what extent do you agree or disagree with the following statements? You should take care of yourself and your family first, before helping other people.	5	✓
Friendship should be of use to self	v61	To what extent do you agree or disagree with the following statements? It is all right to develop friendships with people just because they can be of use to you.	5	✓
Favors should be repaid	v62	To what extent do you agree or disagree with the following statements? When another person does a favour for you, you should feel obligated to pay that person back.	5	✓
Differences in living standards should be small	v12	To what extent do you agree or disagree with the following statements? For a society to be fair, differences in people's standard of living should be small.	5	✓
Frequency of feeling left out	v33	How often in the past 4 weeks have you felt that ... you are left out?	5	×
Frequency of feeling isolated from others	v32	How often in the past 4 weeks have you felt that ... you are isolated from others?	5	×

Continued on next page

Table SI-5: International Social Survey Programme Outcomes (Continued)

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
Frequency of feeling lack of companionship	v31	How often in the past 4 weeks have you felt that ... you lack companionship?	5	×
ISSP 2019 - Social Inequality V				
People can be trusted	v56	Generally speaking, would you say that people can be trusted or that you can't be too careful in dealing with people?	4	✓
It is fair that the rich can buy better education	v31	Is it just or unjust - right or wrong - that people with higher incomes can buy better education for their children than people with lower incomes?	5	✓
It is fair that the rich can buy better healthcare	v30	Is it just or unjust - right or wrong - that people with higher incomes can buy better health care than people with lower incomes?	5	✓
Much conflict exists between the rich and poor	v36	In all countries, there are differences or even conflicts between different social groups. In your opinion, in [COUNTRY] how much conflict is there between ... poor people and rich people?	4	✓
The country's income distribution is fair	v50	How fair or unfair do you think the income distribution is in [COUNTRY]?	4	✓
Differences in income are too large	v21	To what extent do you agree or disagree with the following statements? Differences in income in [COUNTRY] are too large.	5	✓
Government should reduce income differences	v22	To what extent do you agree or disagree with the following statements? It is the responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes.	5	✓

Continued on next page

Table SI-5: International Social Survey Programme Outcomes (Continued)

Outcome	Variable Code	Survey Question	Points in Response	Scale Flipped?
Companies should reduce wage differences	v24	To what extent do you agree or disagree with the following statements? It is the responsibility of private companies to reduce the differences in pay between their employees with high pay and those with low pay.	5	✓
Higher earners should pay more tax share	v28	Do you think people with high incomes should pay a larger share of their income in taxes than those with low incomes, the same share, or a smaller share?	5	✓

Note: This table summarizes the ISSP survey questions used in the analysis. For each variable, it reports the label by which the question is referred to in the figures, the variable code of the question in the survey dataset, the exact question wording from the ISSP surveys, the number of points in the scale of responses for the question, and whether the scale was reversed for interpretability.

Table SI-6: Varieties of Democracy Indices

Index	Variable Code	Description
Clean Election Index	v2xel_frefair	To what extent are elections free and fair?
Rule of Law Index	v2x_rule	To what extent are laws transparently, independently, predictably, impartially, and equally enforced, and to what extent do the actions of government officials comply with the law?
Democracy Index		
Electoral Democracy Index	v2x_polyarchy	To what extent is the ideal of electoral democracy in its fullest sense achieved?
Liberal Democracy Index	v2x_libdem	To what extent is the ideal of liberal democracy achieved?
Participatory Democracy Index	v2x_partipdem	To what extent is the ideal of participatory democracy achieved?
Deliberative Democracy Index	v2x_delibdem	To what extent is the ideal of deliberative democracy achieved?
Egalitarian Democracy Index	v2x_egaldem	To what extent is the ideal of egalitarian democracy achieved?

Note: This table summarizes the V-Dem indices used as controls in our analysis of EC stratification and socio-political outcomes. For each index, the table reports the dataset variable code and a brief description.

2 Supplementary Information on Construction of EC Measures

2.1 Measuring and predicting socioeconomic status (SES)

This appendix provides additional details on our procedure for predicting the socioeconomic status (SES) of the users in our sample. We describe the survey used to measure SES for a subsample of users, the data sources and feature engineering steps used to train the machine-learning model, and details on the algorithm used in the prediction procedure. This discussion complements information in the Methods section of the main paper.

2.1.1 Survey eliciting components of SES

To collect information on Facebook users' SES, we use responses to a survey conducted by Facebook that appeared in users' News Feeds. The survey had two components: the first part elicited aspects of respondents' access to resources that are correlated with SES, and the second asked about household income quintiles. Figure SI-2 shows screenshots of the survey panels.

The survey was conducted in two waves. Wave 1 ran from July 12 to July 26, 2023, and Wave 2 from September 4 to September 18, 2024. The only change between waves was that Wave 1 elicited annual household income quintiles, while Wave 2 elicited monthly household income quintiles. Quintile cutoffs were based on post-tax income percentiles estimated by the World Inequality Database (Blanchet et al. 2021). Reported individual post-tax income was multiplied by 1.5 to account for average household size, and monthly cutoffs in Wave 2 were derived by dividing annual cutoffs by 12.

The list of countries surveyed is presented in the Methods Section. Overall, we observe 10,470 responses from countries in Africa, 30,567 responses from countries in Asia, 10,126 from countries in North America, 863 from countries in Oceania, and 6,985 responses from countries in South America.

2.1.2 Algorithmic Details

To predict users' SES, we train a gradient boosted tree model (XGBoost) on the SES index constructed from survey responses. The model is trained to minimize mean squared error with equal weights for all observations. Missing features are left unimputed during training. Hyperparameters are tuned using Bayesian optimization. After tuning, the optimal model parameters were left as defaults except for the following: number of trees = 250, learning rate = 0.035, and maximum depth = 5.

2.1.3 Features in the Machine Learning Model

We next describe the features used by the machine learning model to predict SES for users in our sample.

Self-Reported Education. Many individuals report their high school, college, and graduate school affiliations on their Facebook profiles. We transform these fields into binary indicators for high school, college, and graduate school attendance. To address misreporting—such as users listing prestigious universities like Harvard or Oxford without having actually attended—we require that an individual has at least five friends who also report attending the same institution for the affiliation to be considered valid. Because users often report only their highest level of education, the education features are constructed cumulatively: if a college is listed, we assume high school attendance as well; if a graduate school is listed, we assume both high school and college attendance.

Self-Reported Marriage Status. Self-reported marital status is encoded as a binary variable, with "one" indicating currently married and "zero" covering all other possible self-reported statuses (e.g., single, divorced, widowed).

Phone Characteristics. When users connect to Facebook through the mobile site or app, we observe characteristics of their phones, including the model and carrier. Each phone model is matched to a price in euros using either a private dataset from GSMArena or web-scraped price data collected with Anthropic's Claude model. In addition to current-day prices, we use historical GSMArena data to construct a "price at release" feature. We also measure "current phone age" by the date a given model is first observed accessing Facebook. Finally, we include "screen size" as a feature.

Facebook Usage Characteristics. We use three categories of user activity as features for the SES model: (1) Facebook usage, (2) marketplace listings, and (3) charity fundraiser donations.

Facebook usage. We capture both the length of time a user has been on the platform and their frequency of use over the past 28 days. The time since account creation is one of the strongest predictors of SES. While the mechanism is unclear, adoption of new technologies is often linked to socioeconomic status: in many developing countries, for example, having had internet access ten years ago versus only five years ago is a strong indicator of SES. In addition to account age, we include the number of days a user has logged into Facebook in the past 28 days and the number of those logins that occurred on a mobile device as SES model features.

Marketplace activity. Facebook Marketplace is one of the world's largest online marketplaces. We include the total number of products listed, the average price of listings, and the total value of listings as features in the SES model. Prices are converted to USD and adjusted for purchasing power parity.

Fundraiser donations. Facebook allows users to create and donate to fundraisers. We include the total monetary amount donated and the time since last donation as features for each user. Prices are converted to USD and adjusted for purchasing power parity.

Country-Level Variables. Cross-country differences in average SES are substantial, so we include three features that capture the wealth of a user's country of residence. Because the SES model must generalize beyond the countries in our survey sample, we avoid country fixed effects or indicator variables. Instead, we include GDP per capita and the Human Development Index (HDI) as features, which allow the model to classify users by the relative wealth of their country. We also include the Facebook penetration rate, defined as the share of a country's population with Facebook accounts.

Sub-National Income from Administrative Data. Subnational estimates of average income by geography from various statistical agencies are also used as features for the SES model where they are available. Table SI-7 details these data sources. All estimates are converted to USD.

Sub-National Income from External Estimates. We also include in our model regional SES estimates from two external datasets: the Relative Wealth Index (RWI) as described in Chi et al. (2022) and a recently published lat-long-gridded GDP dataset by Rossi-Hansberg and Zhang (2025).

The RWI (Chi et al. 2022) estimates the relative wealth of communities in the developing world using a machine learning model trained on DHS household surveys. The dataset provides estimates at the Bing Tile 14 level (roughly 2.5km squares at the equator) for most of the world, though quality varies by country. We include the RWI value for the Bing Tile of each user's residence as a feature.

Rossi-Hansberg and Zhang (2025) provide gridded estimates of GDP and GDP per capita worldwide. From this dataset, we use predicted GDP per capita for the 0.25-degree lat/long cell corresponding to a user's location.

Subnational Population Measures. The population of the GADM region within which a user resides is also used as a model feature. See Supplementary Information Section 1.8 for more information on the construction of population numbers.

Country-Level Encodings for Categorical Variables. Some features of a user's phone are categorical. These include the user's phone model, mobile carrier, and operating system. Directly including categorical features into our model would limit the ability of the model to generalize relationships between these features and the SES index. To get around this issue, we encode these categorical features as the conditional mean of other variables which have a clearer relationship with SES. This allows the model to learn that in a given country Carrier A has users with a higher average phone price than Carrier B, but this relationship need not be constant across countries. This conditional mean encoding approach is common in the machine learning literature. Below we list all conditional mean encodings included as model features. "Mean cell GDPpc" refers to the gridded GDP per capita estimates from Rossi-Hansberg and Zhang (2025). "Mean Cell GDPpc | Phone OS $\times$ Country" suggests that we calculate the Mean Cell GDPpc for each Phone OS by country combination, and provide that average value as a feature for all users from that country with that OS.

1. Mean Cell GDPpc | Phone OS $\times$ Country
2. Mean Cell GDPpc | Mobile Carrier $\times$ Country
3. Mean Cell GDPpc | Phone Model $\times$ Country
4. Mean Phone Price At Release | Mobile Carrier $\times$ Country
5. Mean Phone Price At Release | Phone OS $\times$ Country
6. Mean Phone Price At Release | Mobile Carrier $\times$ GADM Region
7. Mean Phone Price At Release | Phone OS $\times$ GADM Region
8. Mean Days Since Account Creation | Phone Model

2.1.4 Model Performance

Table SI-8 reports measures of fit for the machine learning model trained to predict SES. Overall, model performance is similar for the training and validation samples, and in the k-fold cross-validation approach. Furthermore, model performance declines only slightly when predicting SES for a country that

was dropped entirely from the training data set, suggesting that it should have similarly high predictive power for countries that were not included in the SES survey (and thus in the training data).

2.1.5 Feature Importance

Figure SI-5 summarizes the variables that most strongly influence the SES model. Each bar represents a feature in the gradient-boosted tree, with its length indicating its relative contribution to predicting socioeconomic status. The most important features are those capturing users' country-level economic context (e.g., GDP per capita), phone characteristics (such as price, model, and country- or region-specific encodings), and Facebook usage patterns (e.g., days on platform). Variables reflecting education, marital status, marketplace activity, and donation history play a comparatively smaller role.

2.2 Geographic Aggregation for Maps

GADM regions for maps are generally chosen to be the 2nd-level administrative boundaries collected by GADM – equivalent to a county in the US or a NUTS3 region in Europe. We up-level to 1st-level administrative boundaries or the country level for some countries with small populations or low Facebook usage. The following countries are up-leveled to the 1st-level GADM boundaries: Afghanistan, Algeria, Angola, Belarus, Bhutan, Brunei Darussalam, Bulgaria, Burundi, Central African Republic, Chad, Croatia, Democratic Republic of the Congo, Djibouti, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Eswatini, Gabon, Gambia, Ghana, Guatemala, Guinea-Bissau, Guyana, Honduras, Iceland, Japan, Kazakhstan, Kenya, Liberia, Malawi, Mali, Malta, Mauritania, Mauritius, Mongolia, Mozambique, Namibia, Netherlands, New Caledonia, New Zealand, Niger, Nigeria, Panama, Papua New Guinea, Qatar, Republic of the Congo, Romania, Rwanda, Sierra Leone, Slovenia, Somalia, South Sudan, Suriname, Switzerland, Tajikistan, Tanzania, Timor-Leste, Togo, Tunisia, Uganda, Uzbekistan, Yemen, Zambia, Zimbabwe. The following countries are up-leveled to their country boundaries: Andorra, Bahamas, Kiribati, Maldives, Micronesia, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Solomon Islands, Tonga, Vanuatu. We down-level to 3rd-level administrative boundaries for countries where we have large sample sizes and high confidence in our measures at that level. The following countries are down-leveled to 3rd-level GADM boundaries: Bosnia and Herzegovina, Chile, France, Greece, India, Nepal, Pakistan, South Africa, Spain, United Kingdom. For some countries, the GADM dataset defines identical boundaries across multiple administrative levels (1st-level, 2nd level, and 3rd level). For these countries, we simply retain the 2nd-level administrative boundaries for consistency, though these also correspond to 1st-level boundaries. This applies to the following countries: Armenia, Barbados, Bahrain, Belize, Cape Verde, Cyprus, Eritrea, Grenada, Equatorial Guinea, Israel, Jamaica, Comoros, Kuwait, Lesotho, Libya, Moldova, Montenegro, North Macedonia, Mauritius, and Singapore.

2.3 Comparison with Prior Estimates of Economic Connectedness

To assess the validity of our measure of low-SES EC, we compare it with previously developed metrics for both the United States and the United Kingdom (Fig. SI-7). Low-SES EC is defined as the predicted value at the 25th percentile of the friend rank–rank regression, whereas the prior measures capture the average share of above-median-SES friends among below-median-SES individuals, normalized by 50%. For the U.S., we examine its relationship with estimates from Chetty, Jackson, et al. (2022a) and Chetty, Jackson, et al. (2022b) at the county level; for the U.K., we compare it with estimates from Harris et al.

(2025) at the local authority level. In both cases, we find strong positive correlations.

3 Differential Privacy

During the construction of the aggregate statistics underlying this paper and the public data release, we take several steps to protect the privacy of individuals included in the estimates for each cell. First, we do not release data for any geographic unit containing fewer than 500 users. Second, for the cells that we do release, we add noise to ensure that two cells which differ by one observation are sufficiently difficult for an adversary to distinguish, following a procedure termed μ-Gaussian Differential Privacy (Dong et al. 2022). Specifically, for each operation involved in calculating a statistic, we add noise that is bounded from below by $\mathcal{N}(0, \frac{s}{\mu})$, where μ is set to be approximately 1.02, and s represents the maximum sensitivity of the calculation to the exclusion of a given individual. The sensitivity is defined using the empirical distribution of the variable across individuals in a given cell–that is, the sensitivity is set such that $s = max(x) - min(x)$, where x refers to the statistic being aggregated across individuals.

Some of the statistics are a function of several aggregations from the individual-level data, and thus incorporate several uncorrelated random noise draws. For instance, when calculating the average income centile in a geography, we first calculate the sum of the individual-level centiles in that geography, before adding noise drawn from $\mathcal{N}\left(0, \frac{\max(SES) - \min(SES)}{\mu}\right)$, where $max(SES)$ and $min(SES)$ respectively represent the maximum and minimum SES centiles of the users in the geography. We then calculate the number of individuals in that cell, add noise to that figure from the distribution $\mathcal{N}(0, \frac{1}{\mu})$, and then divide the two figures. In this way, we can ensure that the amount of noise we add to the final figure is bounded from below by the true sensitivity of the complete calculation of a given figure, the direct calculation of which would often involve complicated formulae.

We present summary statistics highlighting the weighted and unweighted correlations between the pre- and post-differential privacy estimates in Tables SI-14 and SI-15. In almost all cases, the correlation between these figures is above 0.99 after weighting cells by the number of users included in each estimate. The generally lower unweighted correlations reflect that the amount of noise we add in cells with fewer users is generally larger as a proportion of the raw value of the statistic we calculate.

4 Supplementary References

Acciari, Paolo et al. (July 2022). "And Yet It Moves: Intergenerational Mobility in Italy". In: *American Economic Journal: Applied Economics* 14.3, pp. 118–163.

Alesina, Alberto et al. (2021). "Intergenerational Mobility in Africa". In: *Econometrica* 89.1, pp. 1–35.

Asher, Sam et al. (2024). "Intergenerational Mobility in India: New Measures and Estimates across Time and Social Groups". In: *American Economic Journal: Applied Economics* 16.2, pp. 66–98.

Blanchet, Thomas et al. (2021). "Distributional National Accounts Guidelines". In: *Methods and Concepts Used in the World Inequality Database. World Inequality Lab*.

Britto, Diogo et al. (2024). "Intergenerational Mobility in the Land of Inequality". SSRN Working Paper 4241590.

Chetty, Raj, John N. Friedman, et al. (2026). "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility". In: *American Economic Review*.

Chetty, Raj, Matthew O. Jackson, et al. (Aug. 2022a). "Social Capital I: Measurement and Associations with Economic Mobility". In: *Nature* 608.79217921, pp. 108–121.

— (Aug. 2022b). "Social Capital II: Determinants of Economic Connectedness". In: *Nature* 608.79217921, pp. 122–134.

Chi, Guanghua et al. (Jan. 2022). "Microestimates of Wealth for All Low- and Middle-Income Countries". In: *Proceedings of the National Academy of Sciences* 119.3, e2113658119.

Cholli, Neil A. et al. (2024). "Understanding the Heterogeneity of Intergenerational Mobility Across Neighborhoods". NBER Working Paper w33035.

Chuard, Patrick and Veronica Schmiedgen-Grassi (July 2020). "Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education". SSRN Working Paper 3662560.

Connolly, Marie et al. (2019). "Intergenerational Mobility Between and Within Canada and the United States". In: *Journal of Labor Economics* 37.S2.

Coppedge, Michael et al. (2025). *V-Dem [Country-Year/Country-Date] Dataset v15*.

Crichton, Sarah and Jess Jenkins (2026). "Stuck or Soaring? A Geography of Intergenerational Income Mobility in Aotearoa New Zealand". Working Paper / Seminar Presentation. Working Paper.

Del Pozo, Diego and Lorena Moreno (2024). "Intergenerational Mobility in Ecuador: A First Approach". National Institute of Statistics and Censuses Working Paper 19.

Deutscher, Nathan and Bhashkar Mazumder (2020). "Intergenerational Mobility Across Australia and the Stability of Regional Estimates". In: *Labour Economics* 66, p. 101861.

Dodin, Majed et al. (2024). "Social Mobility in Germany". In: *Journal of Public Economics* 232, p. 105074.

Dong, Jinshuo et al. (2022). "Gaussian differential privacy". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 84.1, pp. 3–37.

Eurostat (2021). *Income of households by NUTS 2 region*. Eurostat. URL: https://ec.europa.eu/eurostat/databrowser/view/nama_10r_2hhinc__custom_13225473/.

— (2025). *Population density by NUTS 3 region*. Eurostat. URL: https://ec.europa.eu/eurostat/databrowser/product/view/DEMO_R_D3DENS.

EVS/WVS (2022). *European Values Study and World Values Survey: Joint EVS/WVS 2017–2022 Dataset (Joint EVS/WVS)*. Version 5.0.0. ZA7505. GESIS Data Archive, Cologne.

Gallup, Inc. (2024). *Gallup World Poll*. Accessed through Gallup Analytics. URL: https://www.gallup.com/analytics/318875/global-research.aspx.

Haerpfer, Christian et al. (2020). *World Values Survey: Round Seven – Country-Pooled Datafile*. JD Systems Institute & WVSA Secretariat, Madrid & Vienna.

Harris, Tom et al. (2025). "Social Capital in the United Kingdom: Evidence from Six Billion Friendships". SocArXiv Working Paper.

ISSP Research Group (2019). *International Social Survey Programme: Social Networks and Social Resources - ISSP 2017*. ZA6980 Data file Version 2.0.0.

— (2022). *International Social Survey Programme: Social Inequality V - ISSP 2019*. ZA7600 Data file Version 3.0.0.

Kenedi, Gustave and Louis Sirugue (2023). "Intergenerational Income Mobility in France: A Comparative and Geographic Analysis". In: *Journal of Public Economics* 226, p. 104974.

Lam, Helen et al. (2025). "How, When, and Where Does the Opportunity Gap in the Netherlands Open Up?" Working Paper.

Lebakula, Viswadeep et al. (2024). *LandScan Silver Edition*. LandScan Global. CY 2023, raster digital data. Released October 9, 2024. Oak Ridge, TN: Oak Ridge National Laboratory.

Maerz, Seraphine F. et al. (2025). *vdemdata: An R package to load, explore and work with the most recent V-Dem (Varieties of Democracy) dataset*. URL: https://github.com/vdeminstitute/vdemdata.

Milanovic, Branko (2019). *All the Ginis (ALG) Dataset, Version February 2019*. Stone Center on Socio-Economic Inequality, Graduate Center, City University of New York. URL: https://stonecenter.gc.cuny.edu/research/all-the-ginis-alg-dataset-version-february-2019/.

Munoz, Ercio and Roy van der Weide (2025). *Intergenerational Income Mobility Around the World: A New Database*.

Muñoz Saavedra, Ercio A. (2024). "The Geography of Intergenerational Mobility in Latin America and the Caribbean". In: *Economía LACEA* 23.1.

Nybom, Martin and Jan Stuhler (2025). "Geographic Variation in Multigenerational Mobility". In: *Sociological Methods & Research*.

Office of the Registrar General and Census Commissioner, India (2011). *Census of India*. Accessed via SHRUG on September 10, 2025 https://www.devdatalab.org/shrug_download/, v2.1.pakora.

Rossi-Hansberg, Esteban and Jialing Zhang (Feb. 2025). "Local GDP Estimates Around the World". NBER Working Paper w33458.

Solt, Frederick (2020). "Measuring Income Inequality Across Countries and Over Time: The Standardized World Income Inequality Database". In: *Social Science Quarterly*. SWIID Version 9.9, June 2025.

Soria Espin, Pedro Javier and Octavio Medina (2025). "Intergenerational Income Mobility in Spain: Geographic Analysis and Causal Effects of Places". Working Paper.

U.S. Census Bureau (2018). *2018 American Community Survey 1-Year Estimates*. United States Census Bureau. URL: https://data.census.gov/cedsci/.

— (2022). *American Community Survey 5-Year Data (2008-2022)*. URL: https://www.census.gov/data/developers/data-sets/acs-5year.html.

Van der Weide, Roy et al. (2024). "Intergenerational mobility around the world: A new database". In: *Journal of Development Economics* 166, p. 103167.

World Bank (2024). *GDP per capita (current USD)*. URL: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

— (2025). *Gini index (World Bank estimate)*. URL: https://data.worldbank.org/indicator/SI.POV.GINI.

Young, Clifford (2025). *The Ipsos Populism Report*. URL: https://www.ipsos.com/sites/default/files/ct/news/documents/2025-06/ipsos-populism-report-2025.pdf.

Zaman, Qamar et al. (Oct. 2024). *Revealing Social Mobility I: Income mobility across local authorities in England*. Blog post, Behavioural Insights Team.

5 Supplementary Tables and Figures

Table SI-7: External data sources used in SES model

Description	Data Source	Version
GDP per Capita (Current USD)	World Bank	2023 or latest
Median Household Income in the Past 12 Months (in 2023 Inflation-Adjusted Dollars) by ZCTA	US Census Bureau	2023 5-Year Estimates
Gross domestic product per capita at current market prices by NUTS 3 region	Eurostat	2019
Median Household Income before Tax by Census Division in Constant CAD	StatCan	2020 data
Mean Personal Income by SA4 Region	Australian Bureau of Statistics	2021-2022 data
Estimates of annual household income for the four income types for Middle layer Super Output Areas, or local areas, in England and Wales	UK Office for National Statistics	2020 Fiscal Year
Human Development Index	UN Development Programme	2025 version
GDP Estimates by Quarter-Degree Tile	Rossi-Hansberg and Zhang (2025)	2024 version
Relative Wealth Index	Chi et al. (2022)	2023 version

Note: This table documents the external data sources used in the construction of the SES model, and provides a link to the dataset along with the version of the data used.

Table SI-8: Measure of Fit for SES Model

Sample	Correlation	RMSE	MAE	Sample Size
Training	0.64	1.34	1.04	66,764
Validation	0.62	1.36	1.07	7,468
IID K-Fold CV (10 Folds)	0.62	1.37	1.07	74,232
Leave One Country Out CV	0.59	1.41	1.10	74,232

Note: This table reports measures of fit for the machine learning model trained to predict SES for individuals in our main analytic samples. Four different samples and training specifications are shown. Training refers to the sample used to train the SES model. Validation refers to 10% of observations from the survey held-out during training. K-Fold CV refers to a common cross validation procedure where the model is fit on 90% of the data and predictions are generated for a hold-out sample, the procedure is repeated 10 times (folds) to generate out-of-sample predictions for the entirety of the training data. Leave One Country Out CV refers to the procedure where an entire country is held-out of the training data, and predictions are generated from a model fit on the remaining 63 countries.

Table SI-9: EC Summary Statistics by Geographic Aggregation, All Available Countries

Summary Statistics

(a) EC Stratification

Level	Mean	P5	P10	P25	P50	P75	P90	P95	N
Country-Level (Equal Weights)	0.26	0.13	0.14	0.19	0.26	0.32	0.40	0.43	178
Country-Level (Pop. Weighted)	0.30	0.16	0.18	0.23	0.32	0.36	0.41	0.43	178
GADM1 (Pop. Weighted)	0.26	0.12	0.15	0.18	0.28	0.32	0.35	0.39	3,117
GADM2 (Pop. Weighted)	0.23	0.12	0.13	0.16	0.23	0.28	0.32	0.34	36,111
GADM3 (Pop. Weighted)	0.22	0.11	0.13	0.17	0.22	0.27	0.32	0.33	84,138
GADM-Best (Pop. Weighted)	0.23	0.12	0.14	0.18	0.22	0.28	0.32	0.33	30,737

(b) Low-SES EC

Level	Mean	P5	P10	P25	P50	P75	P90	P95	N
Country-Level (Equal Weights)	48.90	44.10	45.34	47.04	49.31	50.81	51.96	52.42	178
Country-Level (Pop. Weighted)	47.63	43.80	45.30	46.59	47.09	49.10	50.77	51.82	178
GADM1 (Pop. Weighted)	48.02	44.76	45.35	46.50	48.40	48.63	51.01	51.85	3,117
GADM2 (Pop. Weighted)	48.33	44.91	45.62	47.02	48.53	48.94	51.42	51.78	36,111
GADM3 (Pop. Weighted)	48.48	45.19	45.96	47.08	48.67	48.91	51.50	51.83	84,138
GADM-Best (Pop. Weighted)	48.30	44.91	45.62	46.98	48.53	48.91	51.26	51.85	30,737

Correlation Matrices

(c) EC Stratification

	Country-Level	GADM1 Mean	GADM2 Mean	GADM3 Mean	GADM-Best Mean
Country-Level	1.00				
GADM1 Mean	0.96	1.00			
GADM2 Mean	0.93	0.99	1.00		
GADM3 Mean	0.92	0.98	0.99	1.00	
GADM-Best Mean	0.94	0.99	0.99	0.99	1.00

(d) Low-SES EC

	Country-Level	GADM1 Mean	GADM2 Mean	GADM3 Mean	GADM-Best Mean
Country-Level	1.00				
GADM1 Mean	0.96	1.00			
GADM2 Mean	0.95	0.98	1.00		
GADM3 Mean	0.94	0.96	0.98	1.00	
GADM-Best Mean	0.95	0.99	0.99	0.97	1.00

This table reports summary statistics and correlation matrices for EC measures across all available countries and regions, at multiple levels of geographic aggregation. **a, b,** Summary statistics for EC stratification and Low-SES EC, respectively. For each variable, we report the mean, percentiles (5th, 10th, 25th, 50th, 75th, 90th, and 95th), and sample size (N) across multiple geographic aggregation levels. These levels include: (i) country-level statistics with equal weighting across countries, (ii) country-level statistics weighted by country populations, and (iii) regional statistics at GADM1, GADM2, GADM3, and GADM-Best levels, where regions are weighted by their populations. **c, d,** Correlation matrices for EC stratification and Low-SES EC, respectively. Each matrix reports unweighted correlation coefficients between country-level measures constructed either directly, or by calculating population-weighted averages of the same measure across regions in a country.

Table SI-10: Correlates of EC Stratification: Cross-Country, All Available Countries

(a) Intergenerational Mobility

	Intergenerational Income Mobility			Intergenerational Educational Mobility		
	(1)	(2)	(3)	(4)	(5)	(6)
EC Stratification	-0.446***		-0.291***	-0.453***		-0.407***
	(0.082)		(0.110)	(0.082)		(0.101)
Inequality		-0.437***	-0.270**		-0.294***	-0.094
		(0.093)	(0.116)		(0.098)	(0.104)
Observations	86	86	86	108	108	108
Adjusted R^2	0.189	0.181	0.230	0.198	0.078	0.197

(b) Interpersonal and Institutional Trust, Pro-Social Norms

	Trust Most People		Confidence in Courts		Cheating on Taxes is Justifiable		Votes are Counted Fairly	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EC Stratification	-0.634***	-0.564***	-0.360***	-0.359**	0.376***	0.408***	-0.580***	-0.428***
	(0.100)	(0.125)	(0.112)	(0.157)	(0.094)	(0.125)	(0.082)	(0.124)
Inequality		-0.124		-0.046		-0.057		-0.046
		(0.095)		(0.123)		(0.121)		(0.102)
Rule of Law Index				-0.058				
				(0.112)				
Clean Election Index								0.374***
								(0.090)
Observations	86	86	85	84	86	86	86	85
Adjusted R^2	0.395	0.398	0.119	0.102	0.131	0.123	0.328	0.432

(c) Political Attitudes, Preference for Redistribution, and Life Satisfaction

	Importance of Democracy		Support for Unchecked Leadership		Incomes Should be Equalized		Overall Life Satisfaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EC Stratification	-0.436***	-0.293**	0.648***	0.444***	-0.249***	-0.373***	-0.085	0.044
	(0.088)	(0.128)	(0.069)	(0.114)	(0.086)	(0.121)	(0.104)	(0.200)
Inequality		-0.006		0.043		0.219*		0.079
		(0.118)		(0.088)		(0.128)		(0.176)
Democracy Index		0.273**		-0.363***				0.346**
		(0.106)		(0.097)				(0.149)
Observations	86	85	86	85	86	86	86	85
Adjusted R^2	0.180	0.211	0.413	0.493	0.051	0.073	-0.005	0.077

This table reports estimates from regressions of country-level outcomes on EC stratification for all available countries. EC stratification is the slope from regressing the average SES percentile of an individual's friends on the individual's own SES percentile. Panel A uses intergenerational income and educational mobility data from the World Bank's Global Database on Intergenerational Mobility (see SI Section 1.2). Inequality is measured using Gini coefficients drawn from the World Bank, the Standardized World Income Inequality Database, and All the Ginis (see SI Section 1.6). Panel A employs lagged measures of inequality to approximate the levels of inequality individuals were exposed to during childhood. In Panels B and C, which examine present-day sociopolitical outcomes, we use contemporary inequality data. Panels B and C use country-level aggregates of survey responses from the Joint European Values Study/World Values Survey (see SI Section 1.5.1). The Clean Election Index, Rule of Law Index, and Democracy Index are expert-coded measures from the V-Dem dataset, averaged over 2018-2024 (see SI Section 1.7). All dependent and independent variables are standardized to have mean zero and unit variance. Robust standard errors are reported in parentheses. Asterisks indicate the level of statistical significance: *10%, **5%, ***1%.

Table SI-11: Cross-Country Socio-Political Outcomes against Inequality

(a) Interpersonal and Institutional Trust, Pro-Social Norms

	Trust Most People (1)	Confidence in Courts (2)	Cheating on Taxes is Justifiable (3)	Votes are Counted Fairly (4)
Inequality	-0.438***	-0.251**	0.285**	-0.355***
	(0.088)	(0.115)	(0.112)	(0.092)
Observations	73	72	73	73
Adjusted R^2	0.181	0.050	0.068	0.113

(b) Political Attitudes, Preference for Redistribution, and Life Satisfaction

	Importance of Democracy (1)	Support for Unchecked Leadership (2)	Incomes Should be Equalized (3)	Overall Life Satisfaction (4)
Inequality	-0.279***	0.450***	-0.013	0.126
	(0.094)	(0.083)	(0.101)	(0.104)
Observations	73	73	73	73
Adjusted R^2	0.065	0.191	-0.014	0.002

Note: This table presents ordinary least squares (OLS) regression estimates examining the relationship between country-level socio-political outcomes and inequality. The unit of observation is the country, restricted to countries with data quality marked green. The data on outcomes are country-level aggregates of survey responses from the Joint European Values Study/World Values Survey. The table assesses the correlation of these outcomes with income inequality, measured by Gini coefficients from the World Bank, the Standardized World Income Inequality Database, and the All the Ginis database (see SI Section 1.6). All dependent and independent variables are standardized to have mean zero and unit variance. Robust standard errors are reported in parentheses. Statistical significance is indicated by asterisks: *10%, **5%, ***1%.

Table SI-12: Cross-Country Regression Estimates of Alternative Samples and Measures of Intergenerational Educational Mobility on EC Stratification

	Intergenerational Educational Mobility					
	(1)	(2)	(3)	(4)	(5)	(6)
EC Stratification	-0.482***	-0.506***	-0.343***	-0.594***	-0.444***	-0.405***
	(0.120)	(0.129)	(0.103)	(0.119)	(0.129)	(0.123)
Inequality	0.061	0.092	-0.060	0.195	-0.011	0.149
	(0.130)	(0.141)	(0.123)	(0.125)	(0.136)	(0.124)
Robustness Check	Sample: 1960s cohort	Sample: 1970s cohort	Sample: 1980s cohort, also co-resident surveys	LHS: Expected Rank | Parents H1	LHS: Probability Q4 | Parents H1	LHS: Beta Schooling $\sim$ Parents' Schooling
Observations	84	84	106	84	84	84
Adjusted R^2	0.182	0.188	0.122	0.221	0.184	0.086

Note: This table reports estimates from ordinary least squares (OLS) regressions of country-level intergenerational educational mobility on EC stratification. It provides robustness checks for the main analysis using alternative samples and mobility measures. The dependent variable is one minus the correlation between children's and parents' years of schooling, unless otherwise noted. The key regressor, EC stratification, is defined as the slope from a regression of average friend (national) SES percentile on individuals' own (national) SES percentile. Inequality is measured by the Gini coefficient, harmonized from the World Bank, the Standardized World Income Inequality Database, and the All the Ginis database. Intergenerational mobility data come from the Global Database of Intergenerational Mobility (Munoz and van der Weide 2025). Columns 1–2 use alternative decade-specific cohorts (1960s, 1970s) with matching inequality data; Column 3 uses the 1980s cohort but includes co-resident surveys to expand coverage beyond retrospective sources. Columns 4–6 employ alternative mobility indicators: MU050 (expected child education rank for those born in the bottom half), BHQ4 (probability of reaching the top education quartile conditional on being born in the bottom half), and BETA (slope coefficient from regressing children's on parents' schooling years). All variables are standardized to have mean 0 and unit variance. Robust standard errors are in parentheses. Asterisks indicate the level of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table SI-13: Preferences for Redistribution

	Incomes Should be Equalized				
	(1)	(2)	(3)	(4)	(5)
EC Stratification	-0.243**		-0.480***		-0.456***
	(0.115)		(0.173)		(0.159)
Inequality		0.077	0.378**		0.397**
		(0.136)	(0.186)		(0.161)
Imp. of Luck for Success				0.358**	0.350**
				(0.150)	(0.154)
Observations	49	49	49	49	49
Adjusted R^2	0.039	-0.015	0.109	0.110	0.217

Note: This table reports estimates from ordinary least squares (OLS) regressions of country-level measures on an outcome from the Joint European Values Study/World Values Survey about whether respondents believe that "incomes should be made more equal." The unit of observation is a country; the sample includes all green data quality countries that are in the World Values Survey. All dependent and independent variables are standardized to have mean zero and unit variance. Robust standard errors are reported in parentheses. Asterisks indicate the level of significance: *10%, **5%, ***1%.

Table SI-14: Weighted Correlations between Pre- and Post-DP Measures by Geography

Variable	Countries	GADM1	GADM2	GADM3
EC Stratification	1.000	1.000	0.999	0.996
Low-SES EC	1.000	1.000	1.000	0.996
Average Friend Centile of Ego Centile 5	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 10	1.000	0.996	NA	NA
Average Friend Centile of Ego Centile 15	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 20	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 25	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 30	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 35	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 40	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 45	1.000	1.000	NA	NA
Average Friend Centile of Ego Centile 50	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 55	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 60	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 65	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 70	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 75	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 80	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 85	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 90	1.000	0.999	NA	NA
Average Friend Centile of Ego Centile 95	1.000	0.999	NA	NA

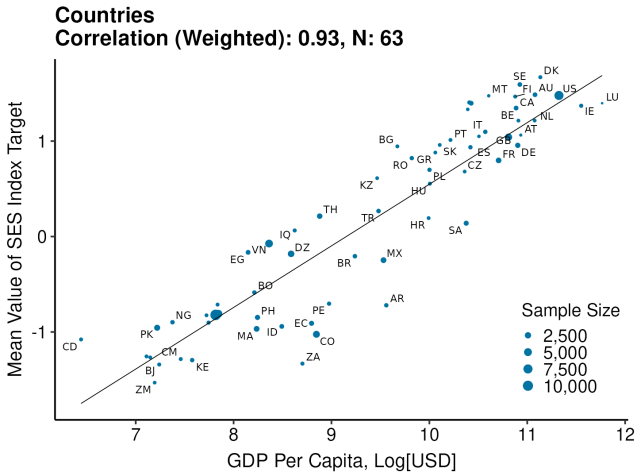
Note: This table reports the population-weighted correlations between the pre- and post-differential privacy values of various EC-related measures at different levels of geographic aggregation.

Table SI-15: Unweighted Correlations between Pre- and Post-DP Measures by Geography

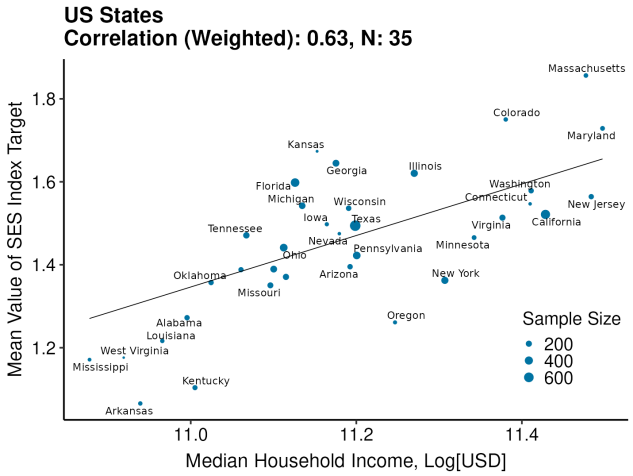
Variable	Countries	GADM1	GADM2	GADM3
EC Stratification	1.000	0.999	0.994	0.989
Low-SES EC	1.000	1.000	1.000	0.999
Average Friend Centile of Ego Centile 5	0.996	0.944	NA	NA
Average Friend Centile of Ego Centile 10	0.996	0.937	NA	NA
Average Friend Centile of Ego Centile 15	0.997	0.895	NA	NA
Average Friend Centile of Ego Centile 20	0.997	0.929	NA	NA
Average Friend Centile of Ego Centile 25	0.995	0.893	NA	NA
Average Friend Centile of Ego Centile 30	0.997	0.847	NA	NA
Average Friend Centile of Ego Centile 35	0.990	0.888	NA	NA
Average Friend Centile of Ego Centile 40	0.996	0.835	NA	NA
Average Friend Centile of Ego Centile 45	0.990	0.850	NA	NA
Average Friend Centile of Ego Centile 50	0.997	0.845	NA	NA
Average Friend Centile of Ego Centile 55	0.985	0.845	NA	NA
Average Friend Centile of Ego Centile 60	0.991	0.826	NA	NA
Average Friend Centile of Ego Centile 65	0.986	0.859	NA	NA
Average Friend Centile of Ego Centile 70	0.964	0.844	NA	NA
Average Friend Centile of Ego Centile 75	0.993	0.834	NA	NA
Average Friend Centile of Ego Centile 80	0.983	0.851	NA	NA
Average Friend Centile of Ego Centile 85	0.996	0.810	NA	NA
Average Friend Centile of Ego Centile 90	0.980	0.787	NA	NA
Average Friend Centile of Ego Centile 95	0.991	0.806	NA	NA

Note: This table reports the unweighted correlations between the pre- and post-differential privacy values of various EC-related measures at different levels of geographic aggregation.

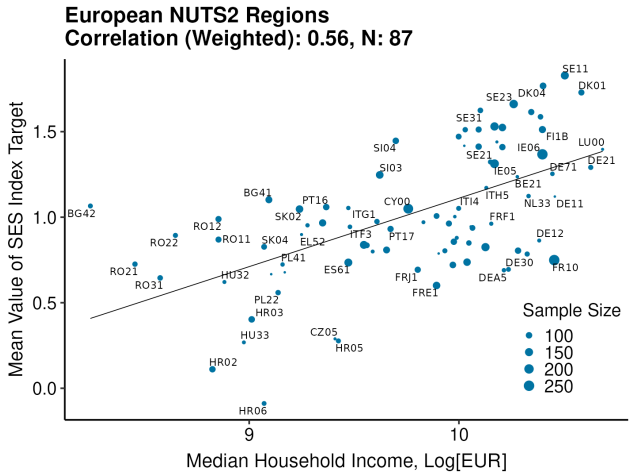
Fig. SI-1: Validation of SES Target with Administrative Data



(a) GDP Per Capita



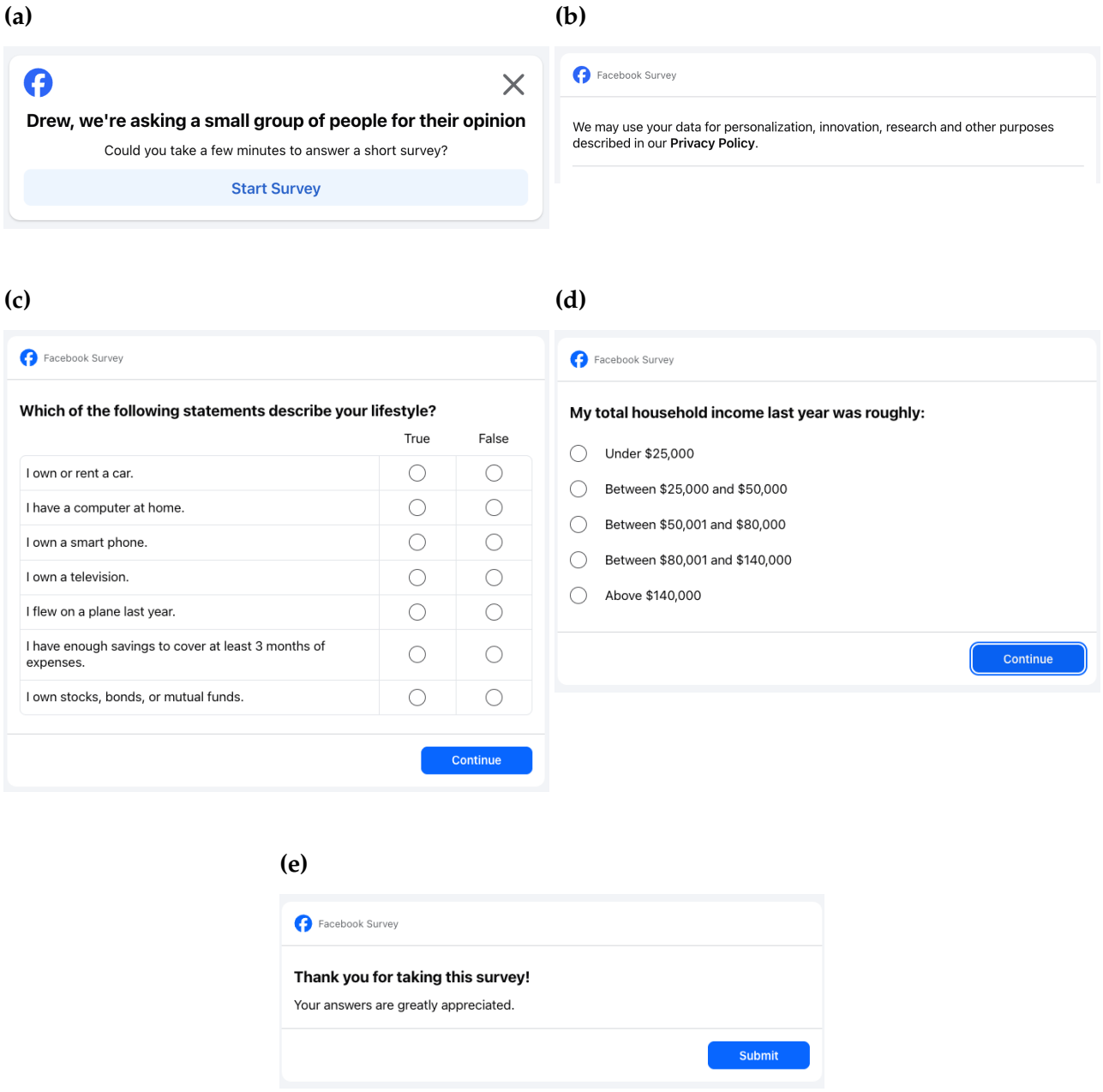
(b) Household Income (U.S. States)



(c) Household Income (NUTS2 Regions)

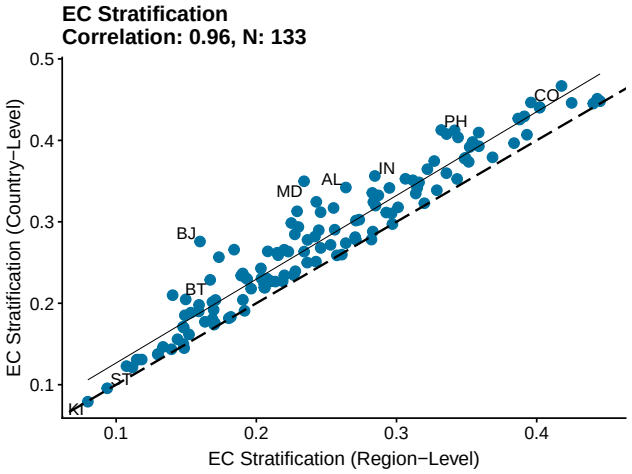
Note: Aggregated values of the survey-based SES index correlated against aggregate measures of regional income. Panel (a) shows country-level averages of the SES index versus GDP per capita in 2021, expressed in current U.S. dollars. GDP data is from the World Bank (2024). Panels (b) and (c) show state/region level averages of the SES Index plotted against the median household income. Household income data for the US comes from the 2022 5-Year American Community Survey (U.S. Census Bureau 2022). Household income data for Europe comes from Eurostat (2021).

Fig. SI-2: Screenshots of SES Survey on Facebook

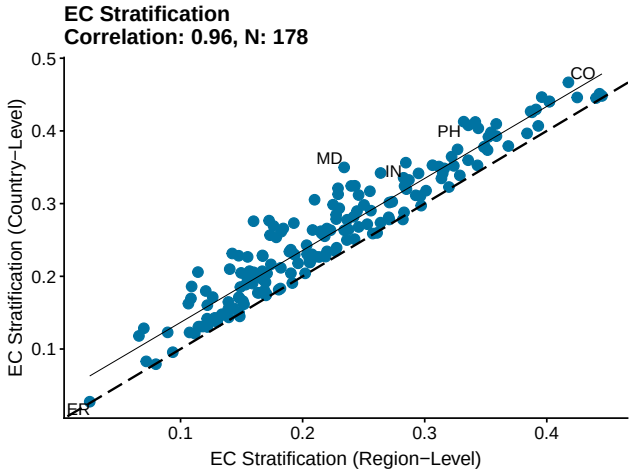


Note: (a) Pop-up message in Facebook news feed inviting participation in survey. (b) Disclaimer stating that user data may be used for personalization, innovation, research, and other purposes. (c) Survey panel on access to resources in the version of the survey used in the United States. (d) Question on annual household income in the version of the survey used in the United States. (e) Message thanking the respondent for their participation in the survey.

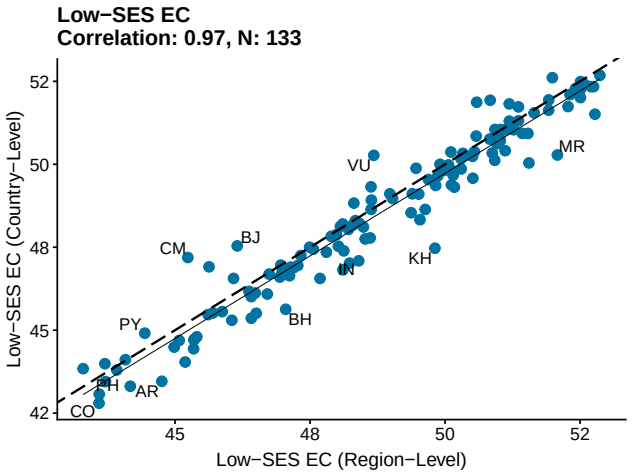
Fig. SI-3: Country-Level EC against Weighted Average of GADM1-Level EC



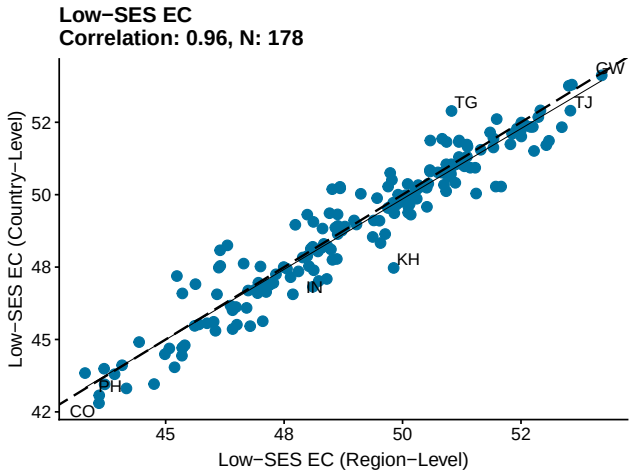
(a) EC Stratification, Green Countries



(b) EC Stratification, All Available Countries



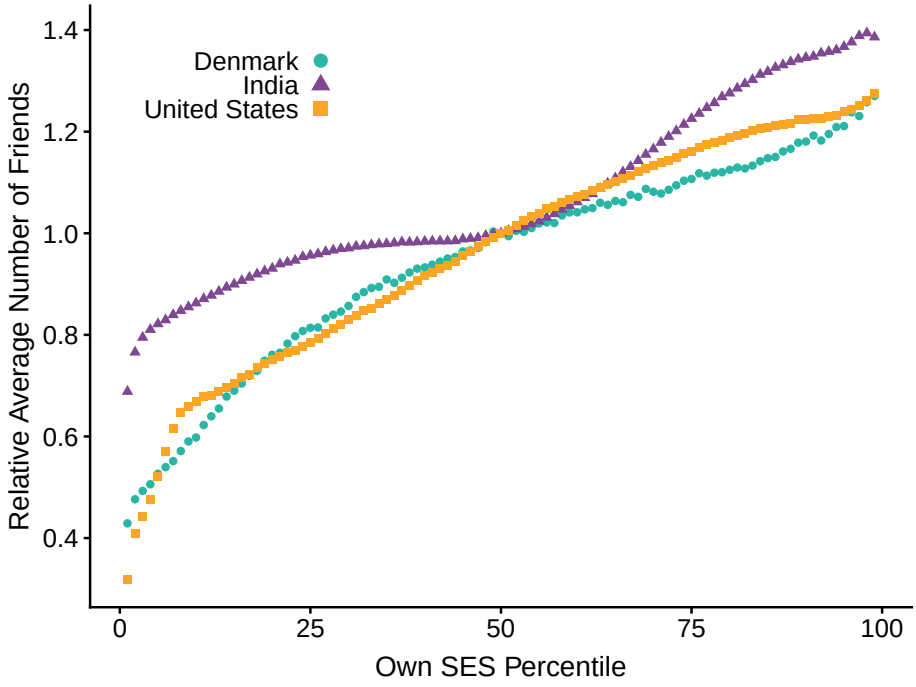
(c) Low-SES EC, Green Countries



(d) Low-SES EC, All Available Countries

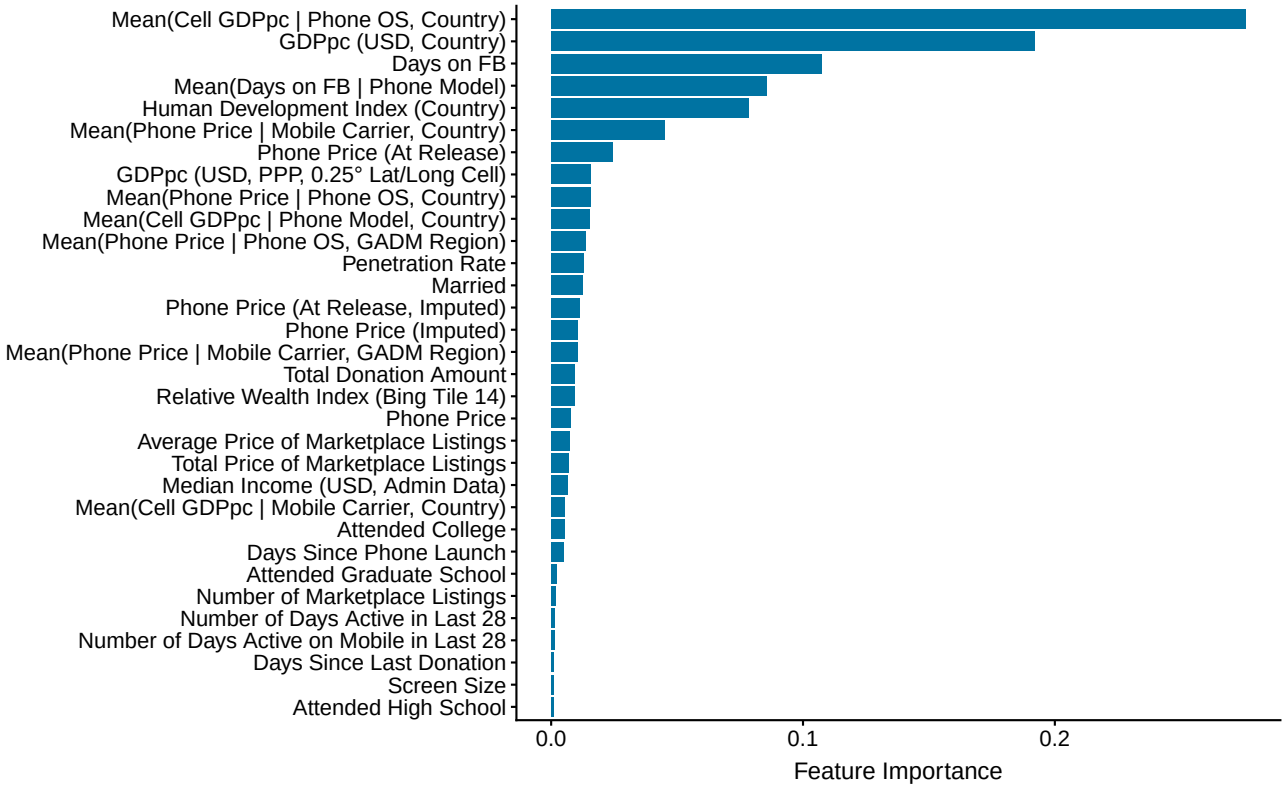
Note: Each scatter plot displays the relationship between a country-level EC measure and the corresponding population-weighted averages of the same measure across GADM1 regions in a country. The top row show this relationship for EC stratification in only countries with green data quality (Panel a) and in all available countries (Panel b), while the bottom row shows this relationship for Low-SES EC in only countries with green data quality (Panel c) and in all available countries (Panel d). The dashed line is the 45-degree line, the solid line the line of best fit from a linear regression.

Fig. SI-4: Relative Number of Friends by Own SES Percentile



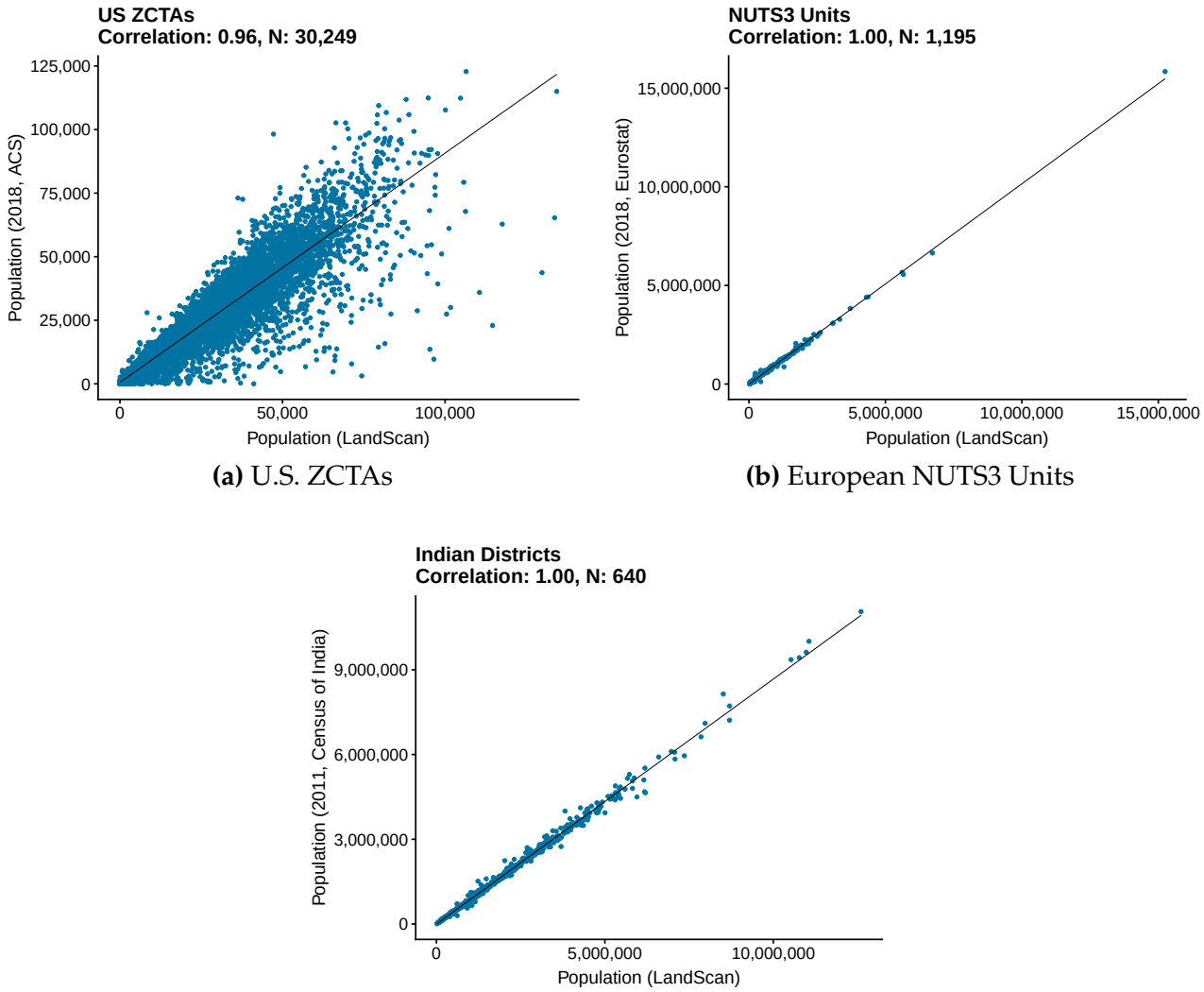
Note: This figure plots the relative average number of friends by own SES percentile as a ratio of number of friends of someone at the 50th SES percentile in Denmark, India, and the United States.

Fig. SI-5: SES Model Feature Importance



Note: Feature importance scores for the gradient boosted tree trained to predicted SES of individuals. Importance scores are the total gain of each feature normalized to add up to 1.

Fig. SI-6: Comparison of calculated population estimates using LandScan with administrative population data for US ZCTAs, European NUTS3 regional units, and Indian districts



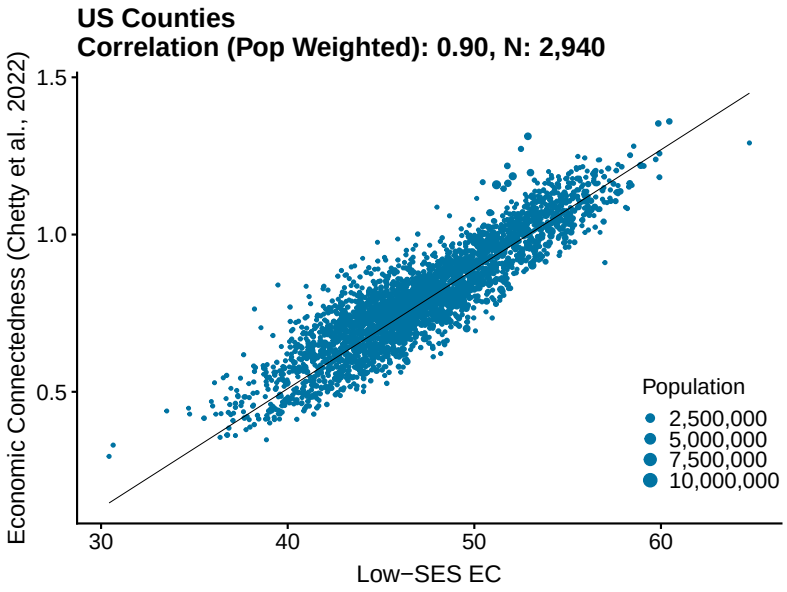
(a) U.S. ZCTAs

(b) European NUTS3 Units

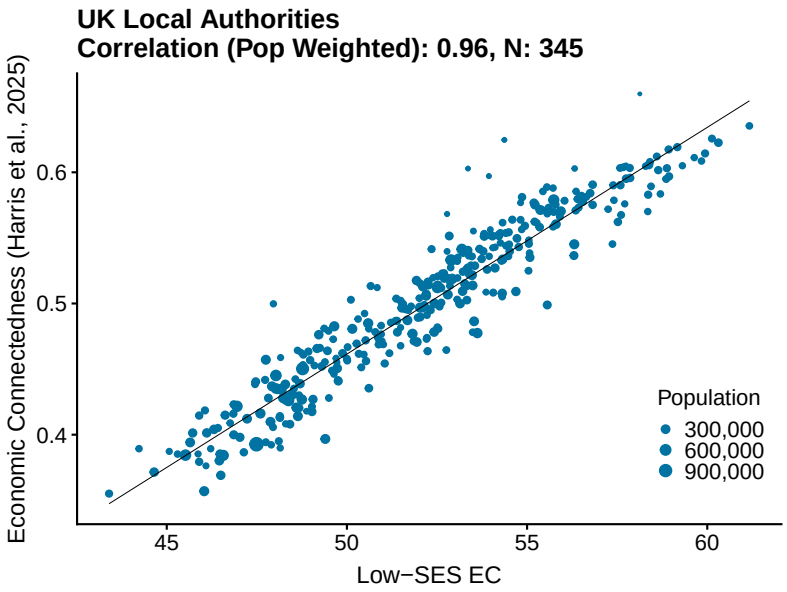
(c) Indian Districts

Note: Population estimates used in our analyses are constructed by aggregating 1km-by-1km LandScan cells (Lebakula et al. 2024) to the required geographic units, splitting cells across boundaries by land-area shares. This figure compares those estimates with administrative data on population. **a**, US ZCTA population estimates from the 2018 American Community Survey (U.S. Census Bureau 2018). **b**, European NUTS3 regional units data on population density in 2018 from Eurostat (Eurostat 2025). Measures are converted from population density to raw population by multiplying with area. **c**, Indian district population data from the 2011 Census of India, accessed via SHRUG (Office of the Registrar General and Census Commissioner, India 2011).

Fig. SI-7: Comparison with existing Economic Connectedness Measures in US and UK



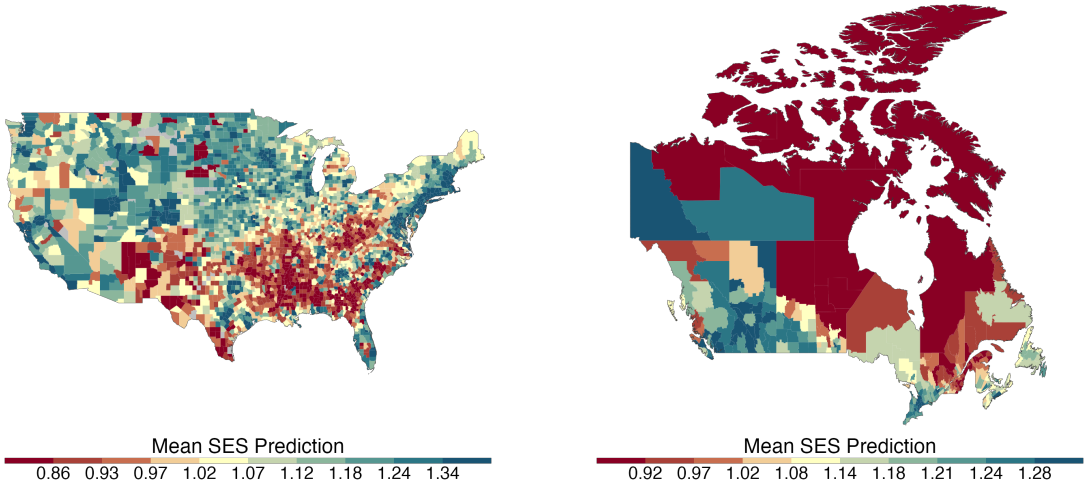
(a) U.S. Counties



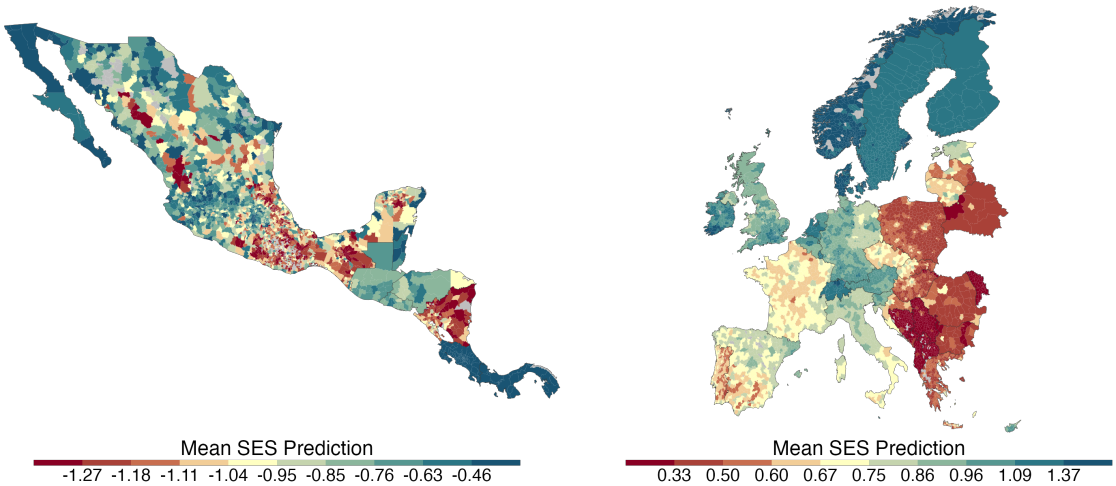
(b) UK Local Authorities

Note: Each scatter plot compares measures of Low-SES EC measures to prior measures of economic connectendess for the corresponding geographic unit. Low-SES EC is defined as the predicted value at the 25th percentile of the friend rank-rank regression. The prior measures, as defined by Chetty, Jackson, et al. (2022a) for US counties and Harris et al. (2025) for UK local authorities, calculate the average share of above-median-SES friends among below-median-SES individuals, normalized by 50%.

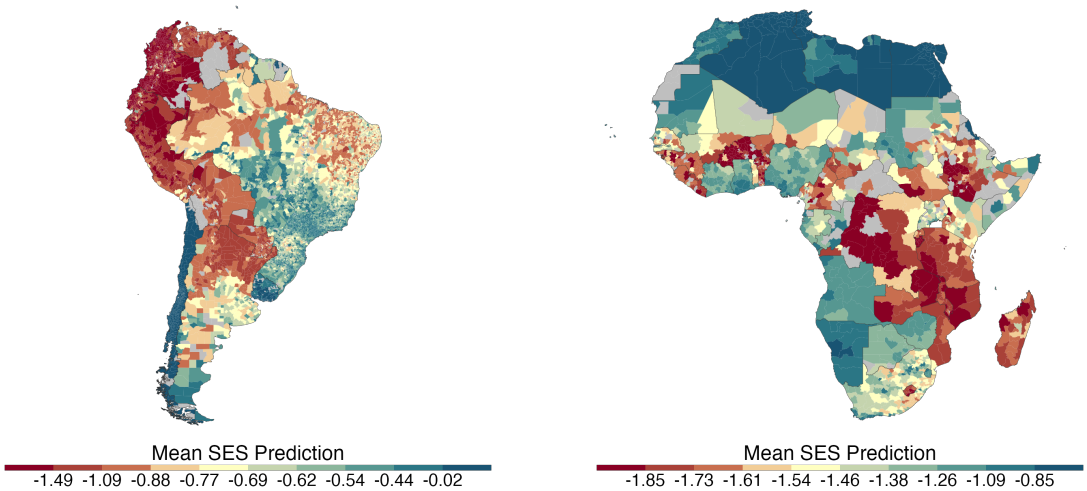
Fig. SI-8: Mean SES Prediction Around the World, Within-Region Scale



(a) United States **(b)** Canada



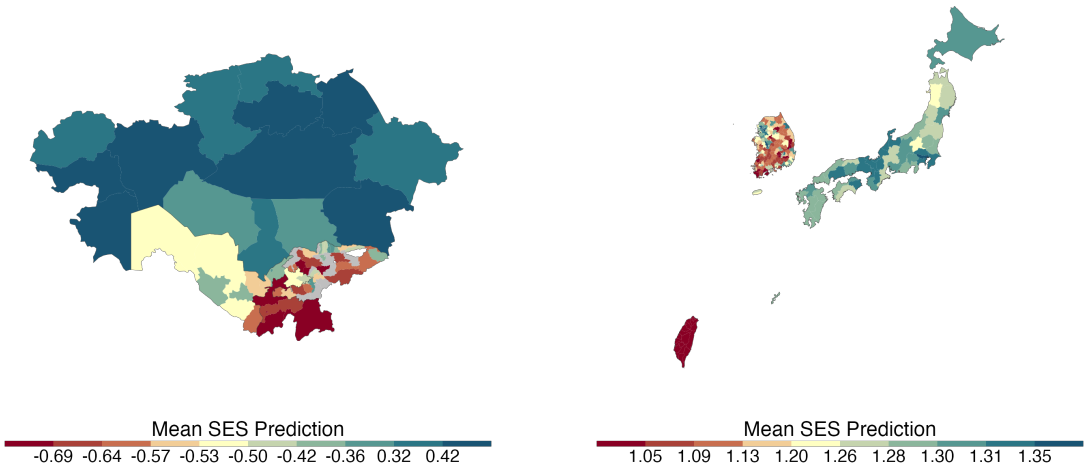
(c) Mexico and Central America **(d)** Europe



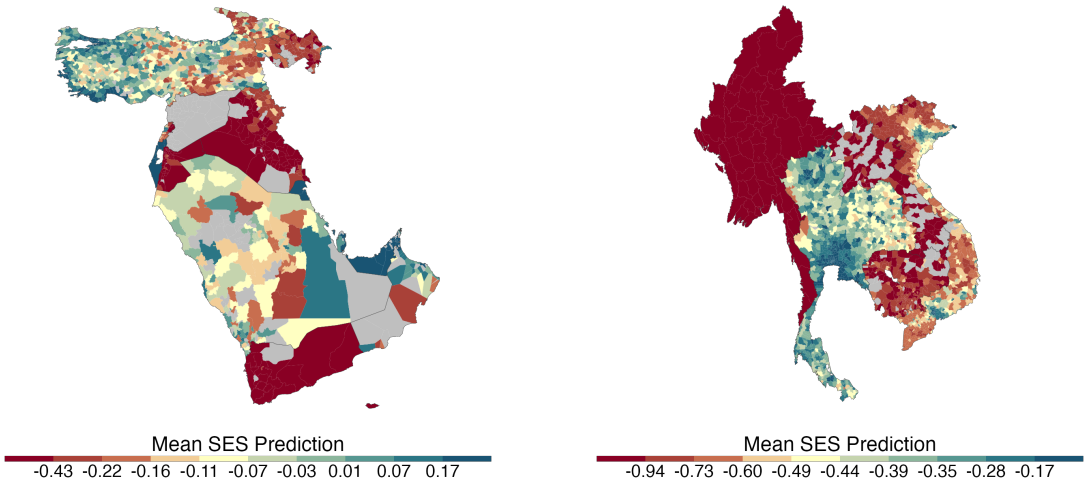
(e) South America **(f)** Africa

Note: The maps show average predicted SES by region using a region-specific color scale to highlight SES variation within that region. Low-population regions in gray.

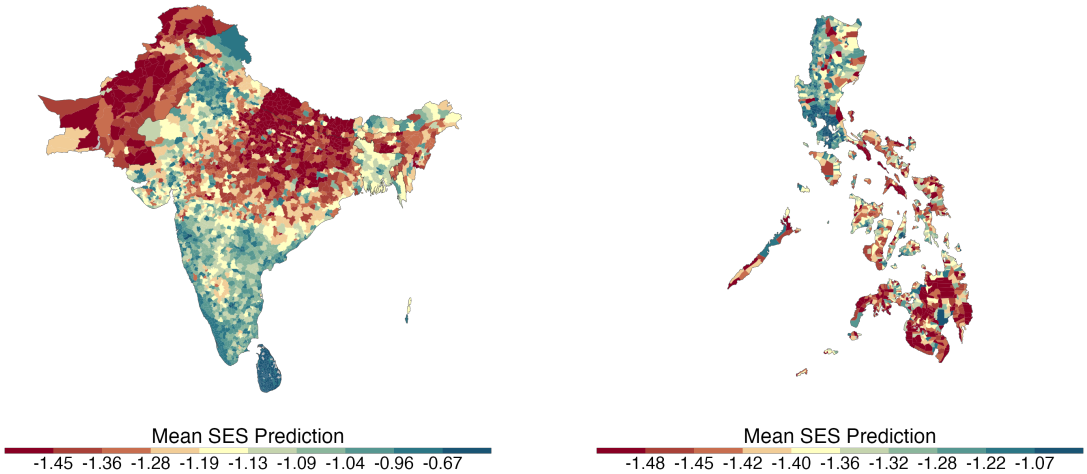
Fig. SI-8: Mean SES Prediction Around the World, Within-Region Scale



(g) Central Asia **(h)** East Asia



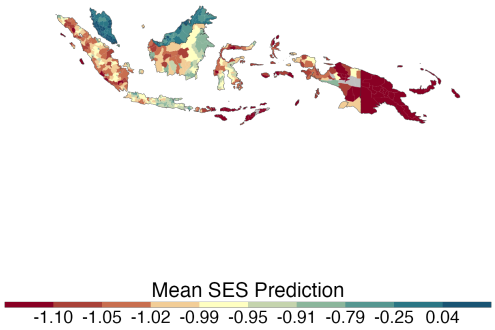
(i) West Asia **(j)** Southeast Asia



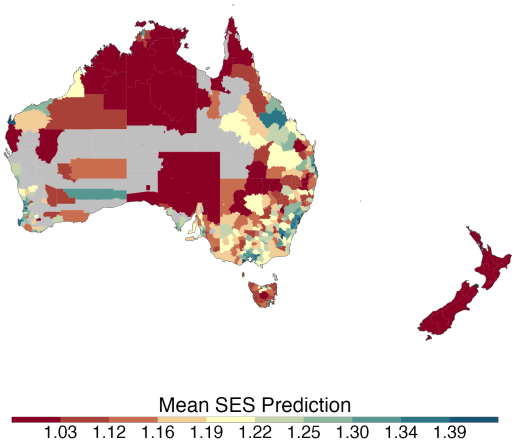
(k) South Asia **(l)** Philippines

Note: See page 49.

Fig. SI-8: Mean SES Prediction Around the World, Within-Region Scale



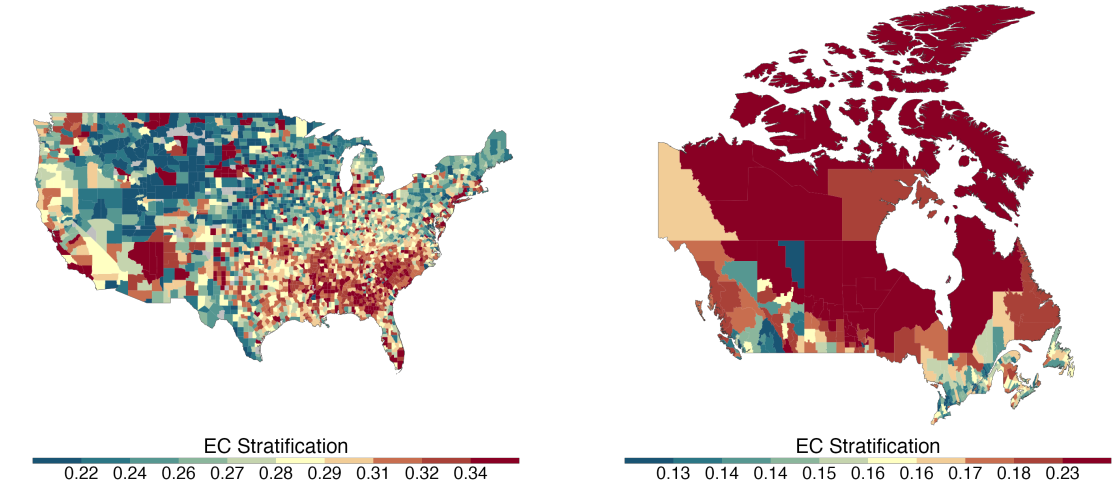
(m) Maritime Southeast Asia



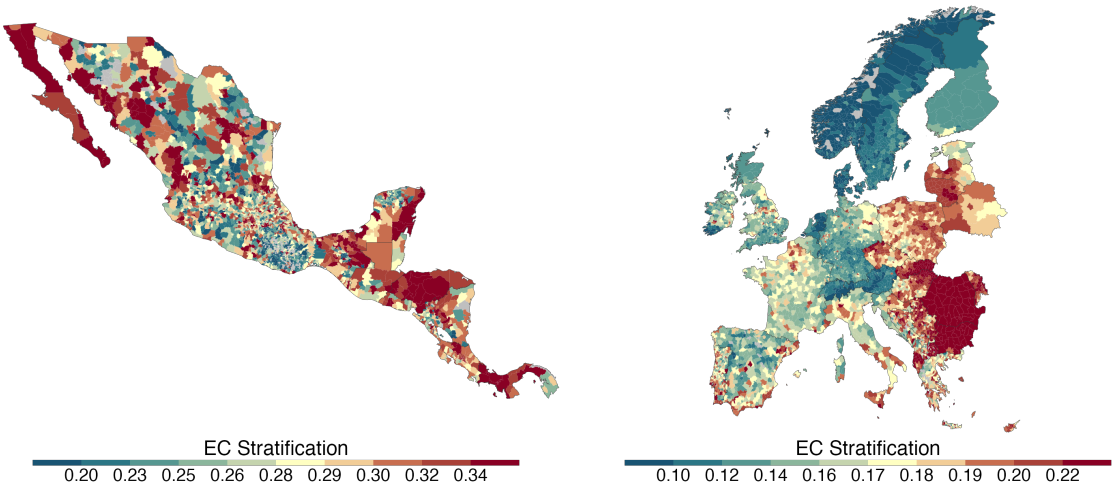
(n) Australia & New Zealand

Note: See page 49.

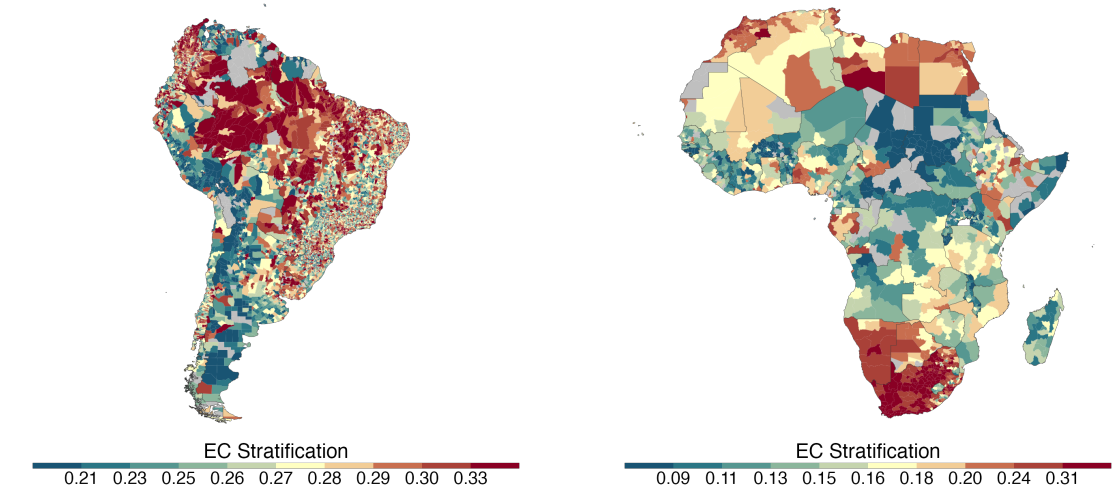
Fig. SI-9: EC Stratification Around the World, Within-Region Scale



(a) United States **(b)** Canada



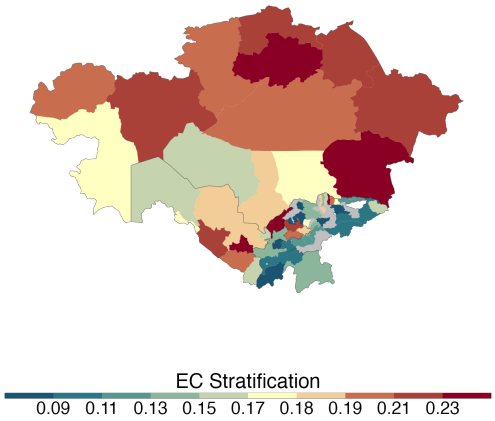
(c) Mexico and Central America **(d)** Europe



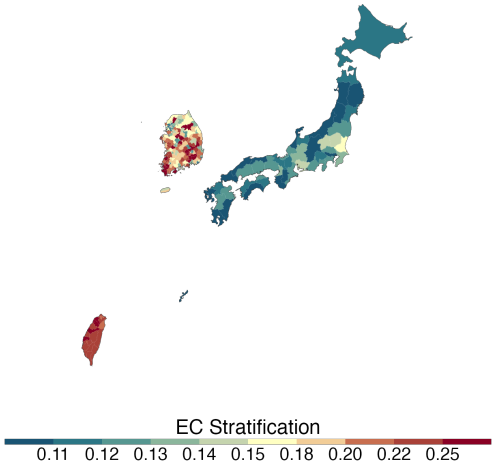
(e) South America **(f)** Africa

Note: The maps show EC stratification across GADM regions. EC stratification is defined as the regression slope of average friend (national) SES percentile versus individuals' own (national) SES percentile. Each map uses a region-specific color scale to highlight EC stratification variation within that region. Low-population regions in gray.

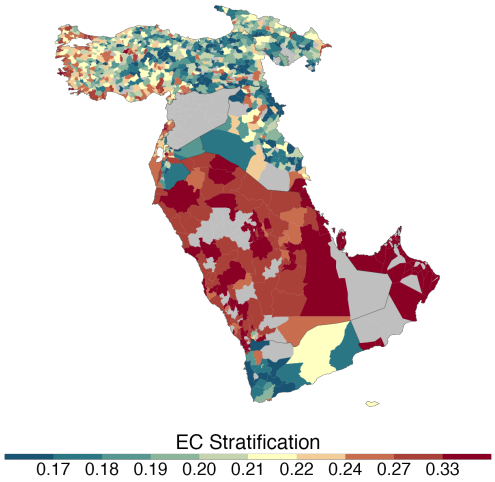
Fig. SI-9: EC Stratification Around the World, Within-Region Scale



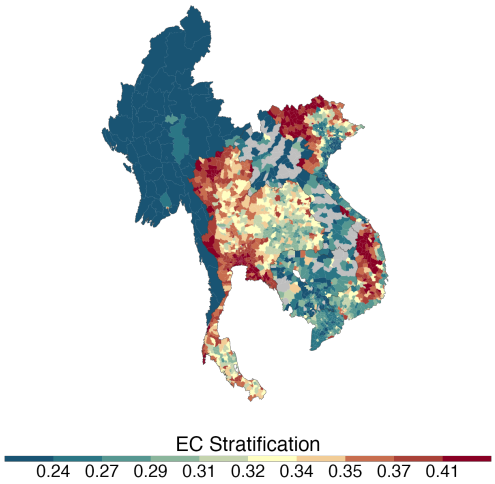
(g) Central Asia



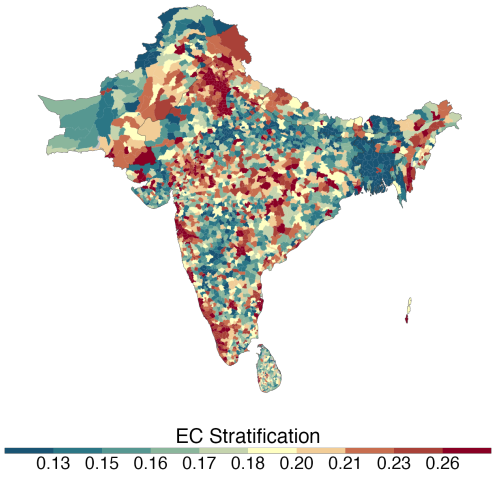
(h) East Asia



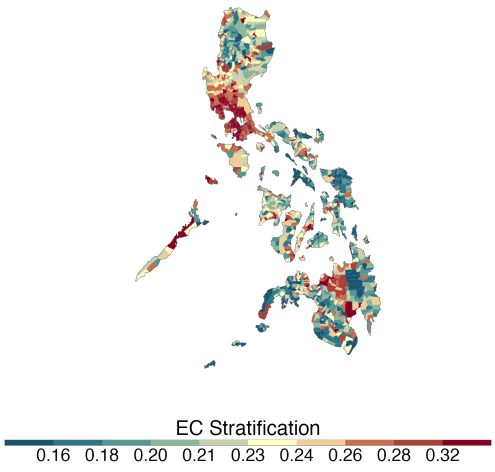
(i) West Asia



(j) Southeast Asia



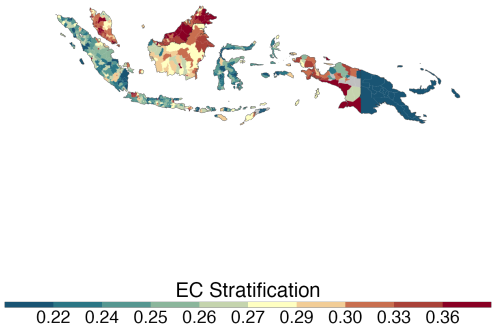
(k) South Asia



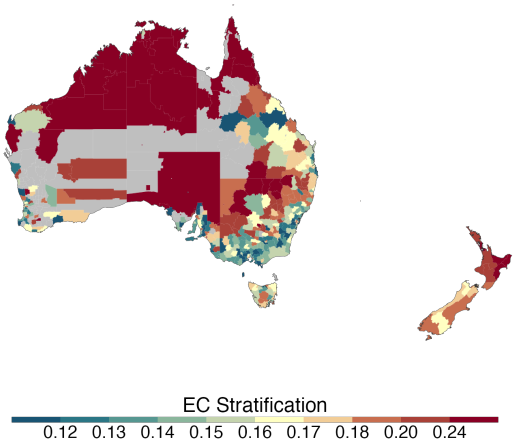
(l) Philippines

Note: See page 52.

Fig. SI-9: EC Stratification Around the World, Within-Region Scale



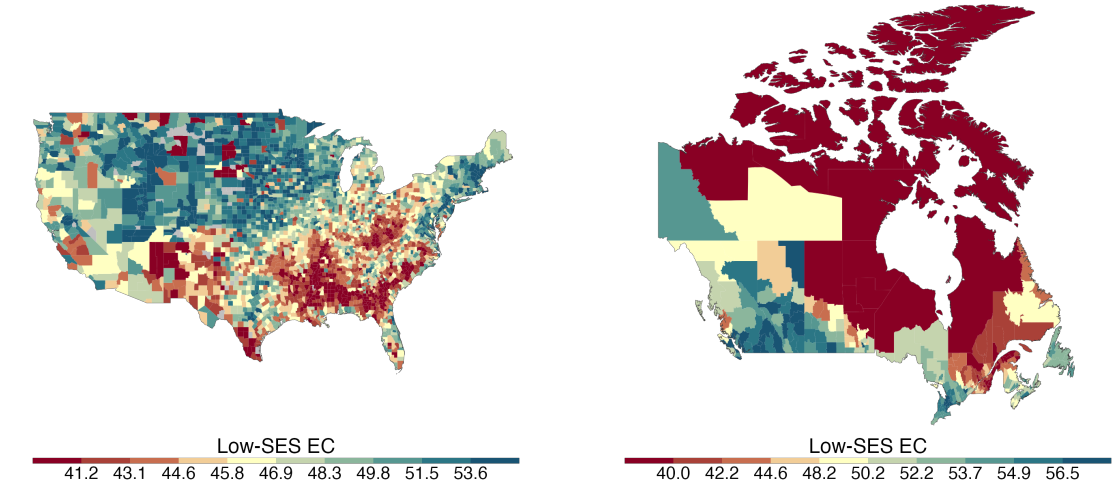
(m) Maritime Southeast Asia



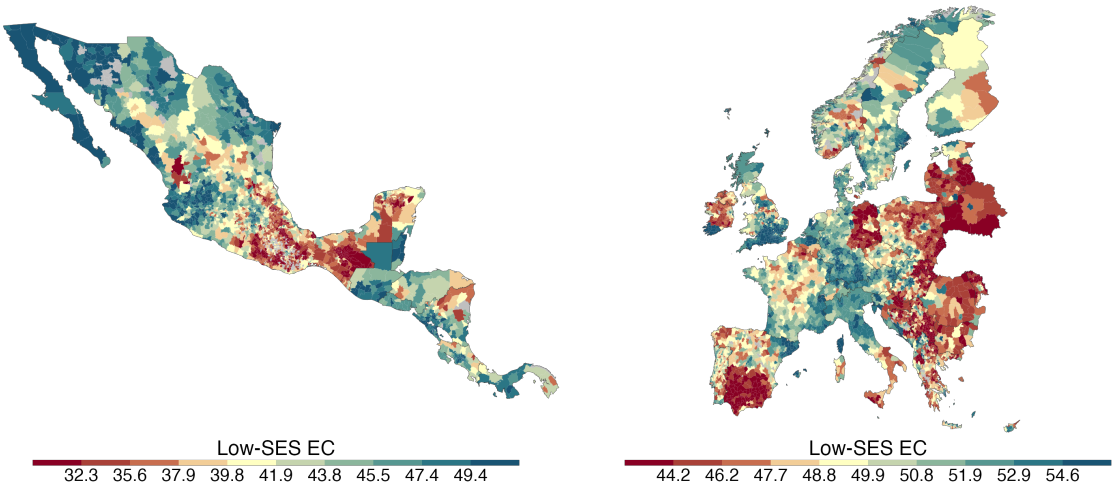
(n) Australia & New Zealand

Note: See page 52.

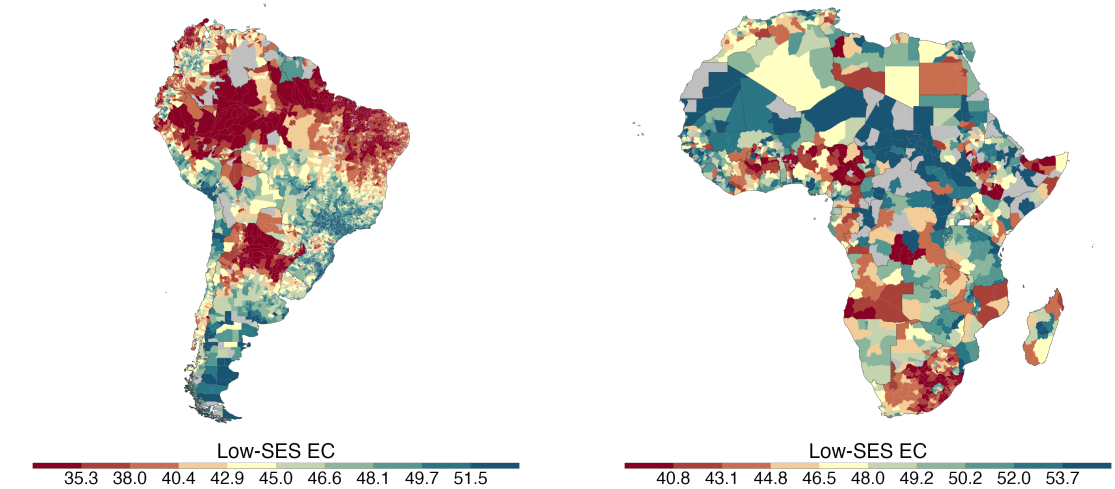
Fig. SI-10: Low-SES EC Around the World, Within-Region Scale



(a) United States **(b)** Canada



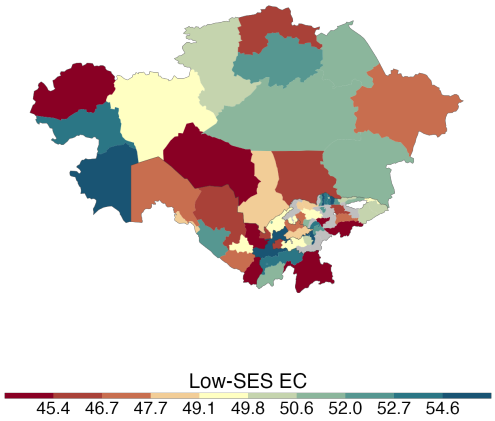
(c) Mexico and Central America **(d)** Europe



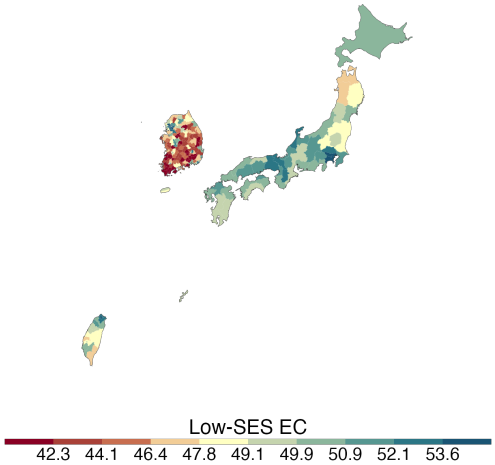
(e) South America **(f)** Africa

Note: The maps show Low-SES EC across GADM regions. Low-SES EC is defined as the expected average friend (national) SES percentile for an individual at the 25th (national) SES percentile. Each map uses a region-specific color scale to highlight Low-SES EC variation within that region. Low-population regions in gray.

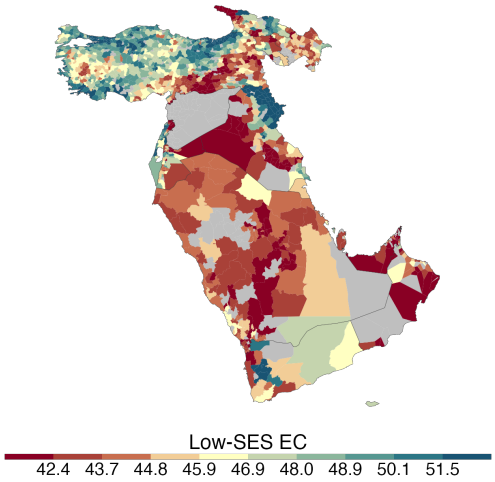
Fig. SI-10: Low-SES EC Around the World, Within-Region Scale



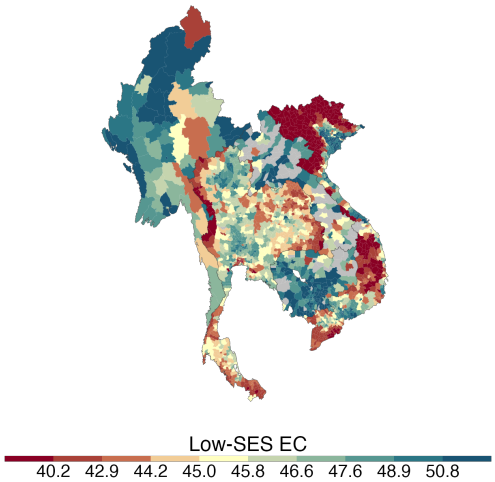
(g) Central Asia



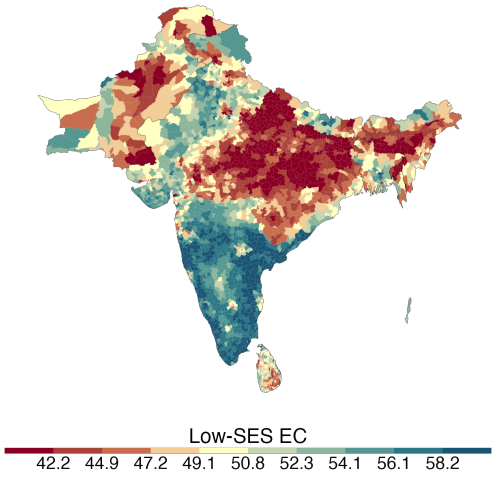
(h) East Asia



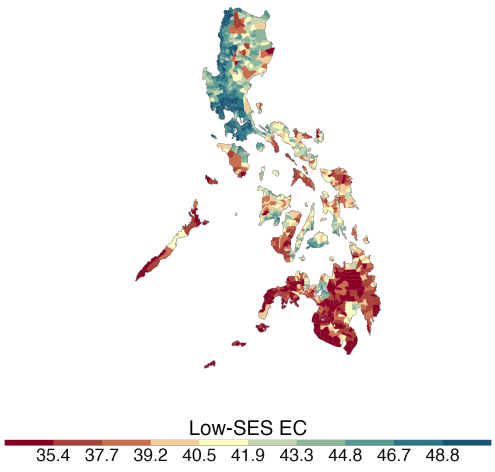
(i) West Asia



(j) Southeast Asia



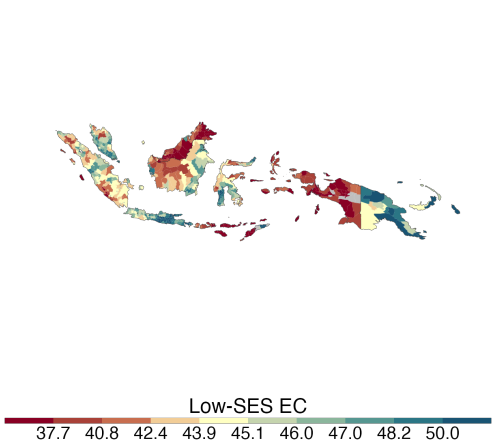
(k) South Asia



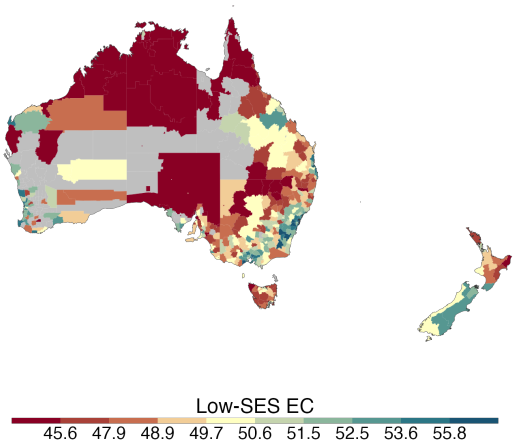
(l) Philippines

Note: See page 55.

Fig. SI-10: Low-SES EC Around the World, Within-Region Scale



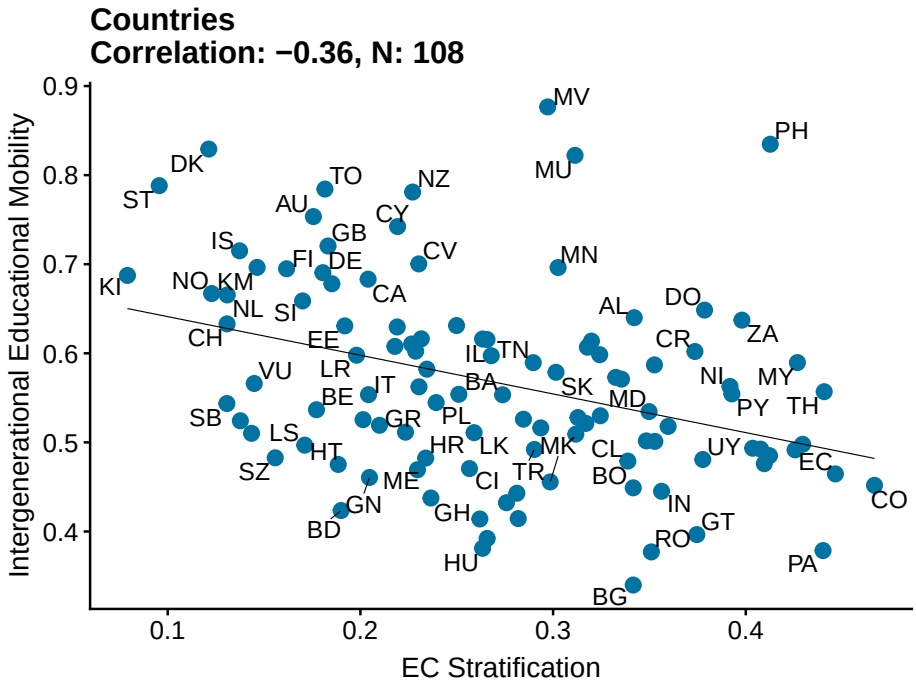
(m) Maritime Southeast Asia



(n) Australia & New Zealand

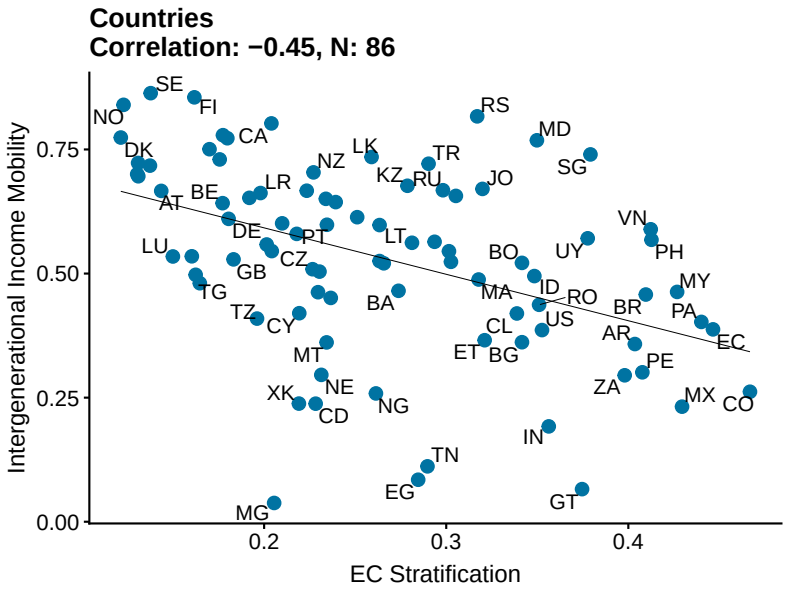
Note: See page 55.

Fig. SI-11: EC Stratification and Intergenerational Educational Mobility

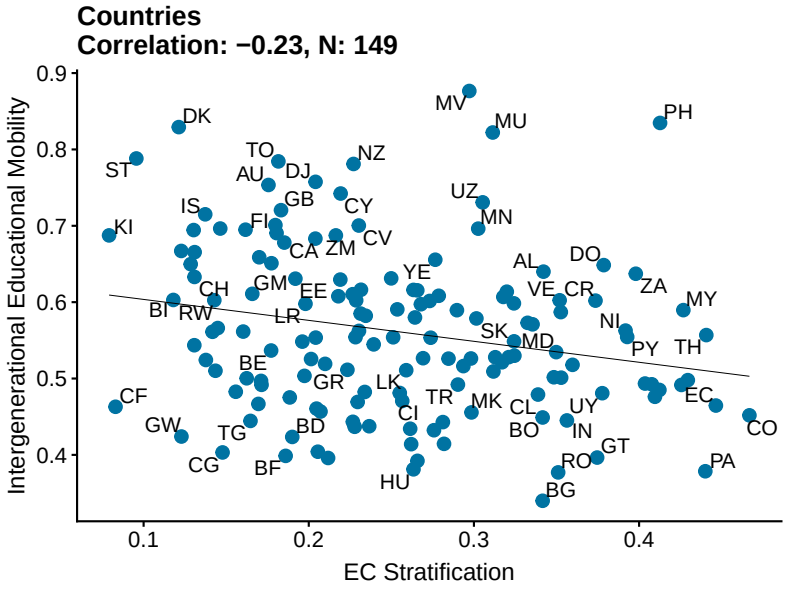


Note: This figure plots country-level estimates of intergenerational educational mobility against EC stratification. Intergenerational educational mobility is measured as 1 − the correlation coefficient between children's and parents' years of schooling, using data from the Global Database of Intergenerational Mobility (Muñoz et al., 2025). Higher values of EC stratification indicate greater social segregation, while higher values of educational mobility reflect weaker associations between parents' and children's educational attainment.

Fig. SI-12: EC Stratification and Intergenerational Mobility for All Available Countries



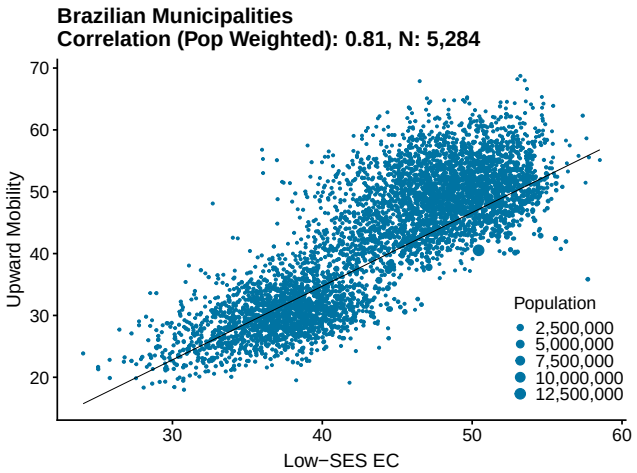
(a) Income Mobility



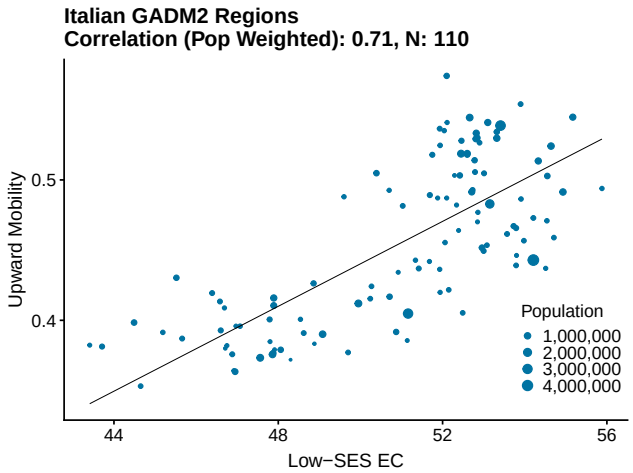
(b) Educational Mobility

Note: This figure plots estimates of intergenerational mobility against EC stratification for all countries (with data quality green, orange, and red). Intergenerational income mobility is measured as 1 − the intergenerational income elasticity. Intergenerational educational mobility is measured as 1 − the correlation coefficient between children's and parents' years of schooling. Both measures use data from the Global Database of Intergenerational Mobility.

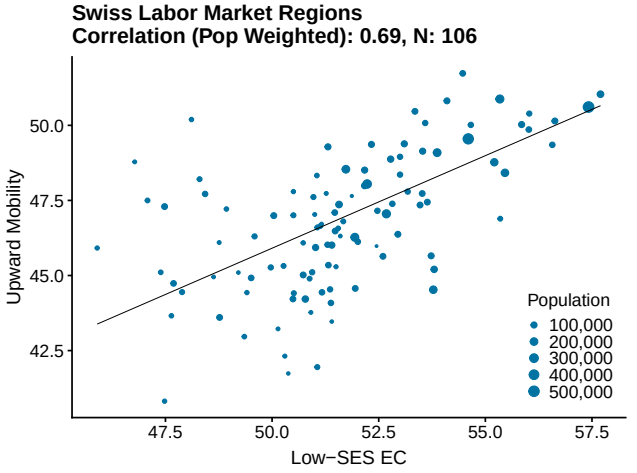
Fig. SI-13: Low-SES EC and Intergenerational Income Mobility, Subnational



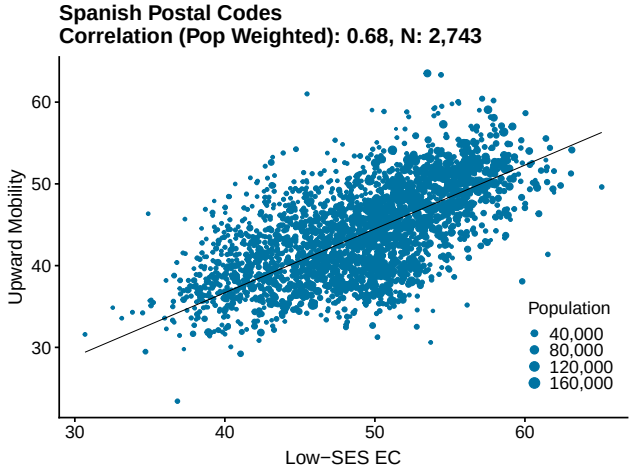
(a) Brazilian Municipalities



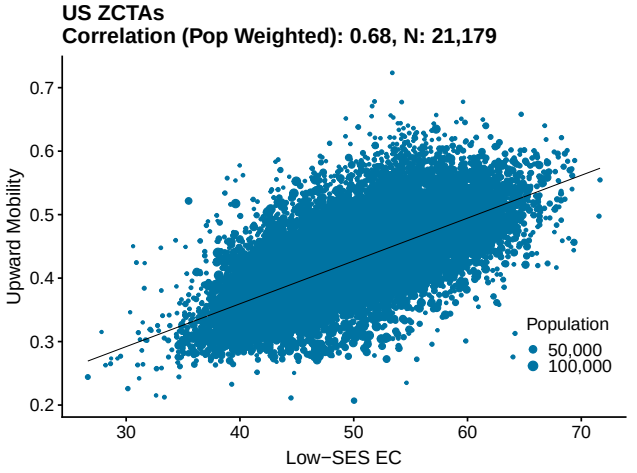
(b) Italian GADM2 Regions



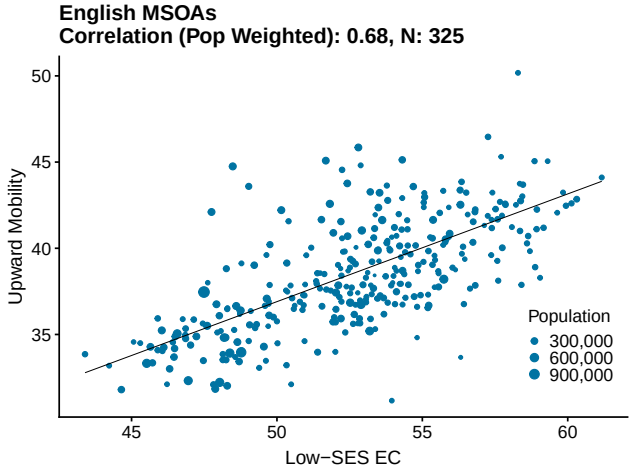
(c) Swiss Labor Market Regions



(d) Spanish Postal Codes



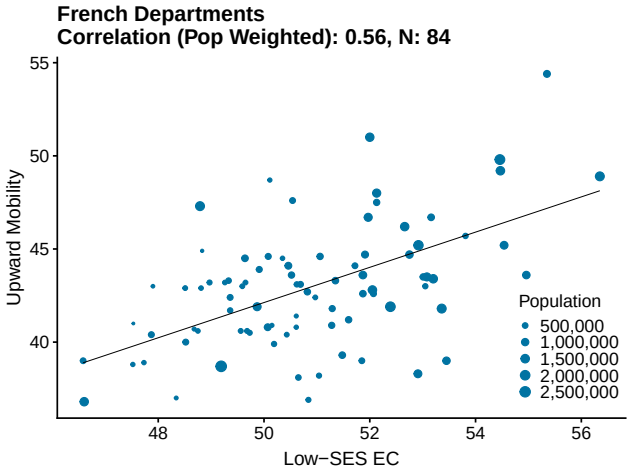
(e) US ZCTAs



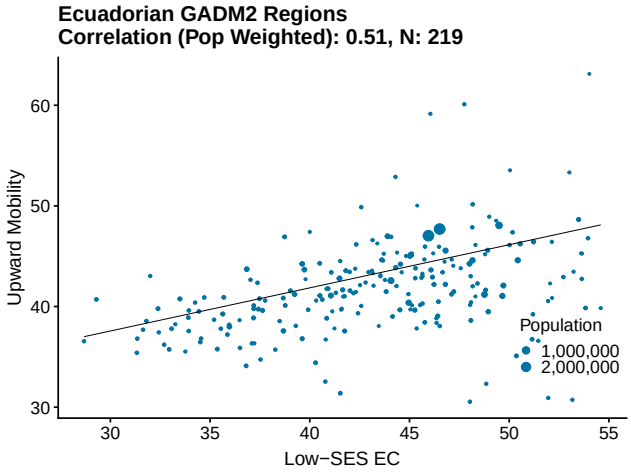
(f) English MSOAs

Note: Each scatter plot displays the relationship between low-SES EC and region-level intergenerational income mobility within a given country. Regional intergenerational mobility measures are drawn from recent administrative and survey-based studies, as documented in the Methods and Supplementary Information sections.

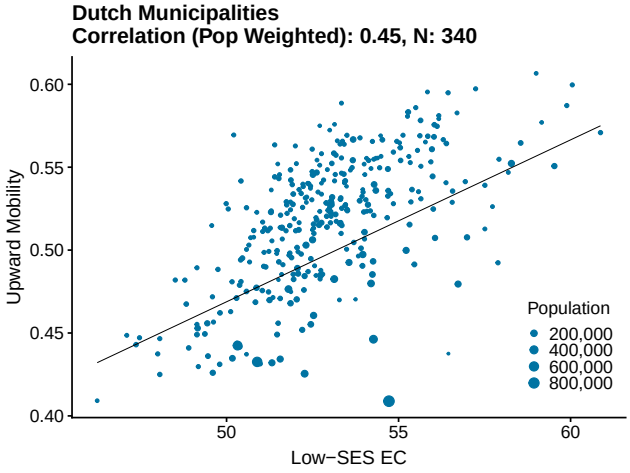
Fig. SI-13: Low-SES EC and Intergenerational Income Mobility, Subnational



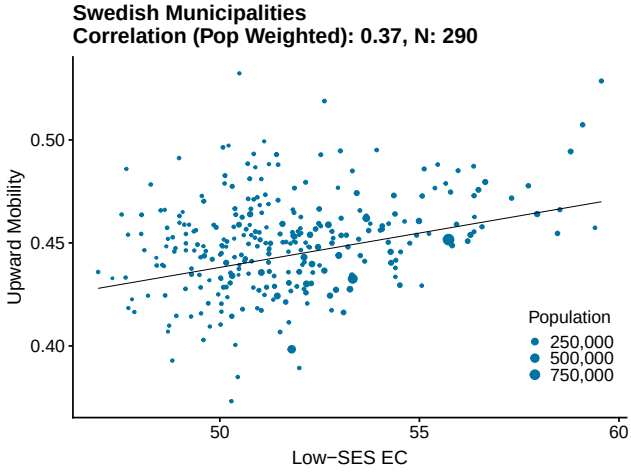
(g) French Departments



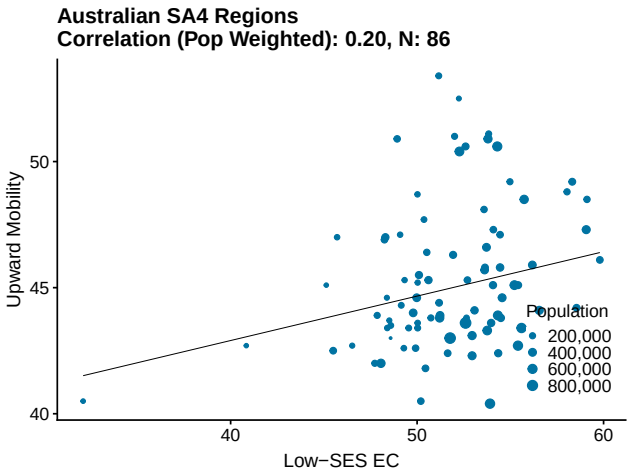
(h) Ecuadorian GADM1 Regions



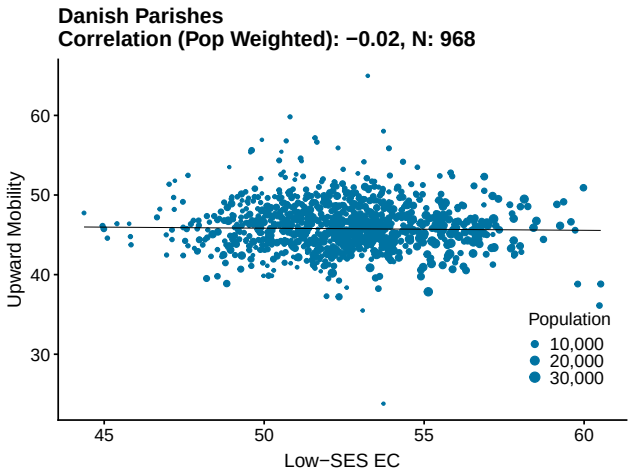
(i) Dutch Municipalities



(j) Swedish Municipalities



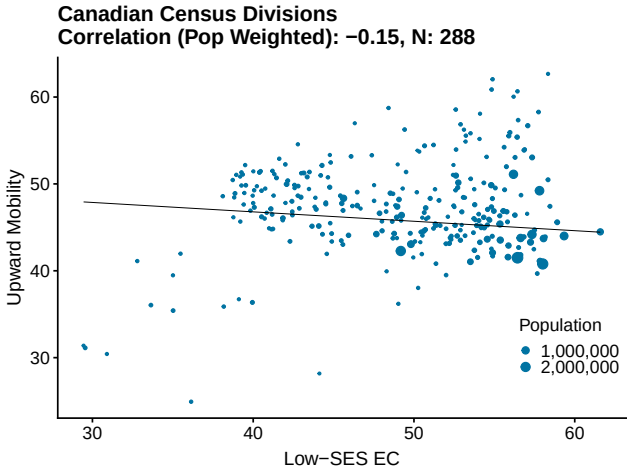
(k) Australian SA4 Regions



(l) Danish Parishes

Note: See page 60.

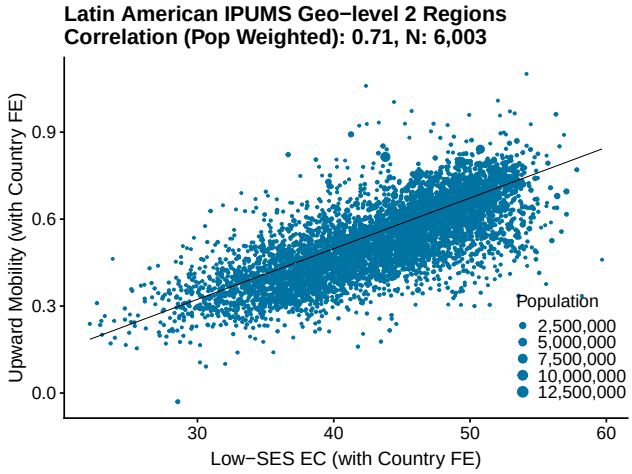
Fig. SI-13: Low-SES EC and Intergenerational Income Mobility, Subnational



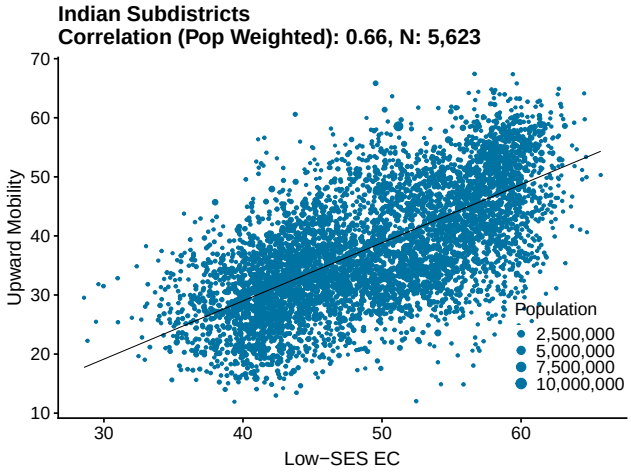
(m) Canadian Census Divisions

Note: See page 60.

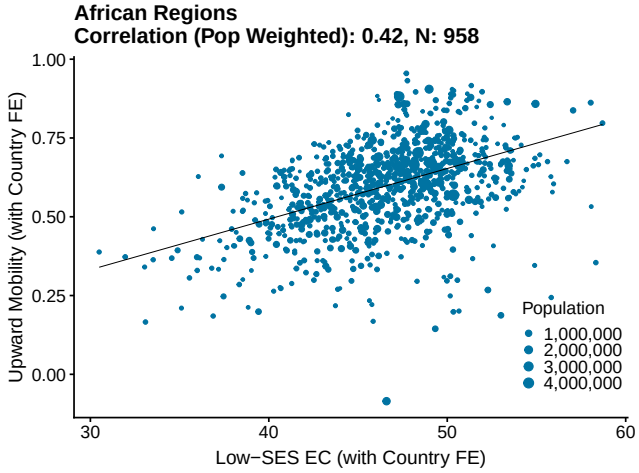
Fig. SI-14: Low-SES EC and Intergenerational Educational Mobility, Subnational



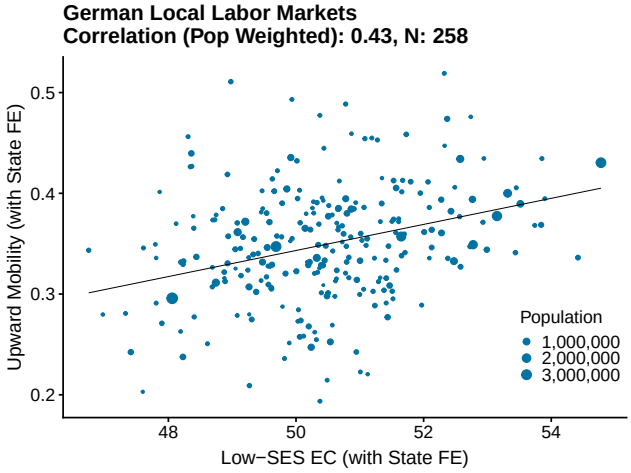
(a) Latin American IPUMS Geo-level 2 Regions



(b) Indian Subdistricts



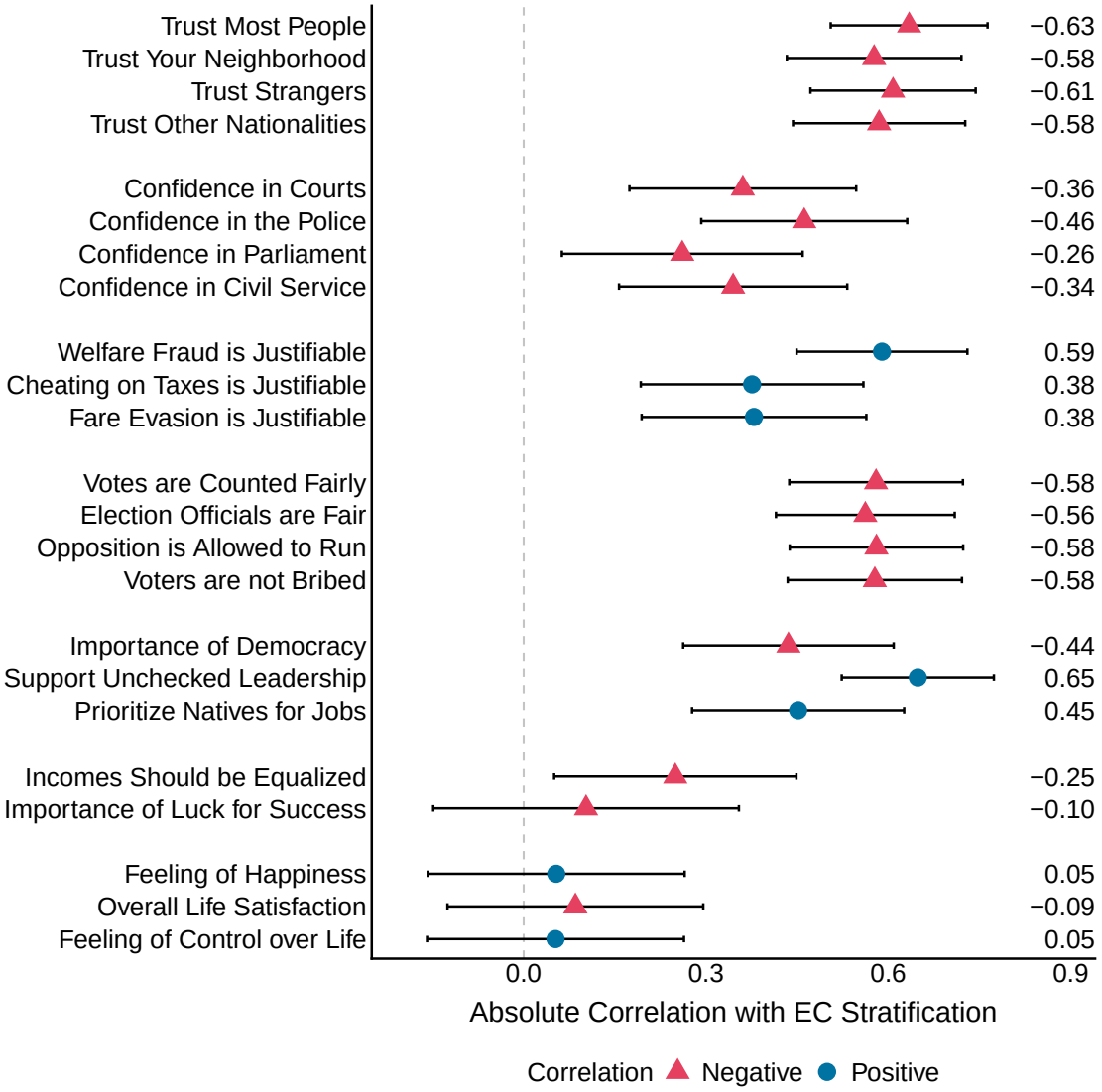
(c) African Regions



(d) German Local Labor Market Regions

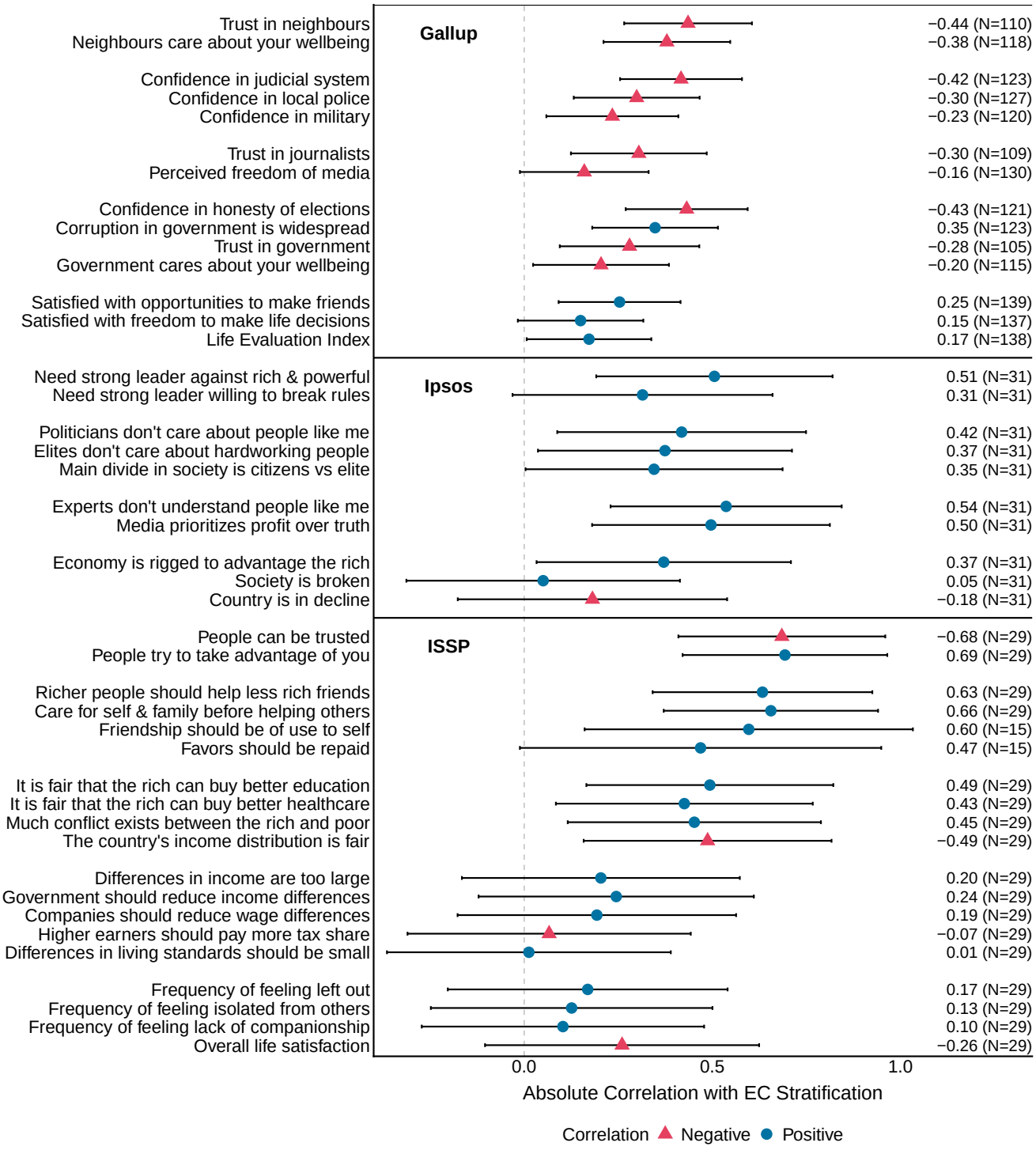
Note: Each scatter plot displays the relationship between low-SES EC and region-level intergenerational educational mobility within a given geographic context: (a) Latin American regions (IPUMS Geo-level 2 regions), (b) Indian subdistricts, (c) African regions, and (d) German local labor market regions. Regional intergenerational mobility measures are drawn from recent administrative and survey-based studies, as documented in the Methods and Supplementary Information sections. Panels (a) and (c) include country fixed effects, and panel (d) includes state fixed effects.

Fig. SI-15: EC Stratification and Joint EVS/WVS Outcomes for All Available Countries



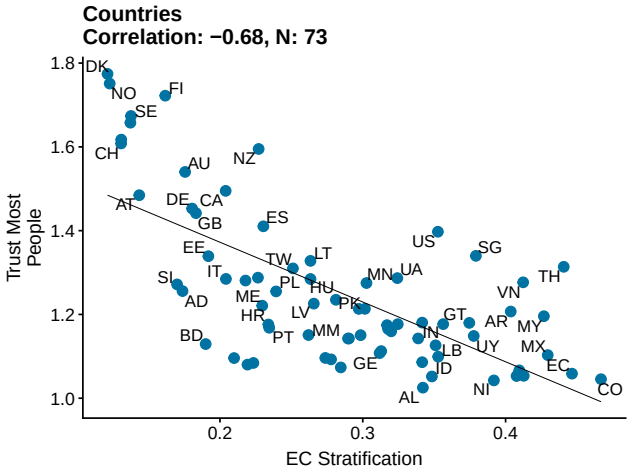
Note: This figure plots the absolute value of the correlation coefficient between EC stratification and country-level outcomes from the Joint European Values Study/World Values Survey. The sample includes all available countries (with data quality green, orange, and red). Correlations are shown in absolute value to facilitate comparison of magnitudes. The direction of the relationship is indicated by the marker: blue circles denote positive correlations and pink triangles denote negative correlations. Error bars represent 95% confidence intervals.

Fig. SI-16: EC Stratification and Gallup, Ipsos, and ISSP Outcomes for All Available Countries

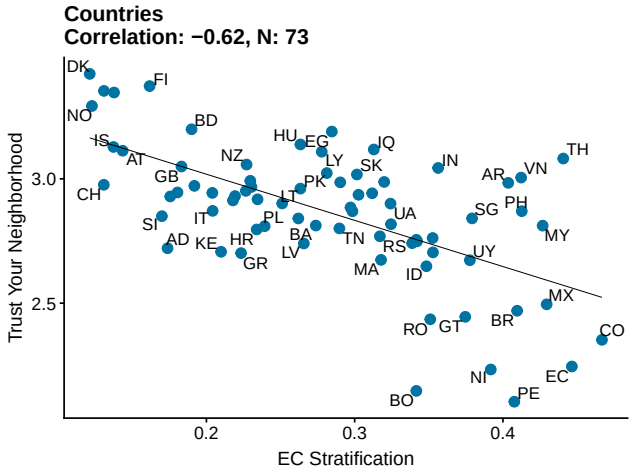


Note: This figure plots the absolute value of the correlation coefficient between EC stratification and country-level outcomes from three survey sources: the Ipsos Populism Report, the International Social Survey Programme (ISSP), and the Gallup World Poll. The sample includes all available countries (with data quality green, orange, and red). Correlations are shown in absolute value to facilitate comparison of magnitudes. The direction of the relationship is indicated by marker color: blue circles denote positive correlations and pink triangles denote negative correlations. Error bars represent 95% confidence intervals.

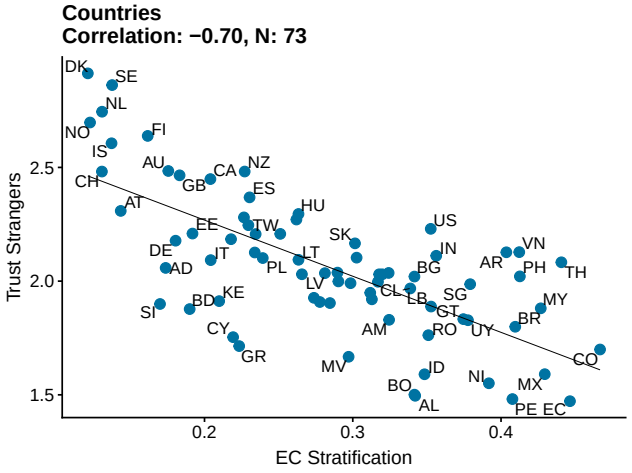
Fig. SI-17: EC Stratification and Socio-Political Outcomes from the Joint EVS/WVS



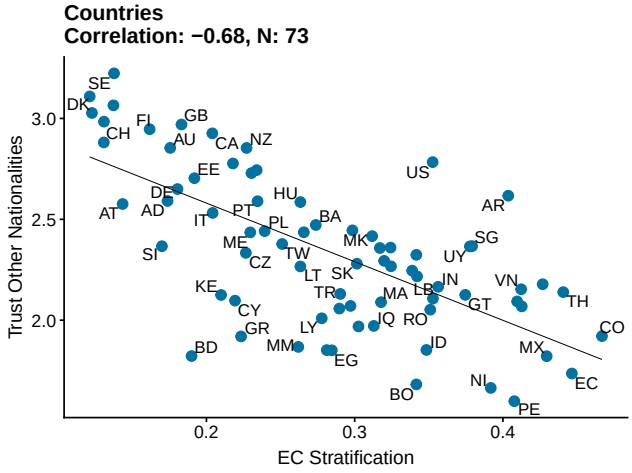
(a) Trust Most People



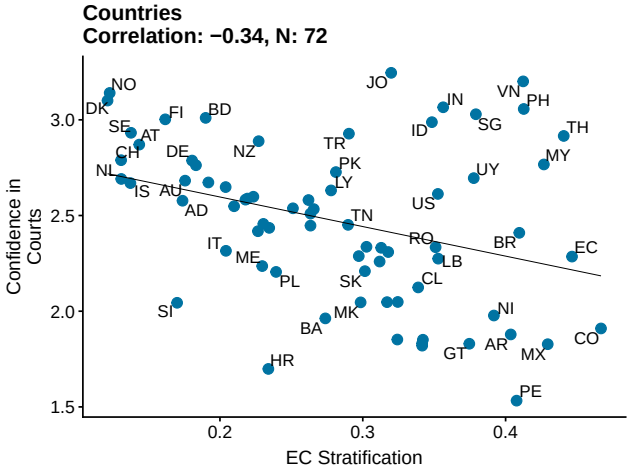
(b) Trust Your Neighborhood



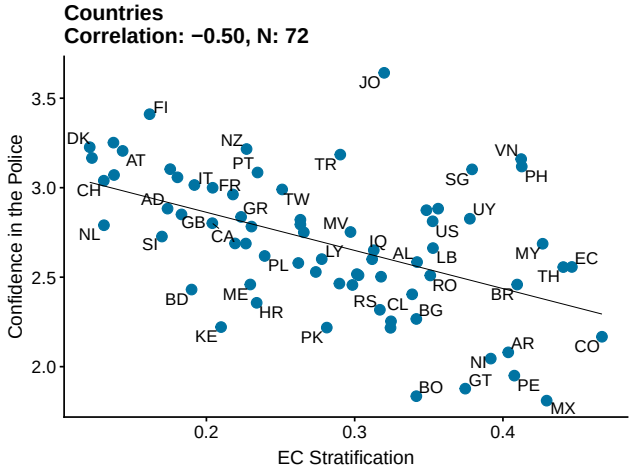
(c) Trust Strangers



(d) Trust Other Nationalities



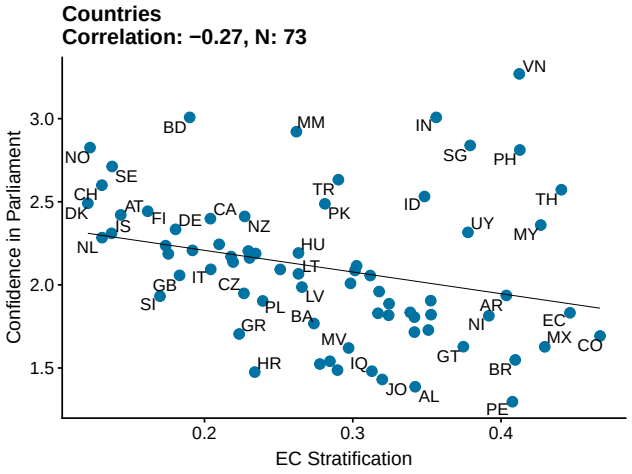
(e) Confidence in Courts



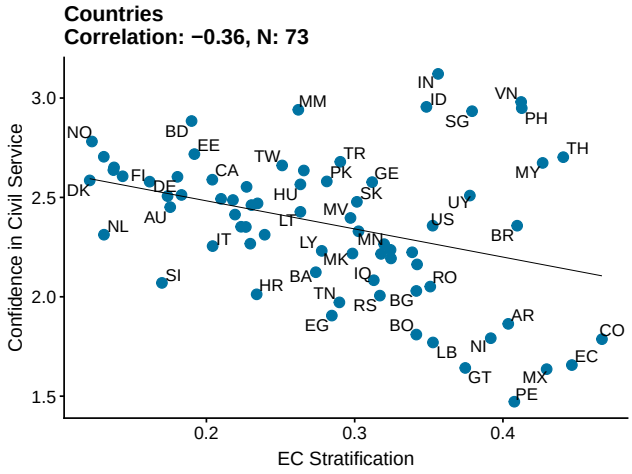
(f) Confidence in the Police

Note: Each scatter plot displays the relationship between EC stratification and country-level measures of various socio-political outcomes. Data on outcomes are drawn from the Joint European Values Study/World Values Survey.

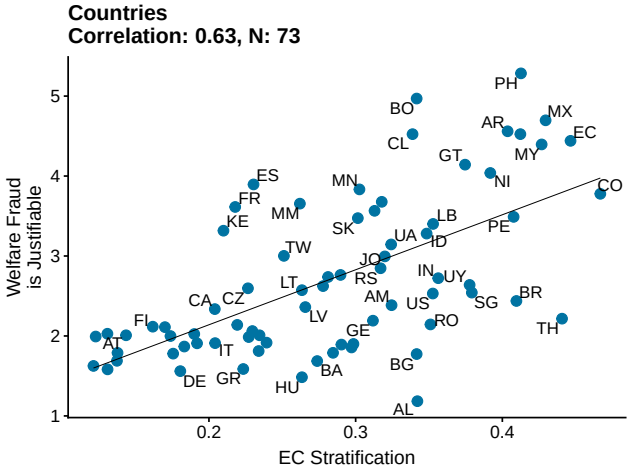
Fig. SI-17: EC Stratification and Socio-Political Outcomes from the Joint EVS/WVS



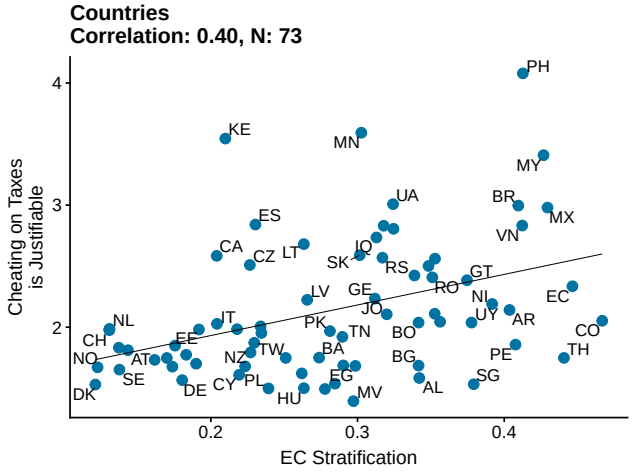
(g) Confidence in Parliament



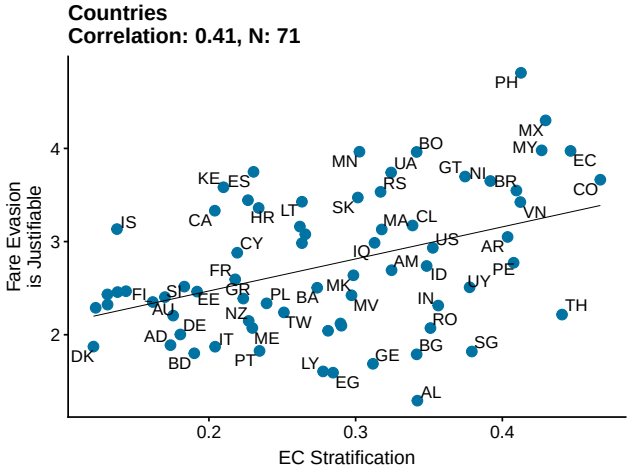
(h) Confidence in Civil Service



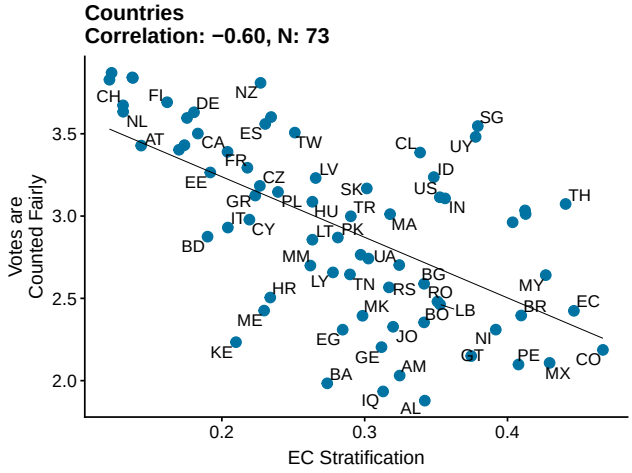
(i) Welfare Fraud is Justifiable



(j) Cheating on Taxes is Justifiable



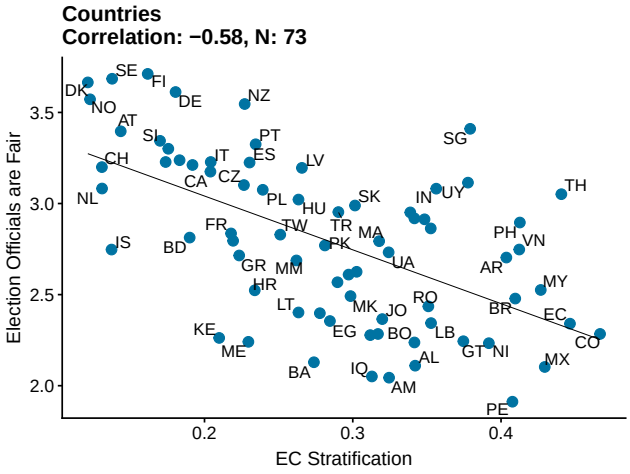
(k) Fare Evasion is Justifiable



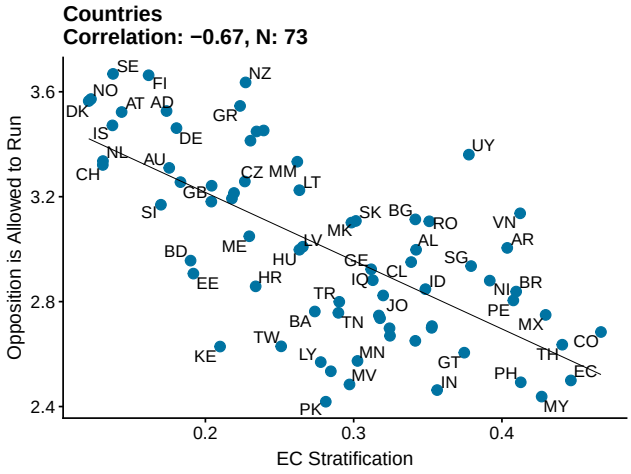
(l) Votes are Counted Fairly

Note: See page 66.

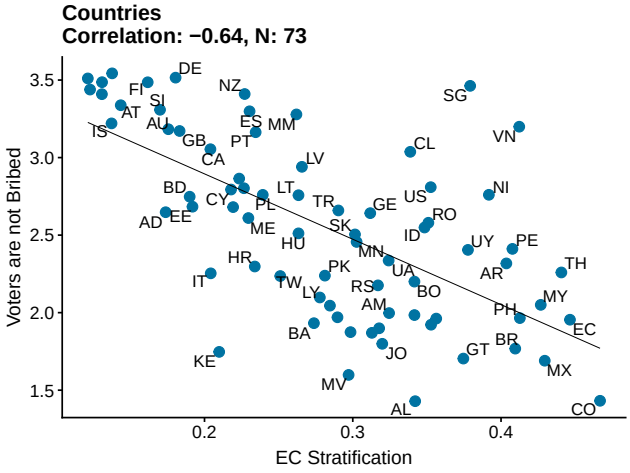
Fig. SI-17: EC Stratification and Socio-Political Outcomes from the Joint EVS/WVS



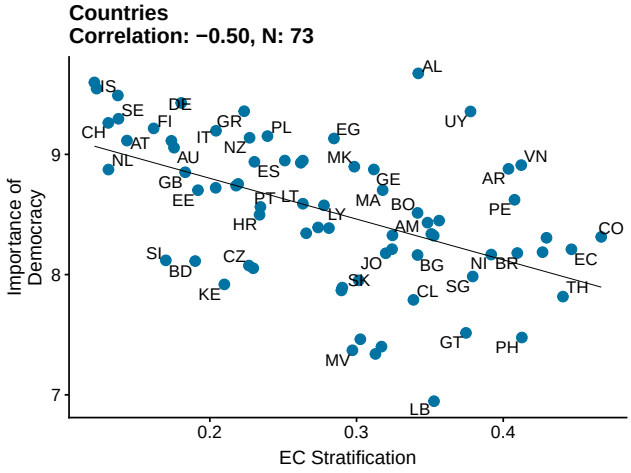
(m) Election Officials are Fair



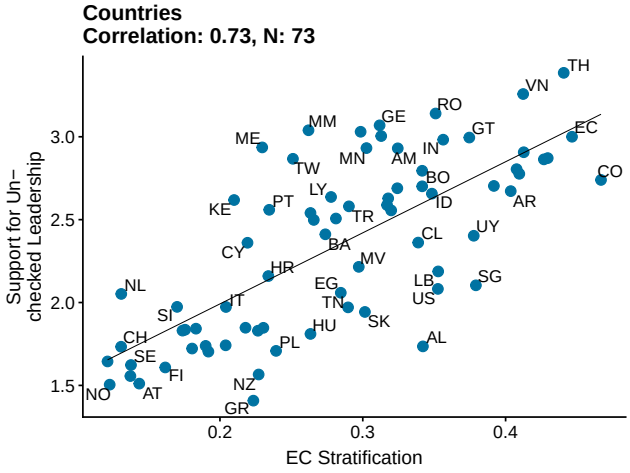
(n) Opposition is Allowed to Run



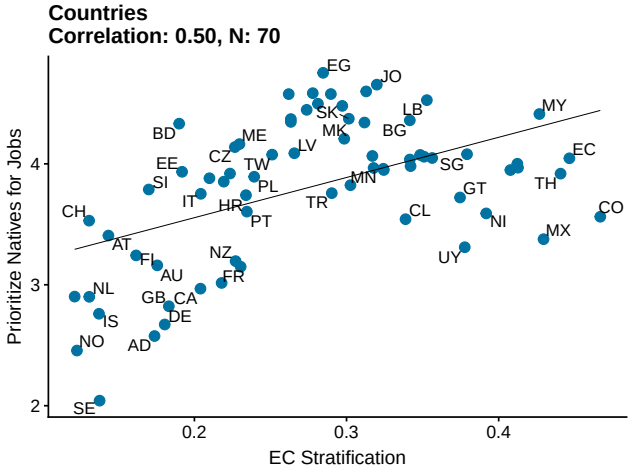
(o) Voters are not Bribed



(p) Importance of Democracy



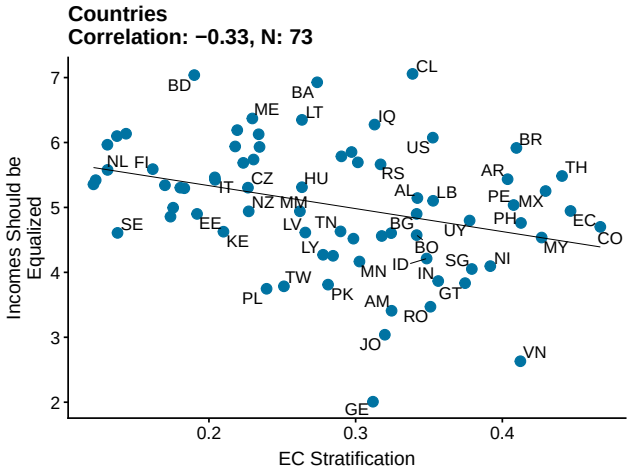
(q) Support for Strong Leadership



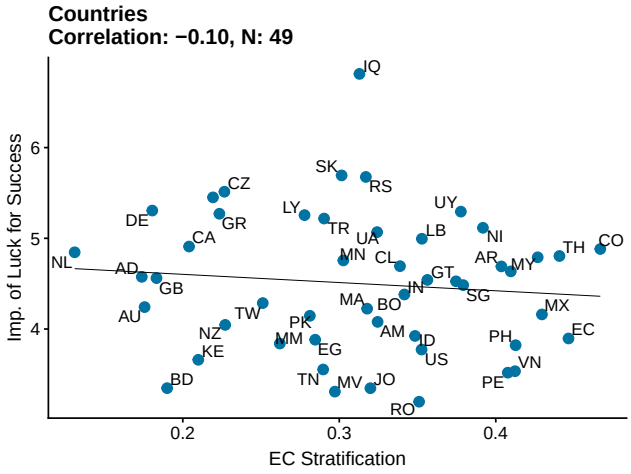
(r) Prioritize Natives for Jobs

Note: See page 66.

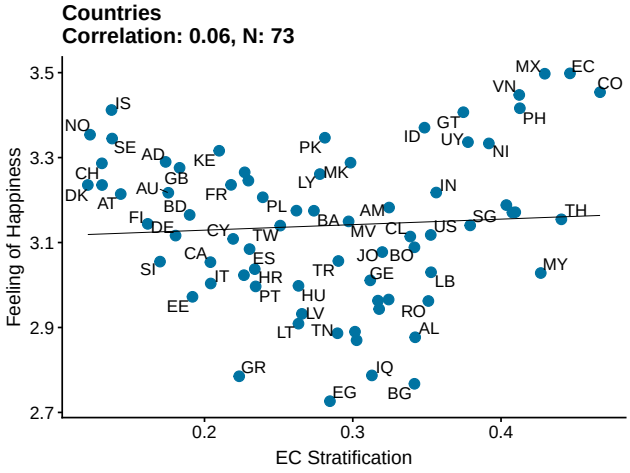
Fig. SI-17: EC Stratification and Socio-Political Outcomes from the Joint EVS/WVS



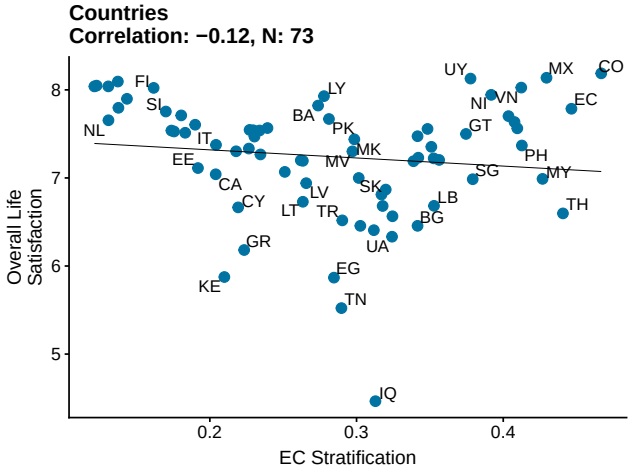
(s) Incomes Should be Equalized



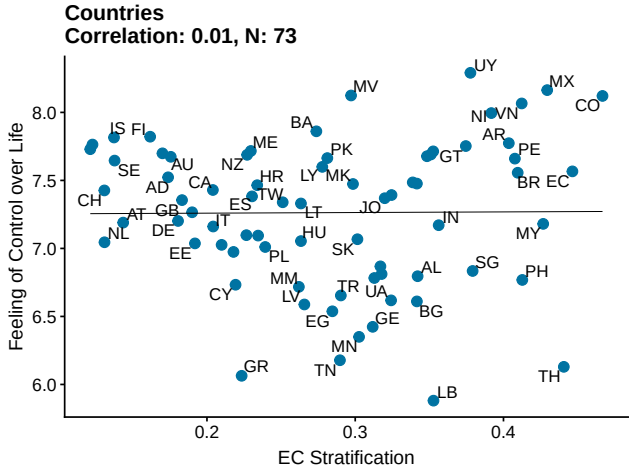
(t) Importance of Luck for Success



(u) Feeling of Happiness



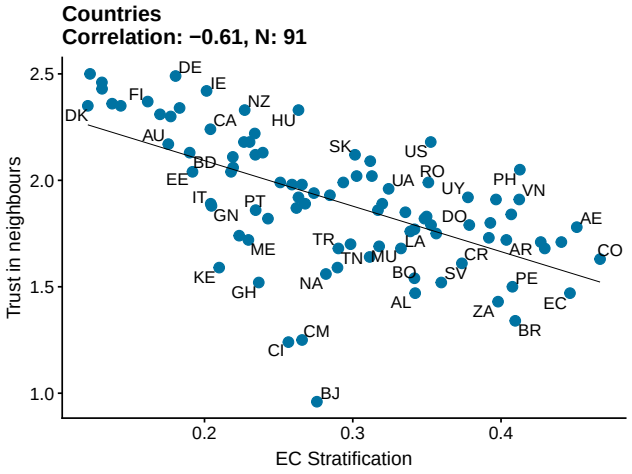
(v) Overall Life Satisfaction



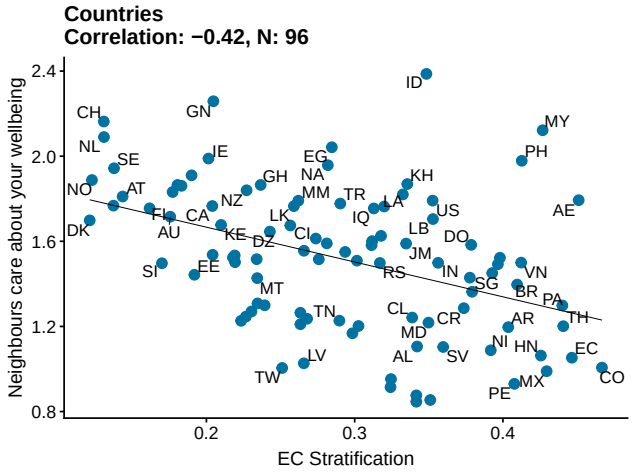
(w) Feeling of Control over Life

Note: See page 66.

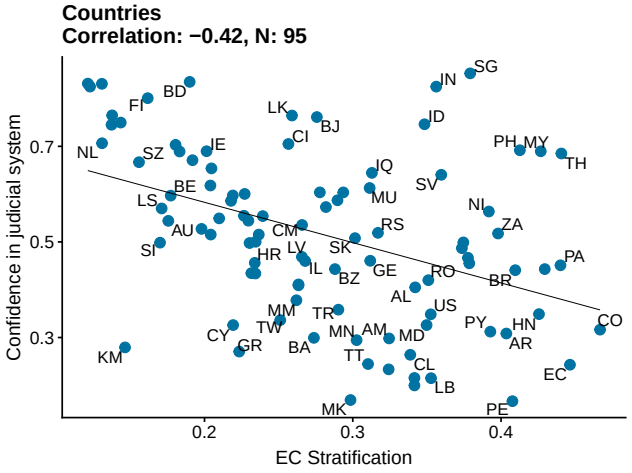
Fig. SI-18: EC Stratification and Socio-Political Outcomes from the Gallup World Poll



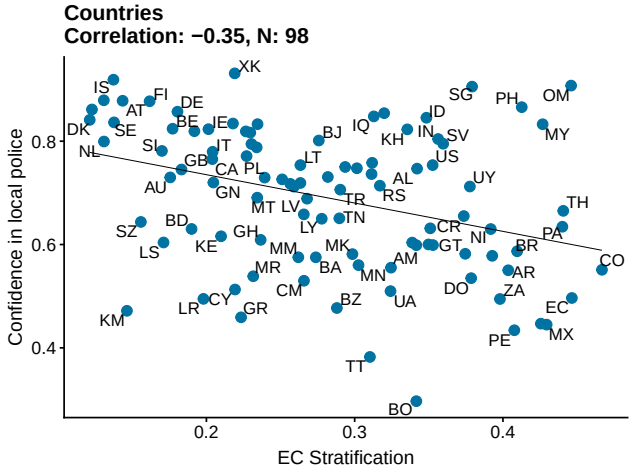
(a) Trust in neighbours



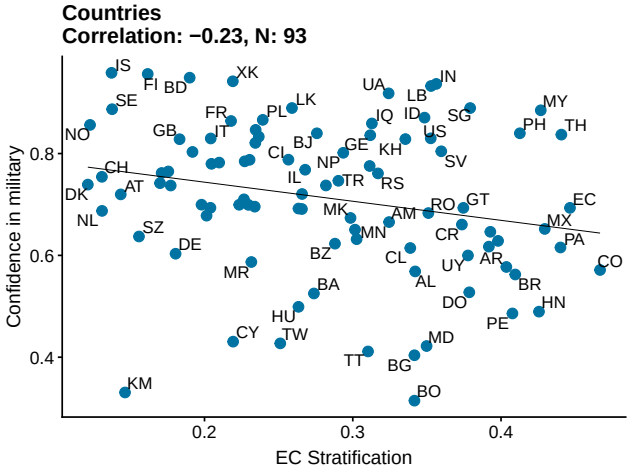
(b) Neighbours care about your wellbeing



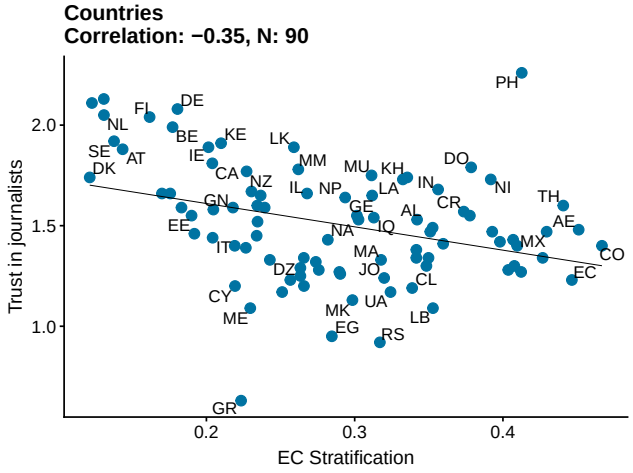
(c) Confidence in judicial system



(d) Confidence in local police



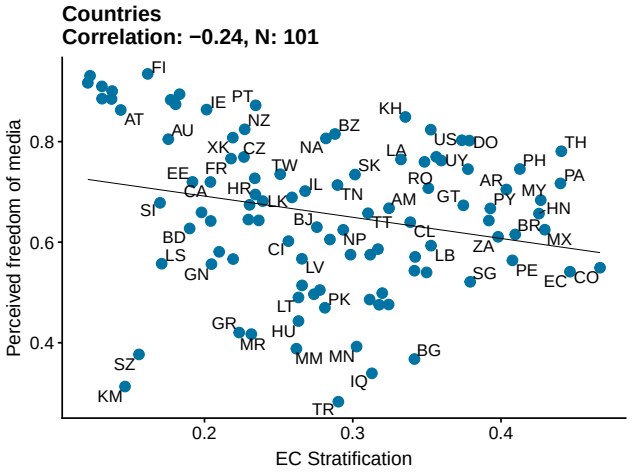
(e) Confidence in military



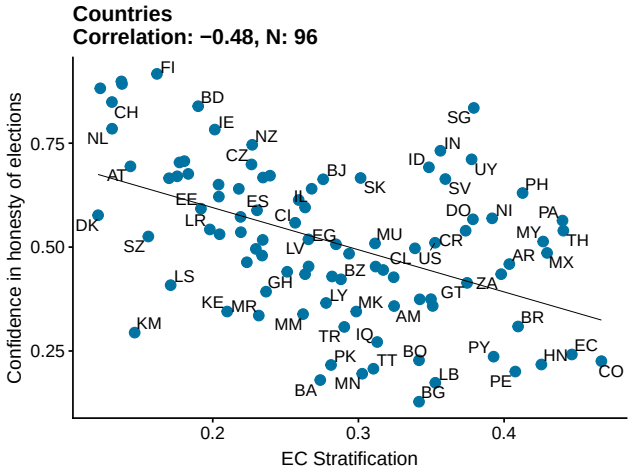
(f) Trust in journalists

Note: Each scatter plot displays the relationship between EC stratification and country-level measures of various socio-political outcomes. Data on outcomes are drawn from the Gallup World Poll.

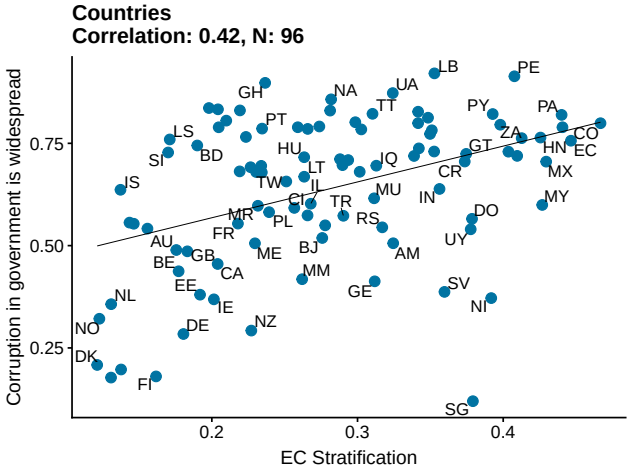
Fig. SI-18: EC Stratification and Socio-Political Outcomes from the Gallup World Poll



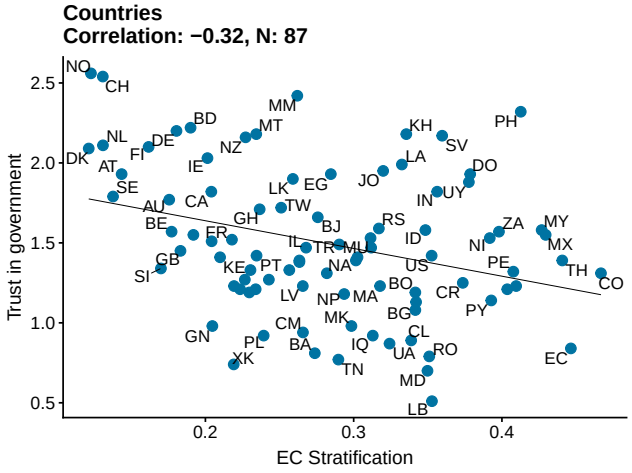
(g) Perceived freedom of media



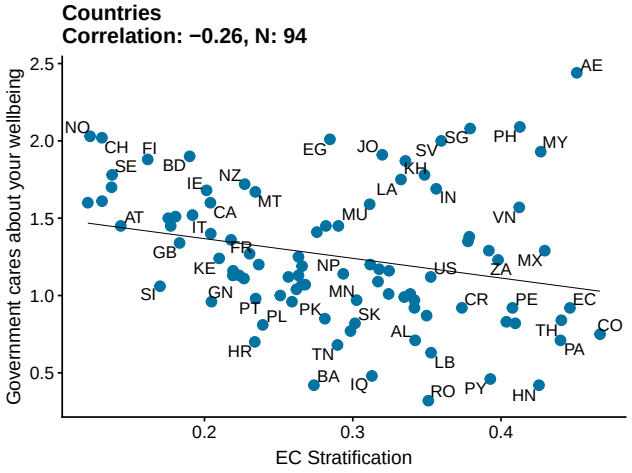
(h) Confidence in honesty of elections



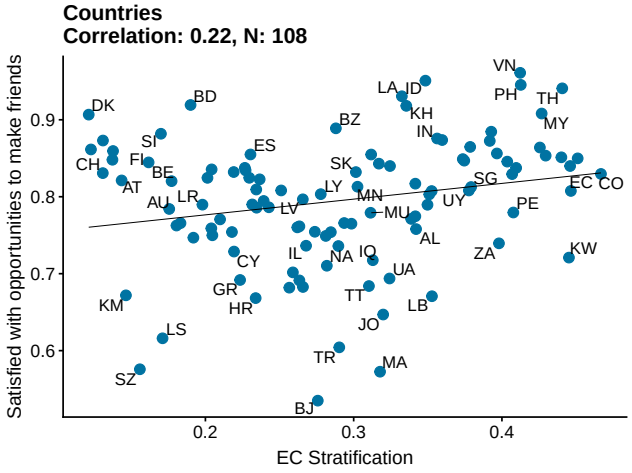
(i) Corruption in government is widespread



(j) Trust in government



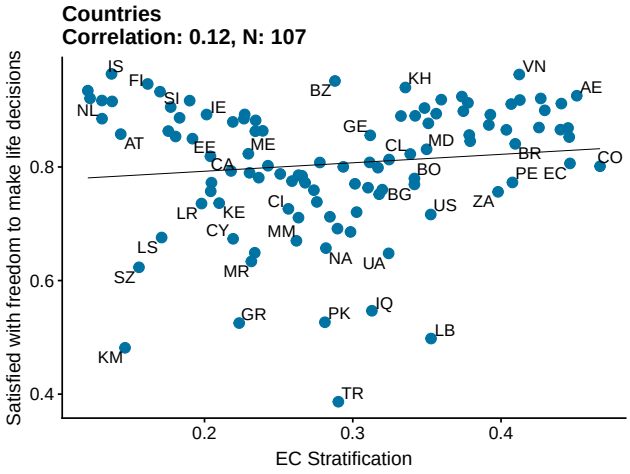
(k) Government cares about your wellbeing



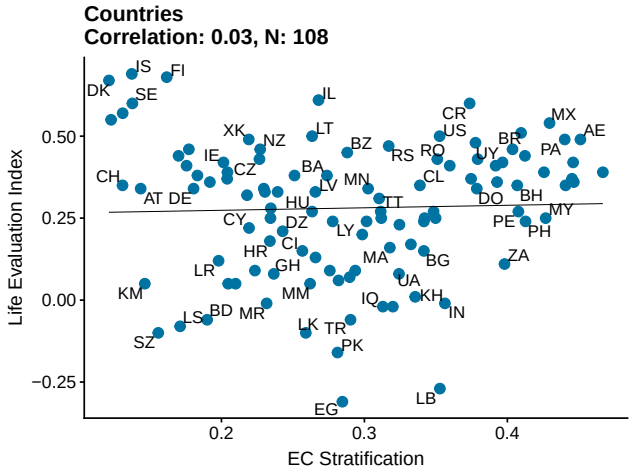
(l) Satisfied with opportunities to make friends

Note: See page 70.

Fig. SI-18: EC Stratification and Socio-Political Outcomes from the Gallup World Poll



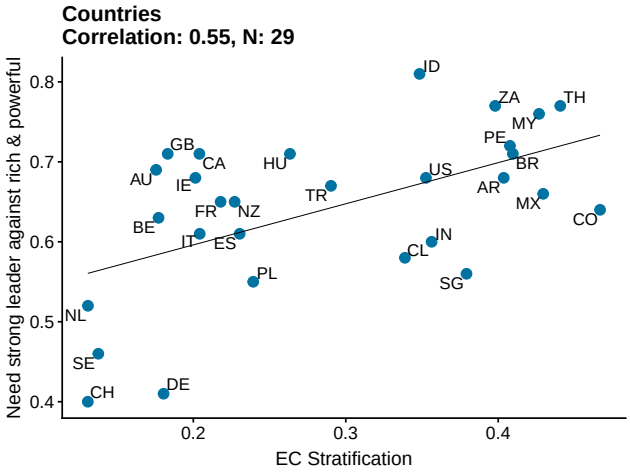
(m) Satisfied w. freedom to make life decisions



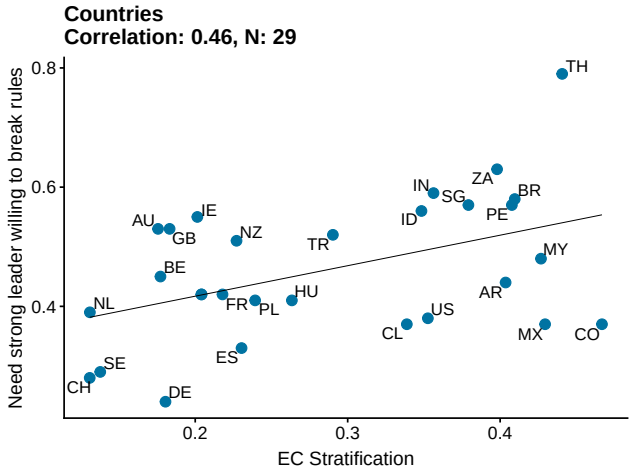
(n) Life Evaluation Index

Note: See page 70.

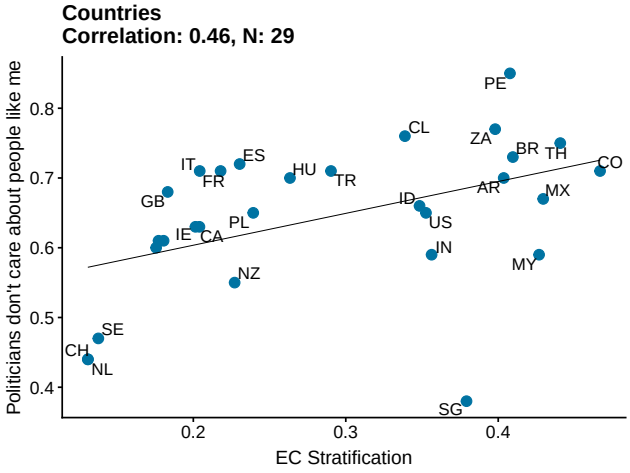
Fig. SI-19: EC Stratification and Socio-Political Outcomes from the Ipsos Populism Report



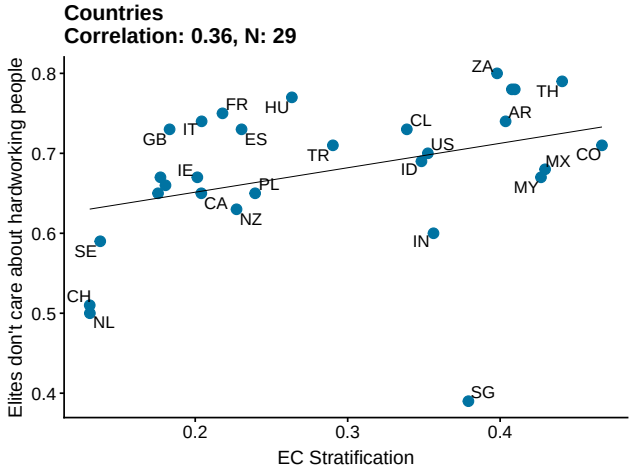
(a) Need strong leader against rich & powerful



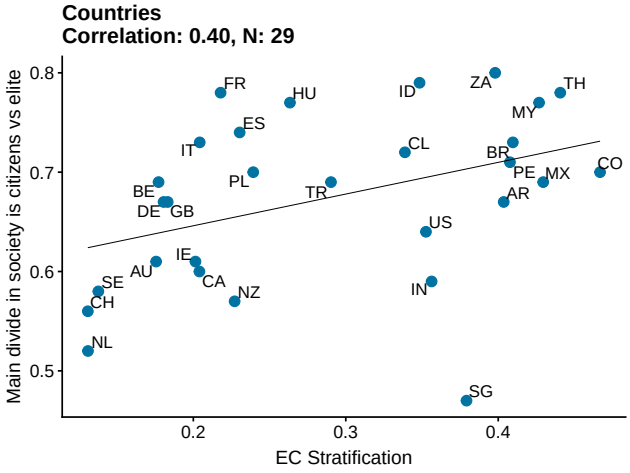
(b) Need strong leader willing to break rules



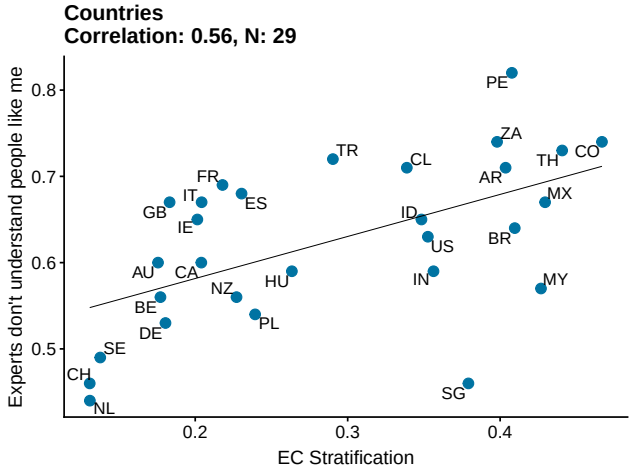
(c) Politicians don't care about people like me



(d) Elites don't care about hardworking people



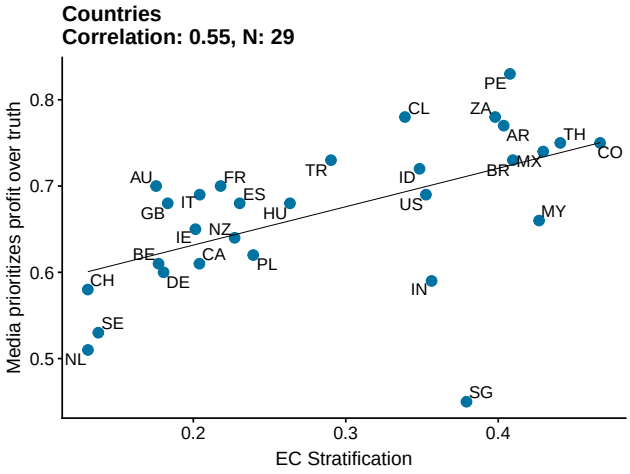
(e) Main divide in society is citizens vs elite



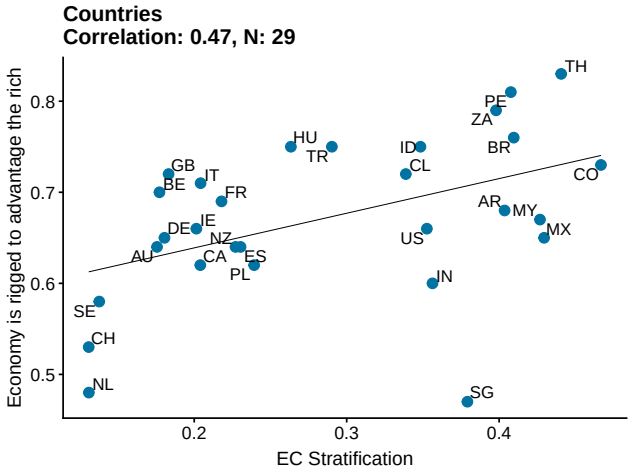
(f) Experts don't understand people like me

Note: Each scatter plot displays the relationship between EC stratification and country-level measures of various socio-political outcomes. Data on outcomes are drawn from the Ipsos Populism Report.

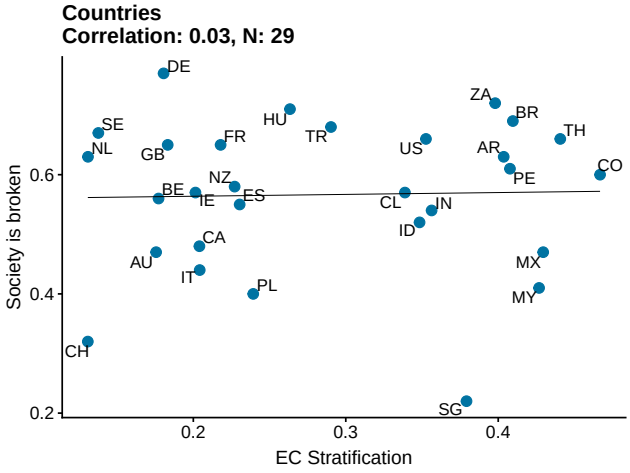
Fig. SI-19: EC Stratification and Socio-Political Outcomes from the Ipsos Populism Report



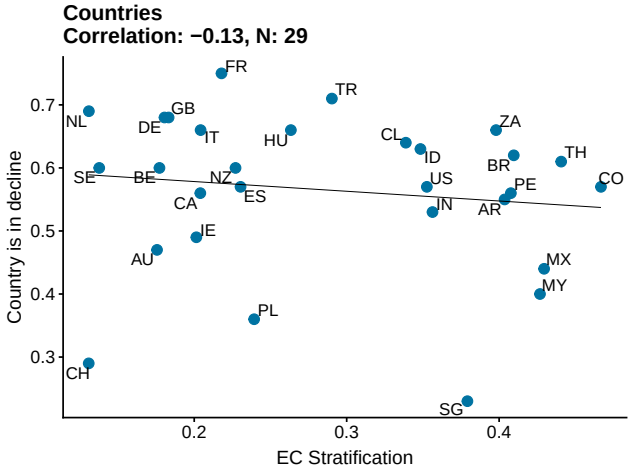
(g) Media prioritizes profit over truth



(h) Economy is rigged to advantage the rich



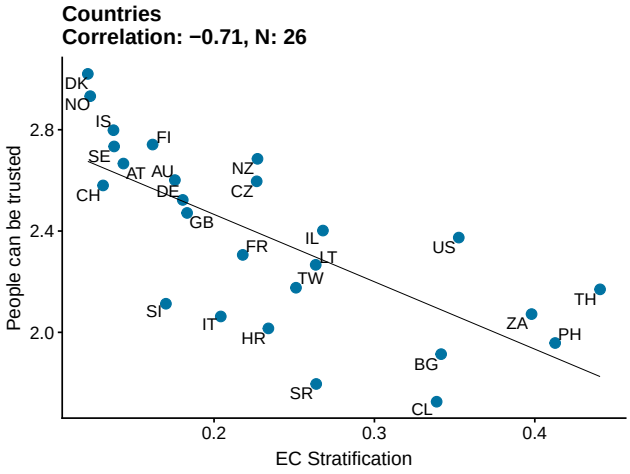
(i) Society is broken



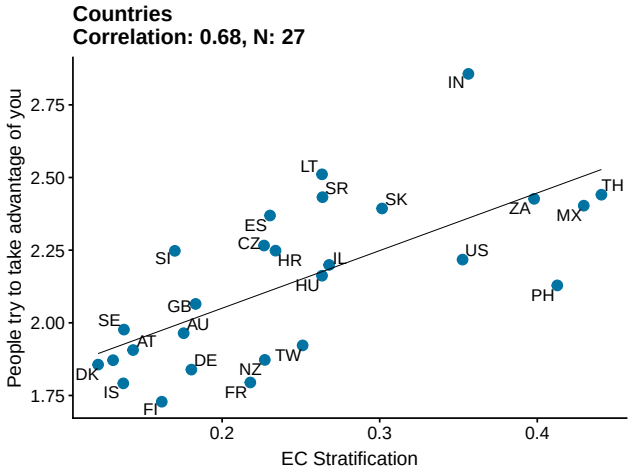
(j) Country is in decline

Note: See page 73.

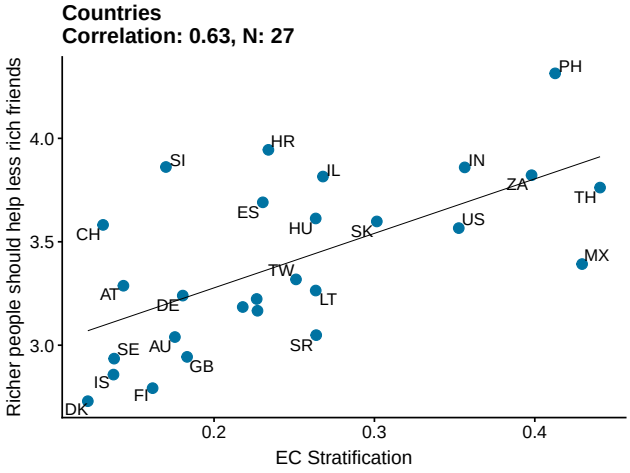
Fig. SI-20: EC Stratification and Socio-Political Outcomes from the ISSP



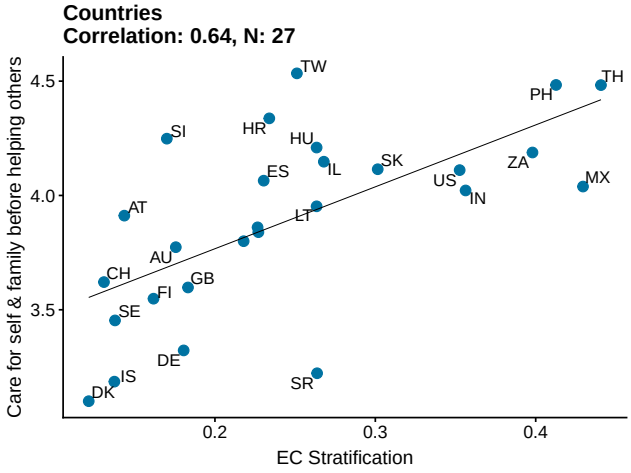
(a) People can be trusted



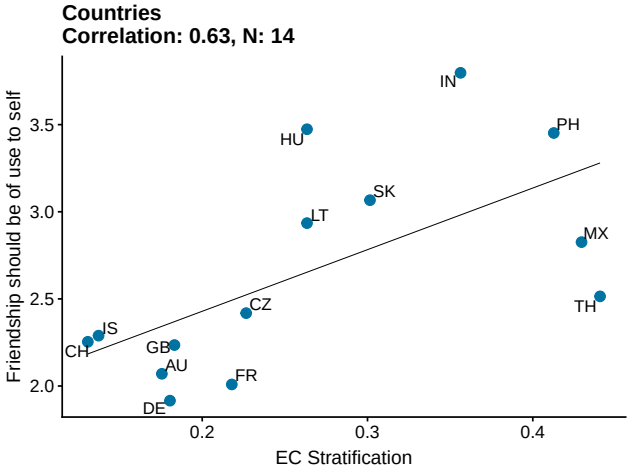
(b) People try to take advantage of you



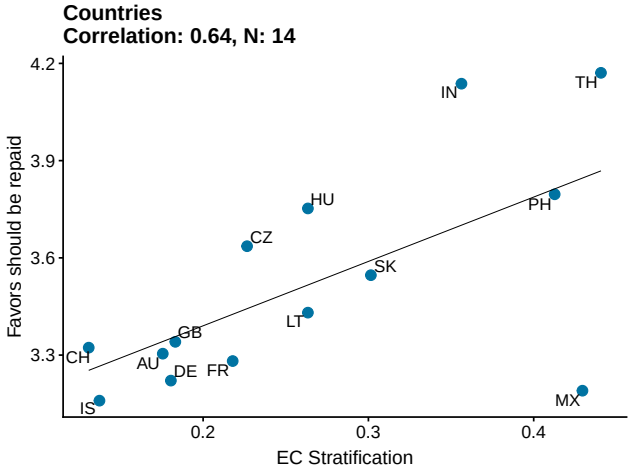
(c) Richer people should help less rich friends



(d) Care for self & family before helping others



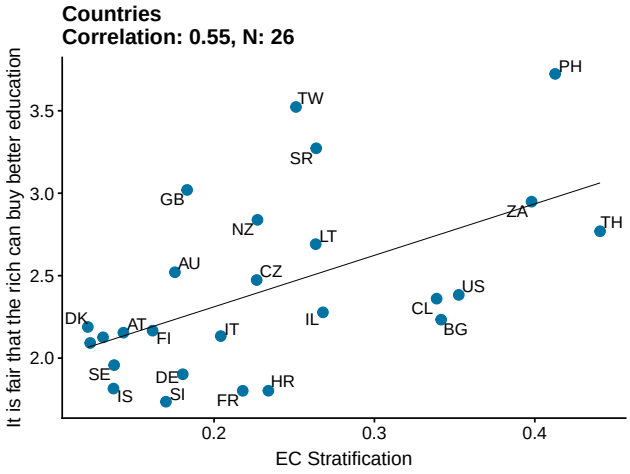
(e) Friendship should be of use to self



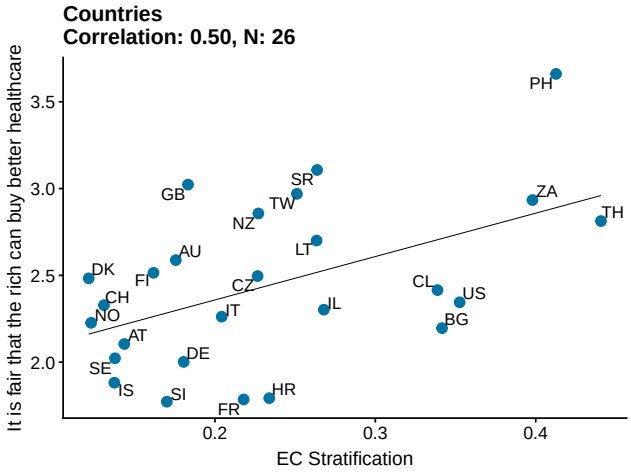
(f) Favors should be repaid

Note: Each scatter plot displays the relationship between EC stratification and country-level measures of various socio-political outcomes. Data on outcomes are drawn from the International Social Survey Programme.

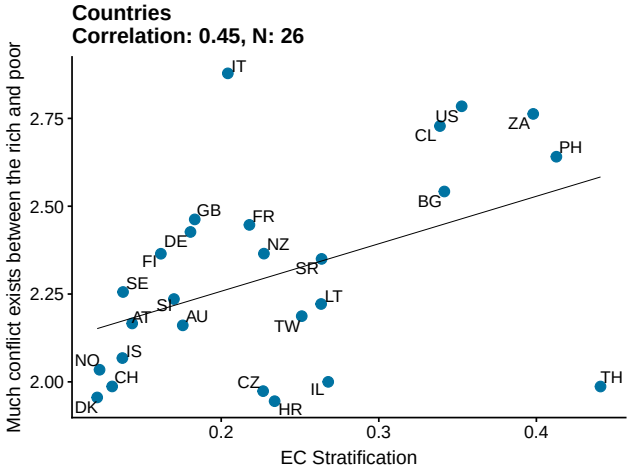
Fig. SI-20: EC Stratification and Socio-Political Outcomes from the ISSP



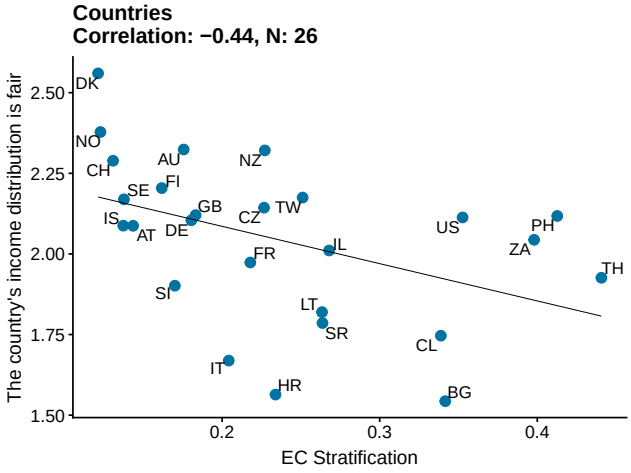
(g) It is fair that rich can buy better education



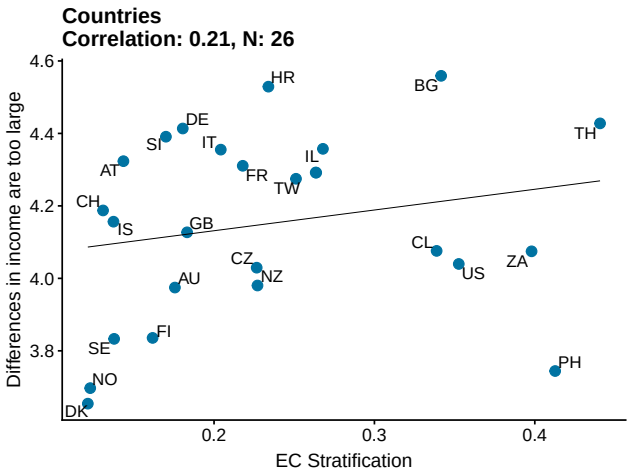
(h) It is fair that rich can buy better healthcare



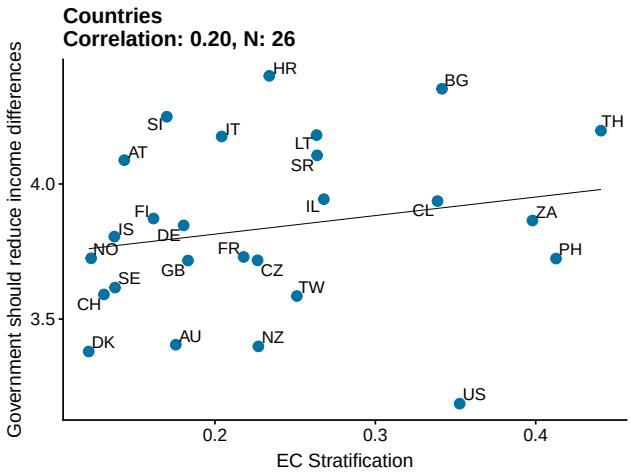
(i) Much conflict exists between rich and poor



(j) The country's income distribution is fair



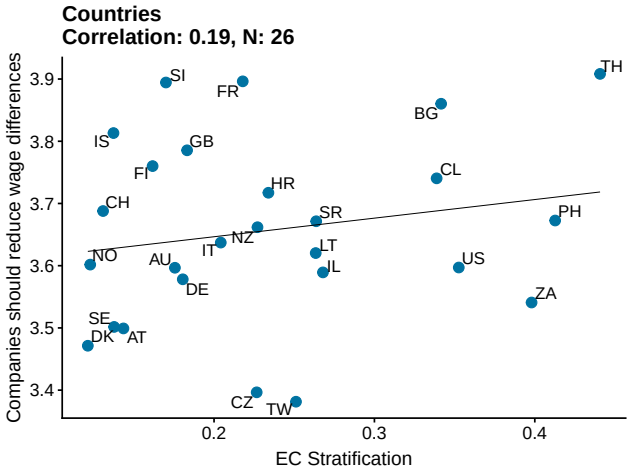
(k) Differences in income are too large



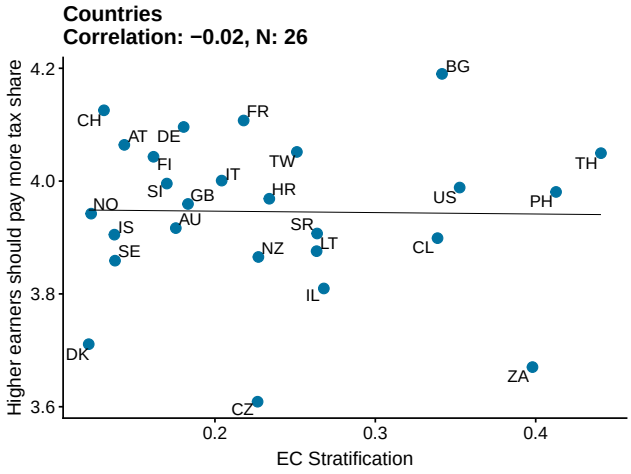
(l) Government should reduce inc. differences

Note: See page 75.

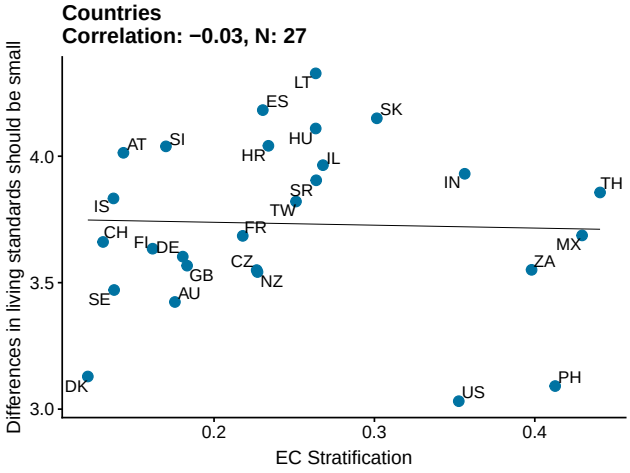
Fig. SI-20: EC Stratification and Socio-Political Outcomes from the ISSP



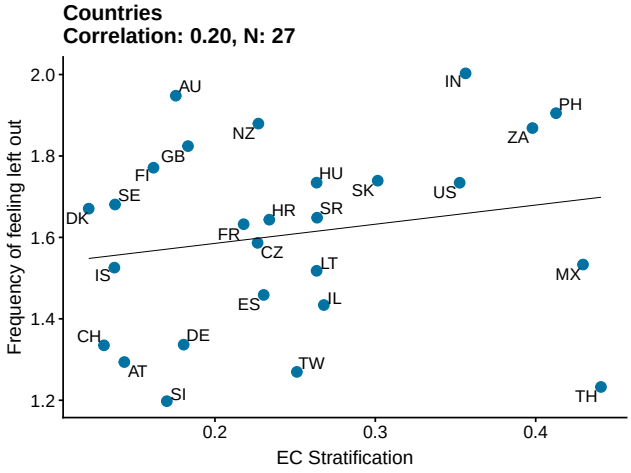
(m) Companies should reduce wage diffs.



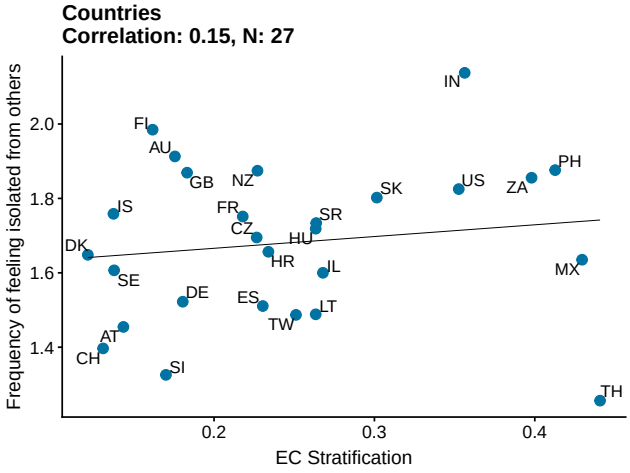
(n) Higher earners should pay more tax share



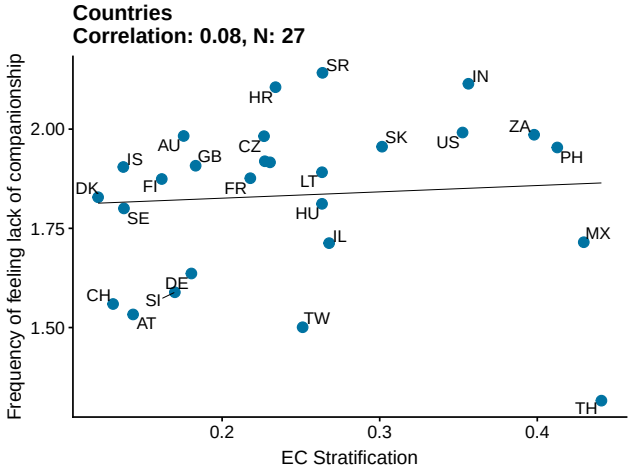
(o) Diffs. in living standards should be small



(p) Frequency of feeling left out



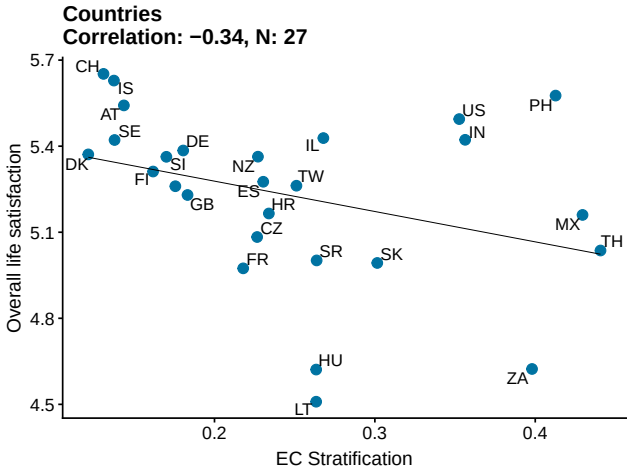
(q) Frequency of feeling isolated from others



(r) Frequency of feeling lack of companionship

Note: See page 75.

Fig. SI-20: EC Stratification and Socio-Political Outcomes from the ISSP



(s) Overall life satisfaction

Note: See page 75.