International Inequalities Institute

WORKING PAPER 99

GEOWEALTH: Spatial wealth inequality data for the United States, 1960-2020

Joel Suss
Bank of England; International Inequalities Institute, LSE

Tom Kemeny
Munk School of Global Affairs & Public Policy, University of Toronto; International Inequalities Institute, LSE

Dylan Connor
School of Geographical Sciences and Urban Planning, Arizona State University

AUGUST 2023

Joel Suss
Bank of England; International Inequalities Institute, LSE

Tom Kemeny
Munk School of Global Affairs & Public Policy, University of Toronto; International Inequalities Institute, LSE

Dylan Connor
School of Geographical Sciences and Urban Planning, Arizona State University

In addition to our working papers series all these publications are available to download free from our website: www.lse.ac.uk/III

For further information on the work of the Institute, please contact the Institute Manager, Liza Ryan at e.ryan@lse.ac.uk

International Inequalities Institute
The London School of Economics and Political Science, Houghton Street, London WC2A 2AE

E Inequalities.institute@lse.ac.uk
W www.lse.ac.uk/III
@LSEInequalities

GEOWEALTH: Spatial wealth inequality data for the United States, 1960-2020

Joel Suss[†], Tom Kemeny[††], and Dylan Connor[†††]

[†]Bank of England and LSE III
[††]University of Toronto and LSE III
[†††]Arizona State University

Abstract

Wealth inequality has been sharply rising in the United States and across many other high-income countries. Due to a lack of data, we know little about how this trend has unfolded across locations within countries. Investigating this subnational geography of wealth is crucial, as from one generation to the next, wealth powerfully shapes opportunity and disadvantage across individuals and communities. Using machine-learning-based imputation to link newly assembled national historical surveys conducted by the U.S. Federal Reserve to population survey microdata, the data presented in this paper addresses this gap. The Geographic Wealth Inequality Database ("GEOWEALTH") provides the first estimates of the level and distribution of wealth at various geographical scales within the United States from 1960 to 2020. The GEOWEALTH database enables new lines of investigation into the contribution of spatial wealth disparities to major societal challenges including wealth concentration, spatial income inequality, social mobility, housing unaffordability, and political polarization.

Introduction

Following a four decade period of sustained growth in wealth inequality in the United States, less than 10 percent of families now possess 70 percent of national wealth (Kuhn et al., 2020; Saez and Zucman, 2016; Piketty and Zucman, 2015). The trajectory of rising national wealth inequality resembles similarly unfavorable long-term patterns of income polarization and declining intergenerational mobility (Goldin and Katz, 2009; Song et al., 2020). For historical income and intergenerational mobility dynamics, there is a growing realization that these prevailing trends have, in fact, arisen from a strongly differentiated subnational geography (Sampson, 2019; Kemeny and Storper, 2022; Connor and Storper, 2020a). In contrast, we still know very little about the geography of wealth inequality and how it has changed over time.

This knowledge gap not only limits our understanding of broader societal trends in inequality, but also the social, economic, political, and even epidemiological consequences of concentrated wealth. (Neckerman and Torche, 2007; Yellen, 2014) Specifically, wealth inequality has previously been linked to the local provision of public goods (Côté et al., 2015; Baumgärtner et al., 2017), social mobility (Hansen, 2014; Acolin and Wachter, 2017; Chetty et al., 2017; Connor and Storper, 2020a), support for populism (Cramer, 2016; Rodríguez-Pose, 2018; Broz et al., 2021), and the health of local economies (Moretti, 2010; Couture et al., 2019). Despite the role of wealth in giving rise to disparities in income (Piketty et al., 2018), wealth and income represent distinctive facets of economic inequality, with potentially different roots and implications (Killewald et al., 2017; Cowell et al., 2017). There is therefore a great need for focused investigation into the changing geography of wealth within and beyond the United States.

This article presents a new source of information on long-term geography of wealth in the United States: The Spatial Wealth Inequality Database ("GEOWEALTH"). GEOWEALTH provides estimates of the level and distribution of wealth at various geographical scales within the United States from 1960 to 2020. These estimates were generated through the application of machine-learning-based imputation to link newly assembled national historical surveys conducted by the U.S. Federal Reserve to population survey microdata. The GEOWEALTH database not only enables new lines investigation into the causes and consequences of spatial wealth inequality in the United States, but also a flexible methodological framework for generating estimates of personal wealth across a range of geographical and historical contexts.

The previous limitations on our understanding of the geography of wealth reflects several key constraints in terms of data and measurement. The stock of a households' wealth is typically measured as the value of its assets net of total debts, across a range of asset types, such as cash holdings, real estate, and financial investments. Unlike income flows, which are reported in the census, few public data sources report on personal assets and debts, or on their constituent components. Our understanding of how individuals' wealth in the U.S. has changed over time comes from confidential administrative data linked to taxes (Piketty et al., 2018) or from a range of smaller household surveys (Killewald et al., 2017). While each has advantages and disadvantages (Kuhn et al., 2020), concerns around confidentiality mean that none of these data sources can be directly used to describe meaningful spatial disparities in wealth. The estimates provided in the GEOWEALTH database do not face confidentiality constraints and thus facilitate detailed spatiotemporal analysis of wealth dynamics.

Importantly, the estimates of wealth inequality provided in the GEOWEALTH database are derived from multidimensional measures of assets and debts. Most of what was previously known about the geography of wealth was confined to the housing market (Gyourko et al., 2013; Ganong and Shoag, 2017). This is because home values and mortgage information are reported in several public data sources, including in tabulations and extracts of the decennial census. But, while important – especially for those who are less affluent – home values are only one among several channels through which wealth can vary across locations. In practice, across American households, home values and net wealth are only moderately correlated ($r = 0.535, p < 0.001$,

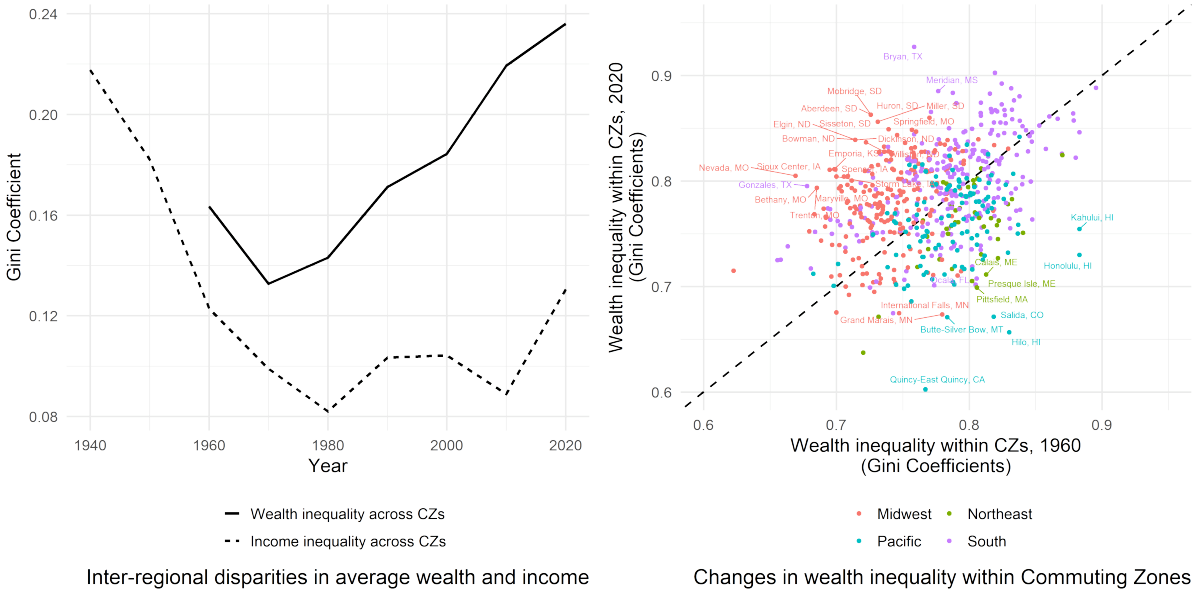

Figure 1: The geography of wealth and local wealth inequality, 1960-2020

Note: For the period 1960-2019, Panel (A) describes the evolution of Gini coefficients tracking inter-regional inequality in terms of average household income and wealth across U.S. commuting zones, defined using 1990-vintage commuting flow data. Wealth estimates come from the GEOWEALTH dataset that is the primary output of this study. Income series is estimated based on U.S. Census population survey microdata (Decennial and American Community Survey), from IPUMS (Ruggles et al., 2022). Note that this panel shows that wealth gaps between places has grown much more sharply than income gaps. Panel (B) visualizes the correlation between 1960 and 2020 measures of local wealth inequality – that is, levels of household wealth inequality within each commuting zone, again measured using Gini coefficients. The positive but only moderately strong correlation between local wealth inequality suggests a mix of continuity and turbulence in the ranks of more and less wealth-unequal locations in the United States.

based on authors' calculations using Bureau of Labor Statistics' Survey of Consumer Finances 2015-16 data). The GEOWEALTH database therefore provides insight on wealth inequality, as measured from a much broader range of asset types. This work also contributes to a broader effort to using big data and computation to inform efforts to tackle issues of inequality and equity (O'Brien, 2022; Chetty, 2021).

The framework used to construct the GEOWEALTH database relies on the application of machine-learning-based imputation. We generate predictive models of household wealth using ensemble learning, applying rich survey information from the Federal Reserve's Survey of Consumer Finances (SCF) as a means to predict wealth among households in Census population surveys that include geographical identifiers. The end result is a dataset that permits description of inter-place variation in wealth ('geography of wealth'), as well as the distribution of wealth within individual local economies ('local wealth inequality'). The resulting data allows researchers to track the evolving geography of wealth and wealth inequality between 1960 and 2020 across more than 700 local labor markets that span the entirety of the U.S.

Our initial investigation of the spatial and temporal patterns of the GEOWEALTH database reveals three key features of the changing geography of wealth in the United States (see Fig.1). First, Panel A reveals that the distribution of wealth *between* regions has become steadily more unequal since 1970. US wealth holdings have become increasingly concentrated in a smaller set of regions. Second, inter-regional *wealth* disparities have growing more sharply than inter-regional *income* disparities. This confirms earlier intuition that spatial wealth inequalities require investigation over and above the study of income inequality. The sharp exacerbation of wealth inequality over this period make this a particularly urgent topic for further research.

Finally, Fig.1 Panel B reveals patterns of both turbulence and persistence in the distribution of wealth *within* regions since 1960. Economies in the South (purple) are particularly notable in having high levels of wealth inequality from 1960 through to 2020. High levels of wealth inequality in the South is therefore a long-term feature of the region. In the Midwest, however, inequality started out relatively low in 1960 but has worsened considerably over the last six decades. The worsening of inequality in the Midwest over this period is consistent with findings from studies of regional income inequality (Manduca, 2019; Kemeny and Storper, 2022) and intergenerational mobility (Connor and Storper, 2020b), perhaps pointing to the interdependence and common underlying sources affecting different facets of spatial inequality. Our publication of the GEOWEALTH database provides new avenues for investigating the causes, consequences, and common coherence, of these patterns.

Methods

We use the public-release files of the Federal Reserve's Survey of Consumer Finances (SCF) that spans the 1989-2019 period. Making use of the multiple household demographic and income attributes that are present in SCF, we predict household total wealth, gross assets and debts for households in successive waves of public-use Decennial and American Community Survey (ACS) microdata from the U.S. Census Bureau, obtained from IPUMS (Ruggles et al., 2022). Using Census households' place of residence information, we are able to generate estimates of the sub-national geography of net wealth and wealth inequality at various scales, including metropolitan areas and commuting zones.[1]

In the absence of directly-observed, geographically-identified wealth data, our approach offers key strengths. Like other surveys that record information on wealth, such as the Survey of Income and Program Participation (SIPP), SCF includes relevant demographic correlates. Unlike SIPP and other surveys, however, SCF includes detailed information on household wealth and incomes. Crucially, SCF-based variables capturing distinct categories of income, including wages, investment, and business income, closely matching those found in Census population surveys. The result is that imputation from SCF to these surveys is not forced to rely chiefly on demographics for which we know there to be meaningful, geographically-conditioned unobserved heterogeneity in relation to income (and potentially, to wealth). Put another way, we know that the economic characteristics of individuals with observably equivalent demographic features and educational attainment differ based on location (Combes et al., 2008). Thus, by capturing not just demographics, but also detailed income information, housing tenure and value, our imputation model generates superior wealth predictions. Our approach also benefits from the fact that missing wealth data in the population surveys are 'missing completely at random': the mechanism driving missingness will not be a source of bias in prediction (Rubin, 1976).

Beyond the unparalleled depth of detail about wealth and income in the SCF, these data offer the best basis for prediction because of their reliability as a source of information about the full range of the wealth distribution (Killewald et al., 2017). While tax-derived data may capture the very top of the distribution, the SCF will better describe wealth for middle-income households whose housing-centered assets are not sources of taxable income, as well as low-wealth households that may pay little or no taxes (Kuhn et al., 2020). Meanwhile, oversamples for rich households should improve coverage at the top of the wealth distribution.

Our construction of the data comprises three steps: (1) build a model of wealth using the SCF; (2) predict wealth using Census population survey data; and (3) estimate wealth and wealth inequality at various spatial scales.

[1] Although the data underlying this study enumerate characteristics of human subjects and their households, they have been fully anonymized by the agencies responsible for the data. The study nonetheless obtained approval from the Social Sciences, Humanities & Education Research Ethics Board at the University of Toronto.

Step 1: Build a model to predict household wealth

Using the SCF data, we build and combine a set of stacked ensemble models ('ensemble combination') to arrive at the most accurate available predictions of household wealth. As a general approach, stacking involves the combination of a number of predictive models (Breiman, 1996b; Zhou, 2012). Typically, a set of base (or Level 1) models are trained on a subset of the data. A second-level model is then fit on a separate subset of the data, using the Level 1 predictions as inputs. The aim in principle is to garner improvements in prediction that result from bringing together a diverse (relatively uncorrelated) and accurate set of models.

We chose to use stacked ensembles based on a careful comparison between different potential approaches. We evaluated the ensemble combination relative to alternative treatments of net wealth, including modelling the inverse hyperbolic sine transformation of net wealth, and taking the net difference between models separately predicting gross wealth and debts. We also evaluated the performance of each ensemble relative to the individual constituent models that comprise it. Details of comparisons between our preferred approach and other models are found in the Performance Analysis section; these demonstrate that our stacked ensembles jointly outperform available alternatives.

Our stacked ensemble is made up of seven base models: generalized linear regression (GLM), elastic net regression (EN), random forest (RF), gradient boosted trees (GB), neural network (NN), support vector machine (SVM), and K-nearest neighbors (KNN). Note that some of these 'standalone' models are themselves ensembles. In particular, random forests are known as 'bagged ensembles' – bootstrapped aggregations of individual decision or regression trees, while boosted models are ensembles of sequentially grown trees (Breiman, 1996a; Hastie et al., 2009). For the Level 2 model, we estimate a simple linear regression.

To produce a final predicted value of wealth for each household, we combine the outputs of four separate ensembles. First, we estimate the probability of having positive wealth; 8.9% of households in the SCF sample have either no wealth or are in debt. If the predicted probability is at or above the decision threshold (which we vary by Census year in the imputation step), the fitted value of a positive wealth ensemble model is chosen, built using only data on households with some positive value of wealth. If the predicted probability is below the threshold, a further binary stacked ensemble predicts whether the household in question has zero or some quantum of negative wealth. If the latter, a stacked ensemble estimates the quantum, built using data only for households with negative wealth. The combination of different ensembles also has an advantage over other approaches in that we can calibrate levels of inequality by altering the decision threshold at which we classify households as either having wealth or not. As we discuss in the Technical Validation section below, adjusting the threshold can alter some of the inequality estimates, allowing us to match existing benchmarks at aggregate geographical levels.

To fit the ensembles, we train Level 1 models on a random 80% of the SCF data (N = 42,748), and the Level 2 regression model on a validation set consisting of 10% (N = 5,341). We then evaluate performance on a test set consisting of the remaining 10% of data (N = 5,341). For Level 1 models which have hyperparameters to select (i.e. EN, RF, GB, NN, SVM and KNN), in each fold we employ 5-fold cross-validation with random grid search (length of 10). For binary classification models – those modelling whether household has positive wealth or none/negative wealth – we up-sample the negative class within each cross-validation fold such that there a balanced number of cases. That is, we sample with replacement from the subset of observations with negative/no wealth to ensure a 50/50 split. Figure 2 depicts the stacked ensemble structure.

Variable selection and transformations

To build the ensembles, we select the set of variables that are available in both SCF and the Decennial Census/ACS. The complete set of variables (the 'full model') is available from 2008

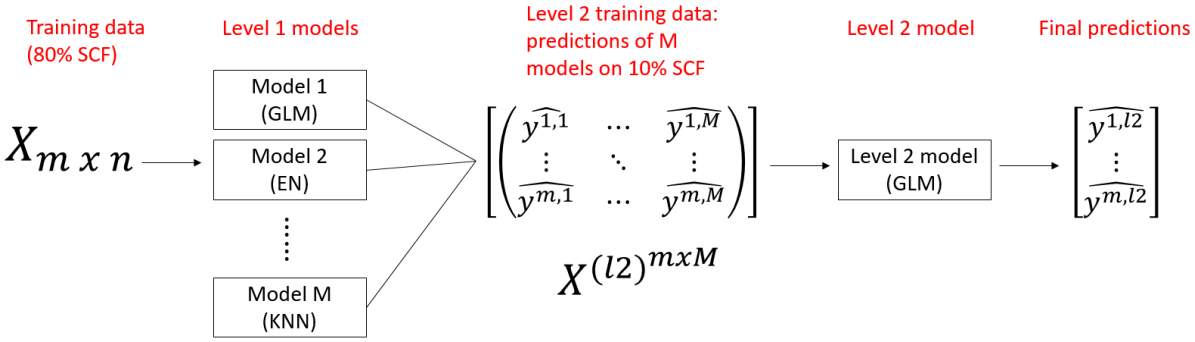

Figure 2: Structure of stack ensemble

onwards in the ACS and from 1989 onwards in SCF. Table 1 describes these inputs, listing variable names from the ACS, as well as corresponding variables in the SCF.

Certain potentially relevant variables are not available in some early years of the Decennial. Therefore, to match Census data availability, beyond years in which the full model is estimable, we leverage the available data and build unique ensembles. Table 2 provides a breakdown of the availability over the study period of each of the relevant variables.

Before splitting the data between training, validation and test samples, we transform some variables in the SCF dataset. In particular, we use the inverse hyperbolic sine transformation for continuous variables, which have a pronounced right-skew and contain zeroes. This issue mainly applies to the home value and income variables. We also harmonize several of our categorical variables by aggregating values into a coarser set of common categories. We do not remove any observations that might in other circumstances be considered outliers; such extreme values are potentially important observations for our exercise, indicating the presence of very wealthy or high-earning households.

Variable importance

Although machine learning approaches to prediction tend to perform better than simple linear regression, weak explainability can be a limitation in these models. To better understand the relative importance of each variable, we compute Shapley values for each test set observation, a well-known approach to model interpretability that originated from research in game theory. (Lundberg and Lee, 2017) Shapley values provide a local interpretation gauging the relative importance of each variable for each specific prediction. We can then take the mean absolute Shapley value per variable to make global assessments of importance. Figure 3 presents these results separately for the dominant model (random forest) for each component of the ensemble combination – the binary, positive wealth, and negative wealth ensembles.

Figure 3 rank orders the importance of each input variable across our three models. Home values (VALUEH) are the most important variable for the binary and positive wealth ensembles followed by investment income (INCINVST). However, the variables of important to the negative wealth ensemble look quite different. Specifically, the year of observation in the census is single most important predictor, followed by vehicle availability (NVEHIC) and again home values (VALUEH). The relationship between year of observation and negative wealth is a reflection of the rising levels of personal debt over recent decades.

Because Shapley values give us local (i.e. prediction-by-prediction) information on explanatory power, we can also investigate how characteristics may contribute differently to wealth and debt depending on a household's net worth. We examine the mean absolute Shapley values for

Category	Variable (Decennial/ACS)	Variable (SCF)	Description
Housing information			
	VALUEH	houses	Value of primary residence
	OWNERSHP	houses, hdebt	Housing tenure – own outright, own with a mortgage, or rent
	MORTAMT1	paymort1	Monthly payment for 1st residential mortgage
	MORTAMT2	paymort2	Monthly payment for 2nd residential mortgage
	TAXINCL	x810	Whether tax is included as part of mortgage payment
	INSINCL	x810	Whether insurance is included as part of mortgage payment
	PROPTX99	x721	Annual property tax amount
	RENT	rent	Usual monthly rent
Detailed income information			
	INCWAGE	x5702	Wage and salary income
	INCBUS	x5704	Business income
	INCSS	x5722	Social security income
	INCWELFR	x5716, x5720, x5724, x5725	Income from welfare receipts
	INCINVST	x5706, x5708, x5710, x5714	Investment, interest and dividend income
	INCRETIR	x5724, x5725	Retirement income, e.g. IRA and 401k
	INCOTHER	x5712, x5718	Other income not included in available categories
Demographic information			
	AGE	age	Age
	RACE	x6809	Race
	EDUC	x5901, x5902, x5904, x5905	Educational attainment
	SEX	hhsex	Sex
	MARST	x8023	Marital Status
	FAMSIZE	x101	Number of own family members in household
	YEAR		Year
Employment information			
	OCC	x7401, x7411	Occupation
	IND	x7402, x7412	Industry
	EMPSTAT	x4100, x4700	Employment status
	CLASSWKR	x4106, x4706	Class of worker
	UHRSWORK	x4110, x4710	Usual hours worked per week
	WKSWORK2	x4111, x4711	Weeks worked last year, intervalled
Other information			
	VEHICLES	nvehic	Number of vehicles available
	HCOVANY	x6341	Any health insurance coverage

Note: Detailed definitions of variables available at usa.ipums.org/usa-action/variables/group. Codebook for SCF (2019) found at www.federalreserve.gov/econres/files/codebk2019.txt

Table 1: Census and SCF variables and definitions by category

high and low net-worth households separately, which we define as those with $10mn or more and between $0 and $25k respectively – see Figure 4. When comparing these two groups we see that the most important variables differ depending on whether a household has high or low net worth: investment income is of greater consequence at high net worth and home values matter more at lower wealth levels. This is in accordance with a study by Saez and Zucman (2016), that finds that increases in investment income are primarily responsible for increases in top wealth shares. We also see that rent is an important predictor for only the low net-worth households, whereas income derived from other sources is an important predictor for high net-worth households.

Step 2: Impute wealth using Census population survey data

Armed with trained stack ensembles, we then impute wealth for each household observed in the census microdata for the years from 1960 to 2020. As noted above, we train different ensembles to impute wealth depending on the year of the census (to account for whether housing variables, or income sub-components are included). We first filter the Census/ACS data as follows: we

Variable (Decennial/ACS)	2020	2010	2000	1990	1980	1970	1960	1950	1940
VALUEH	X	X	X	X	X	X	X		X
OWNERSHP	X	X	X	X	X	X	X		X
MORTAMT1	X	X	X	X					
MORTAMT2	X	X	X	X					
TAXINCL	X	X	X	X	X				
INSINCL	X	X	X	X	X				
PROPTX99	X	X	X	X					
RENT	X	X	X	X	X	X	X		X
INCWAGE	X	X	X	X	X	X	X	X	X
INCBUS	X	X	X	X	X	X	X	X	
INCSS	X	X	X	X	X	X			
INCWELFR	X	X	X	X	X	X			
INCINVST	X	X	X	X	X				
INCRETIR	X	X	X	X					
INCOTHER	X	X	X	X	X	X	X	X	
AGE	X	X	X	X	X	X	X	X	X
RACE	X	X	X	X	X	X	X	X	X
EDUC	X	X	X	X	X	X	X	X	X
SEX	X	X	X	X	X	X	X	X	X
MARST	X	X	X	X	X	X	X	X	X
OCC	X	X	X	X	X	X	X	X	X
IND	X	X	X	X	X	X	X	X	X
EMPSTAT	X	X	X	X	X	X	X	X	X
CLASSWKR	X	X	X	X	X	X	X	X	X
UHRSWORK	X	X	X	X	X				
WKSWORK2	X	X	X	X	X	X	X	X	X
VEHICLE	X	X	X	X					
HCOVANY	X	X							

Note: Detailed definitions of variables in the Decennial and American Community Survey available at usa.ipums.org/usa-action/variables/group.

Table 2: Census variables and availability by year

remove group quarters and institutionalized individuals. To match SCF, we take the household head to compute demographic information. We then adjust all income and housing values to account for inflation, bringing these to 2019 dollars (to match the inflation-adjusted SCF).

We also adjust top and bottom-coded values in the census, given the importance of censoring for inequality estimates (Burkhauser et al., 2011; Fichtenbaum and Shahidi, 1988). For each variable that is censored, we compute a new maximum or minimum value by multiplying given top and bottom-codes by 25. We then adjust top and bottom-coded observations by sampling from a truncated Pareto distribution with values between the top (bottom) code and new maximum (minimum).[2] Pareto distribution parameters are estimated using distributional information from the SCF (which is not censored). As expected, adjusting top and bottom-codes has a large effect on our wealth predictions and inequality estimates – see Figure 5 for the estimates of the Gini coefficient by State pre- and post-adjustment for the 2020 ACS.

Step 3: Estimate inequality at varying geographies using imputed wealth from census

Once we have imputed household wealth in the census, we compute inequality estimates at varying levels of geography. In order to ensure that the population survey microdata reflects the population, we use Census-provided household weights provided in computing inequality.

We estimate inequality (Gini coefficients, wealth shares), as well as mean and median wealth per area, at multiple spatial scales: Public Use Microdata Areas (PUMAs), 1990 Commuting

[2]The Pareto distribution has been shown to reasonably approximate the upper tail of the income distribution. (Fichtenbaum and Shahidi, 1988)

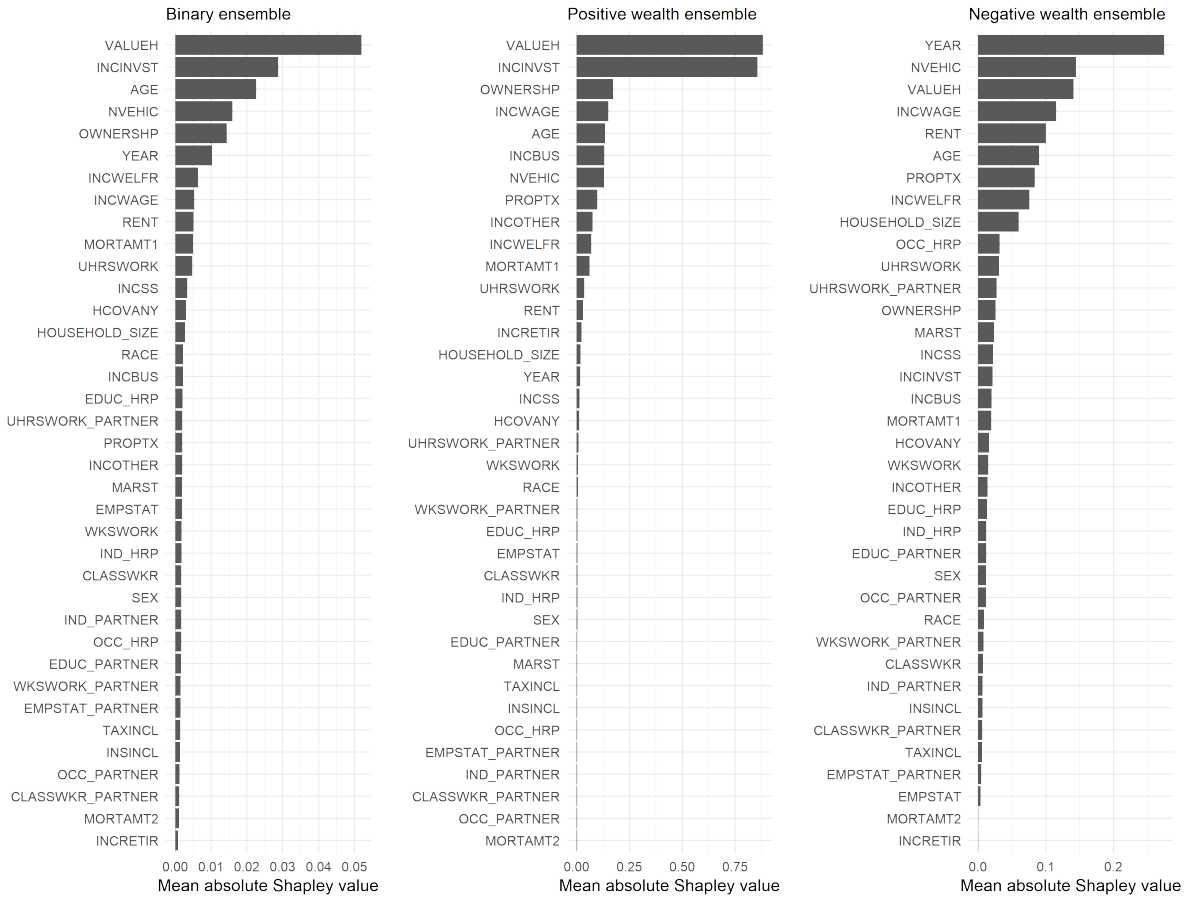

Figure 3: Mean absolute Shapley values for ensemble combination.

Note: Displayed are mean absolute Shapley values for household variables with predictive power on household wealth, with higher Shapley values indicating greater predictive power. Here we include Shapley values for separate predictions of the binary ensemble (whether positive or negative wealth); and individual models predicting levels of positive or negative wealth.

Zones (CZs), Metropolitan Areas, States, Regions, and the country as a whole. PUMAs are available for the 1960 census and from 1990 onwards. We use the crosswalks provided by Dorn (2009) to infer a households' CZ of residence based on their reported PUMA (or state economic areas and county groups for the years 1950, 1970, and 1980). This requires multiplying the household weights by a factor which represents the probability of belonging to a given CZ (which is 1 where PUMAs or county groups lie entirely within a CZ, and less than 1 when split across multiple CZs). Table 5 provides details of data availability at different spatial scales.

Given the imputation procedures used to estimate local wealth levels and distributions, it is appealing to capture the uncertainty around these estimates. To do so, we bootstrap a distribution of 100 inequality estimates, sampling with replacement from the distribution of imputed household wealth, using a 5% confidence level. A main advantage to this simulated approach is that it does not require any assumptions regarding the normality of the distribution of inequality estimates (Efron, 1992).

Comparisons of the resulting dataset against published wealth data for the United States are reported in the Technical Validation section.

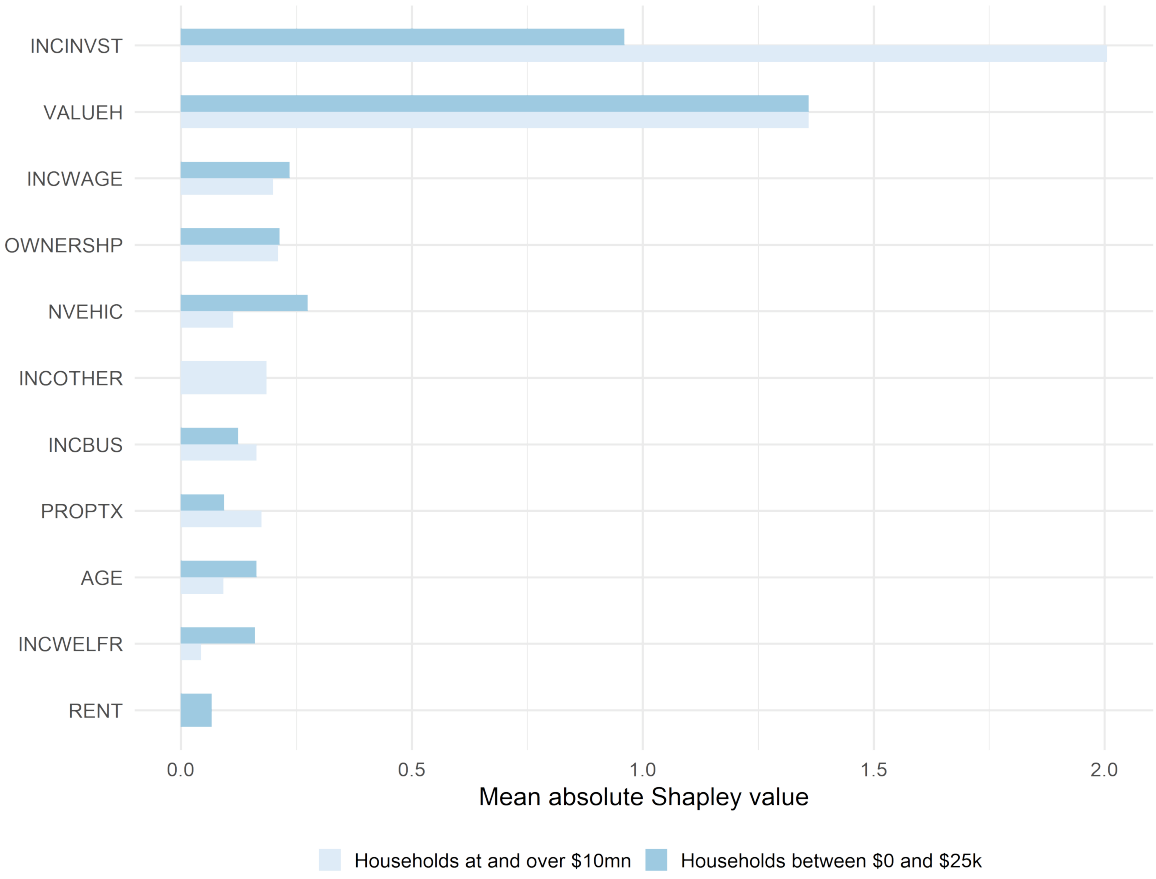

Figure 4: Mean absolute Shapley values for high- and low-income households

Note: Displayed are mean absolute Shapley values for household variables generated from the positive wealth ensemble on sub-samples of the test data. Higher values indicate greater variable importance. The figure compares variable importance for households with income between 0 and $25,000, and at and over $10 million.

Technical Validation

In this section, we present the exercises and analyses undertaken to validate the technical quality of the dataset. Three kinds of validation are reported. First, we describe the performance analysis on the SCF test sample used to arrive at our final model of wealth. Second, we conduct out-of-sample validation to verify that our model performs well in predicting household wealth using data beyond the SCF. Third, we report on exercises that compare our imputed estimates of aggregate wealth and wealth inequality against other published estimates.

Performance evaluation, SCF test sample

In order to select our final model of household wealth, we run a 'horse race' to evaluate competing approaches. The winner is the approach which dominates in terms of predictive performance on the SCF hold-out sample – the 10% test sample not used for fitting any model.

As reported in Step 1 of the Methods section, we fit a variety of well-known models and architectures, also exploring combinations of these in a stacked ensemble. We also consider the implications of different transformations of the outcome measure. Specifically, we explore outcomes as follows:

- 'ENS': binary model predicting whether a household has positive wealth; a model that predicts positive wealth; a model that predicts negative wealth; and a binary model whether

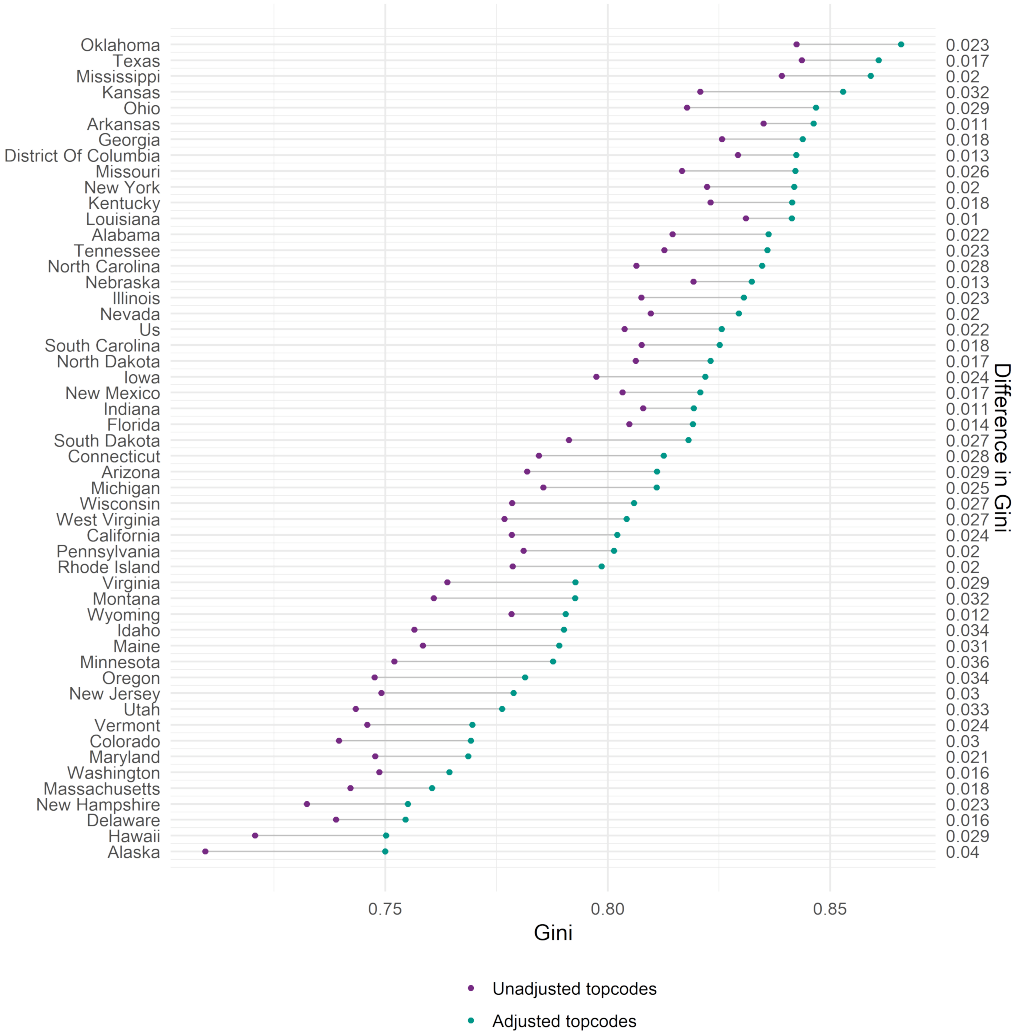

Figure 5: Inequality estimates before and after topcode and bottomcode adjustments, 2020
Note: Displayed is the impact on the estimated Gini Coefficients when adjusting top and bottom-coded household variables in the Census/ACS. The right-hand axis for the figure provides the difference in the Gini Coefficient between adjusted and unadjusted estimates.

a household has zero or some negative value of net wealth

- 'IHS': the inverse hyperbolic sine transformation of net wealth
- 'WD': the net difference of separate models predicting gross wealth and debt

We test these alternative approaches in order to arrive at a final predicted net wealth value that adequately captures the proportion of households with zero or negative net wealth, while also accurately predicting quantitative values of net wealth.

To compare these permutations, we employ a number of performance metrics, shown in Table 3 for the full model. First, we evaluate the ability of each approach to discriminate between binary classes, that is, households that have some positive value of net wealth versus those that do not, and households that have zero wealth versus those that have negative net wealth. For this discriminant ability, we focus on three key measures: the Brier score, the Kappa statistic, and the Area Under the Receiver Operating Characteristic Curve (AUC). For illustrative purposes, we also report overall accuracy, which reports the proportion of correct cases; the true positive rate (TPR) that captures the proportion of households with some

Model	Brier	Kappa	AUC	Accuracy	TPR	TNR	RMSE (+)	RMSE (-)
ENS	0.058	0.426	0.901-0.921	0.887	0.910	0.637	0.999	1.480
IHS	0.282	0.000	0.639-0.686	0.916	1.000	0.000	12.953	15.200
WD	0.078	0.000	0.538-0.566	0.916	1.000	0.000	1.351	19.239
GLM	0.146	0.394	0.884-0.907	0.877	0.900	0.623	1.273	1.475
EN	0.146	0.392	0.883-0.906	0.901	0.939	0.480	1.320	1.457
RF	0.064	0.396	0.887-0.909	0.879	0.903	0.617	1.030	1.457
XGB	0.129	0.420	0.901-0.921	0.887	0.910	0.626	1.003	1.458
NET	0.148	0.384	0.882-0.905	0.863	0.880	0.677	1.064	1.744
SVM	0.145	0.361	0.864-0.89	0.875	0.904	0.558	1.309	1.490
KNN	0.209	0.316	0.84-0.871	0.814	0.821	0.740	1.894	1.505

Note: ENS is the full 4-component ensemble model; IHS is the model predicting the inverse hyperbolic sine transformation of net wealth; WD is the model predicting the net difference between estimates of gross wealth and debt; GLM is the generalized linear model (logistic transformation for binary model); EN is the elastic net model; RF is the random forest; XGB is the gradient boosted trees model; NET is the artifical neural network; SVM is the support vector machine; and KNN is the K nearest neighbors model.

Table 3: Comparison of performance across models for the full model (i.e. 2010-2020 variables).

positive net wealth that are predicted to have positive net wealth; and the true negative rate (TNR), which describes the proportion of households with some negative net wealth that are predicted to have negative net wealth. While the accuracy measure is typically used to measure performance, in situations where there is a large imbalance between classes – such as our own – it can be misleading. Specifically, the result can be high levels of accuracy while the minority class is not well predicted. The TPR and TNR can reveal whether there is an imbalance in accuracy across the different classes. The Kappa statistic overcomes the insensitivity to imbalance, comparing the observed accuracy versus the expected accuracy that would result from random change. AUC provides the probability of correctly discriminating between classes for a randomly selected observation, and is therefore also sensitive to imbalance. The Brier score – which is simply the difference between predicted probability minus the actual outcome (1 or 0) squared – is threshold-agnostic, and therefore provides an indication of the quality of a model's predictions. For continuous predictions, we focus on the root mean squared error statistic (RMSE).

We find that the stacked ensemble (ENS) tends to outperform all other approaches. In particular, ENS has the lowest Brier score and the highest Kappa (at the optimal decision threshold). The benefits of the ensemble are evident when looking at the performance of the RF and XGB models, which are the top performing level one models on the Brier score and Kappa statistic respectively. The binary ensemble is able to encompass the relative benefits of both these approaches while overcoming the deficiencies of each (i.e. the poor Brier score for XGB, and relatively low Kappa for RF). We are also able to see why the ensemble combination is superior to a single ensemble where net wealth is transformed using the inverse hyperbolic sine (IHS), or when gross wealth and debts are modeled separately (WD). In these cases, overall accuracy is high but the approaches are completely insensitive to zero or negative wealth values, each completely missing true negative cases (i.e. TNR = 0%). Note that an ensemble which models the raw, untransformed value of net wealth performs similarly to when the inverse hyperbolic sine is taken. The same pattern holds for the positive wealth models – ENS has the lowest RMSE for positive wealth. The negative wealth ensemble model performs slightly worse than some of the level one learners.

To visually inspect predicted versus actual values, Figure 6 shows the performance of the full ensemble models separately for households with positive and negative wealth in the held-out SCF data (i.e. 10% of data). The figures show the predicted versus actual net wealth for households with a line of symmetry as the diagonal. There is an evident strong fit for households with positive wealth (91% of the sample; $RMSE = 0.99$– the stacked ensemble errors are symmetrical and highly accurate at the household-level. The fit is clearly less good

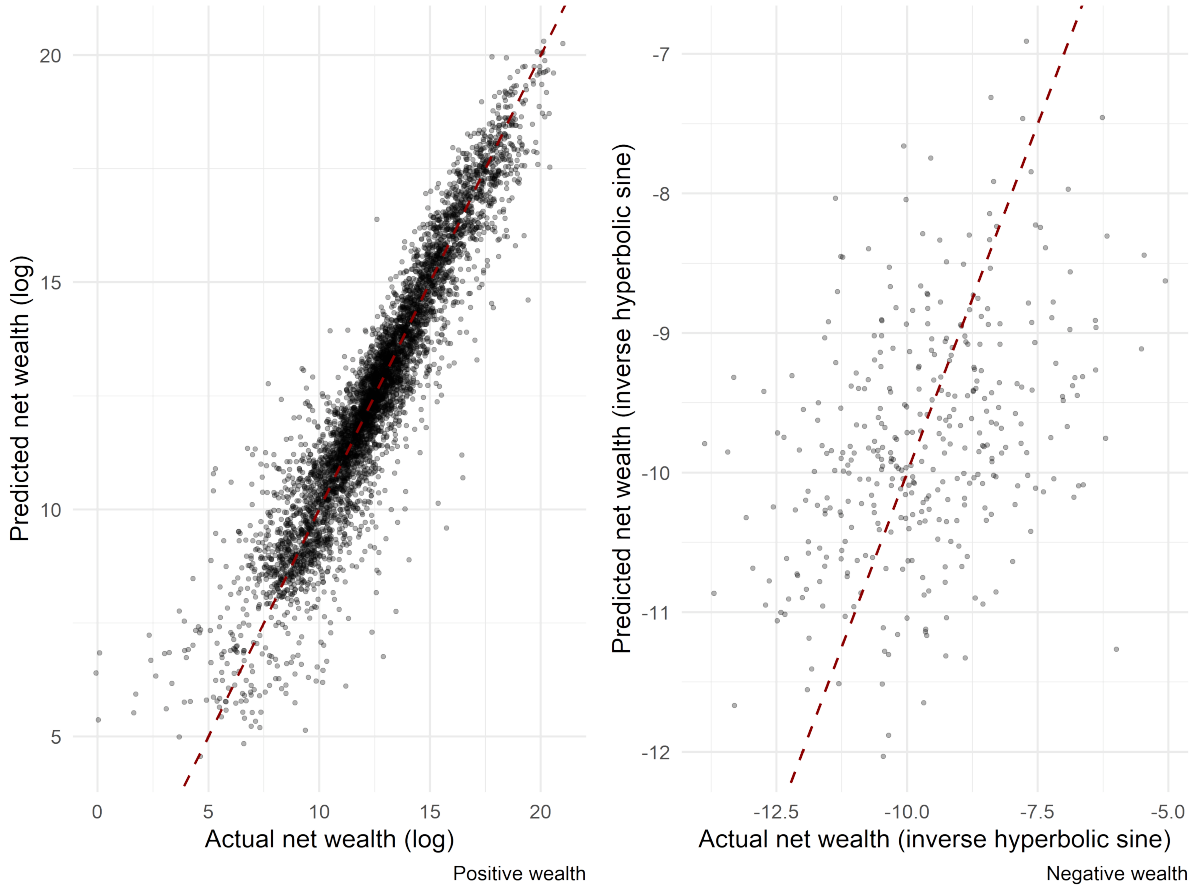

Figure 6: Test sample performance (SCF), positive and negative wealth stacked ensembles
Note: Separately for households with positive and negative wealth in the SCF test-sample data (N=5,341), this figure describes the correlation between actual and predicted values of net wealth. Root mean squared error (RMSE) for positive wealth estimate equal to 0.99. RMSE for negative wealth estimates is 1.48.

for those households with negative wealth (7.5% of sample; $RMSE = 1.48$) due to a relative lack of information which can be used to quantify negative wealth – e.g. income items provide relatively less information as to the quantum of negative wealth.

Predictive performance unsurprisingly degrades for models which are missing some key items, in particular housing tenure and value (missing from the 1950 census), and as income items become less detailed and coarser (for example, from 1940 to 1970, rather than being separately identified, investment income is incorporated into 'other' income). The relative decline is, however, quite small in size, even for the 1950 sample. See Table 4 for year-specific performance results for the 'winning' ensemble approach.

Performance evaluation, out of sample dataset (PSID)

To further validate the performance of the ensemble models, we next examine the fit on a completely separate dataset – the 2019 wave of the Panel Study of Income Dynamics (PSID). The PSID provides an identical set of variables to those used to build the models that are trained from the SCF. There is one important difference, however – the PSID definition of net wealth excludes the value of pensions while SCF includes it (Cooper et al., 2019). Figure 7 shows performance separately for households with positive and negative wealth. We see similar patterns as with the test sample – the performance is relatively stronger for positive wealth. There are higher absolute error levels for both, ($RMSE = 1.26$ and $RMSE = 1.77$ respectively)

Year	Brier	Kappa	AUC	Accuracy	TPR	TNR	RMSE (+)	RMSE (-)
2010-2020	0.058	0.426	0.901-0.921	0.887	0.910	0.637	0.999	1.480
1990-2000	0.058	0.408	0.891-0.913	0.894	0.925	0.554	1.002	1.497
1980	0.063	0.379	0.878-0.901	0.861	0.881	0.653	1.058	1.432
1970	0.063	0.389	0.877-0.9	0.860	0.876	0.685	1.151	1.445
1960	0.063	0.384	0.873-0.897	0.859	0.876	0.679	1.206	1.438
1950	0.066	0.335	0.84-0.87	0.853	0.878	0.588	1.581	1.462
1940	0.063	0.386	0.873-0.897	0.881	0.912	0.560	1.313	1.440

Note: Using performance metrics described in the text, this table compares performance of the ensemble combinations across different years, corresponding to different variable sets available in the Census/ACS.

Table 4: Performance of ensemble combination by year.

due to the clear overestimation of net wealth relative to the test model, and to be expected given PSID excludes pension wealth in its definition of net wealth.

Since PSID includes State identifiers, we can also assess the role of location in generating prediction errors. For each ensemble, we run a simple regression model in which a State identifier is the regressor and error as regressand. For binomial ensemble models, the regressand indicates either correct or inaccurate prediction; for ensembles predicting levels of positive or negative wealth, the dependent variable captures the residuals. In each of these regression estimates, the State predictor has negligible value in explaining the variation in errors, with Pseudo- and Adjusted-R-squared of less than 1%. Since neither the PSID nor any other known publicly-available data on wealth offers geographic identifiers below the level of an individual state, we are unable to assess our estimates of inequality at finer spatial scales. One of the main contributions of the GEOWEALTH database will be in enabling research into wealth dynamics at these finer spatial scales.

Validation against aggregate published measures of wealth inequality

As a further validation exercise, we compare aggregates of our imputed Census wealth data against widely-recognized published data. Available data enables such comparisons to national- and state-level indicators.

The top panel of Figure 8 compares two estimates of changes in national wealth inequality. The solid line, labelled 'Ensemble' presents a series of Gini coefficients generated from our imputed Decennial and ACS wealth data. The dashed line is a series of national-level Gini coefficients estimated using distributional macroeconomic accounts, obtained from Saez and Zucman (2020). Note that, while our base units are households, Saez and Zucman (2016, 2020) use, respectively tax-units, based on capitalizing income reported to the tax authorities, and individuals. The dotted line presents estimates from Kuhn et al. (2020), which uses a harmonized version of the SCF. The bottom panel compares our estimates of the top 1% share of total wealth, the top 10% share, and the bottom 50% share with those provided by Saez and Zucman (2016, 2020).

Our national estimates and those of both Saez and Zucman and Kuhn et al, while differing slightly in levels, broadly agree in terms of trends. This provides support for the idea that our estimates derived from imputing household wealth in the census is picking up important aggregate level dynamics. There are some differences to highlight. The declines in inequality in the Gini and top 1% share from 1960 to 1980 that we estimate lie somewhere between those described in Saez and Zucman (2020) and Kuhn et al. (2020). Meanwhile, though Kuhn et al suggest that Gini coefficients capturing wealth inequality have grown between 2010 and 2016, both Saez and Zucman (2020) and our ensemble estimates indicate a moderate decline in overall national wealth inequality over the last decade. Differences between our ensemble estimates and those of Saez and Zucman (2020) appear to be driven wealth at and above the top one percent.

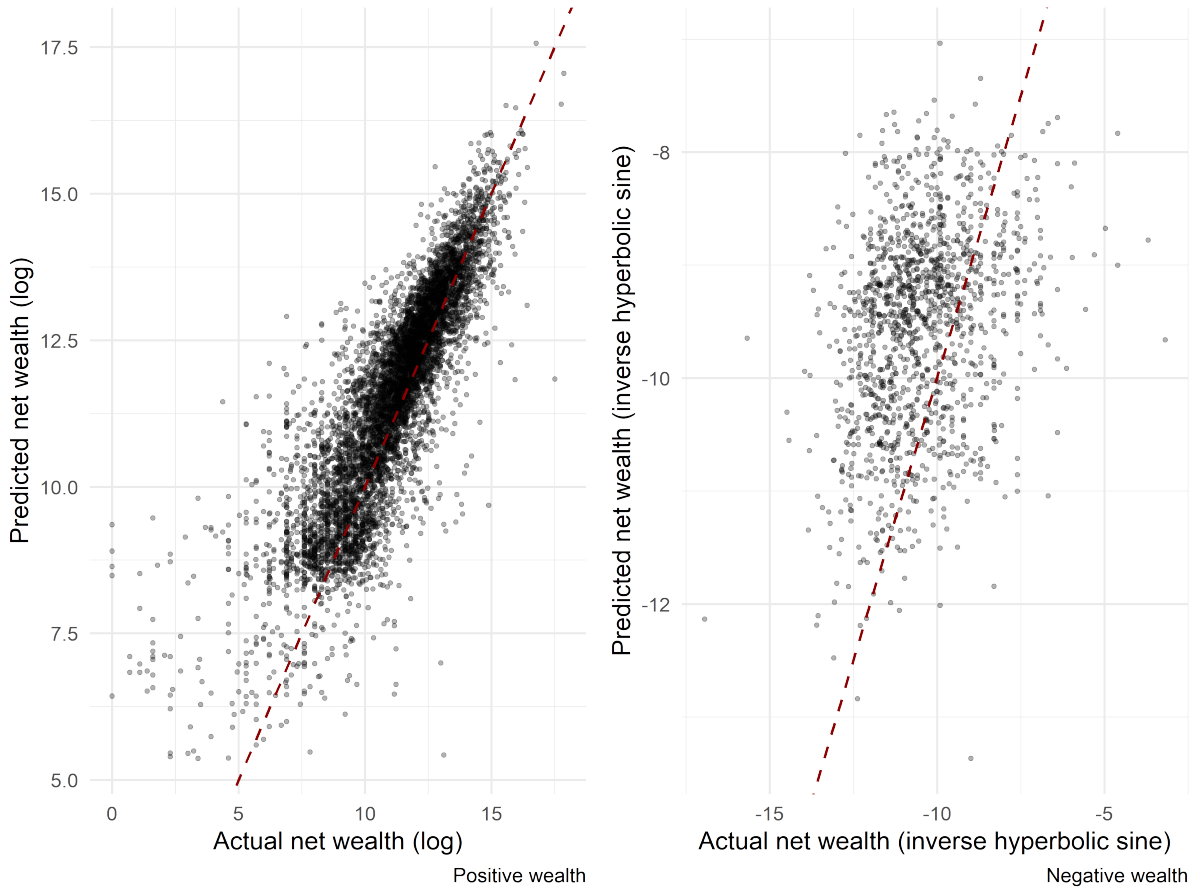

Figure 7: Out of sample performance (PSID), positive and negative wealth stacked ensembles
Note: Separately for households with positive and negative wealth in 2019 Panel Study of Income Dynamics (PSID) data, this figure describes the correlation between actual net wealth and predicted values using our ensemble model. Root mean squared error (RMSE) for positive wealth estimate equal to 1.26. RMSE for negative wealth estimates is 1.77.

For recent years, the U.S. Census Bureau's Survey of Income and Program Participation (SIPP) – a nationally representative survey tracking household economic outcomes and government program participation – provides estimates of mean and median wealth at the state-level. We are thus able to correlate our imputed state-level estimates with those provided by SIPP.

The scatterplots in Figure 9 describe the relationship between SIPP state-level measures of mean and median wealth and our Census imputations in 2020. These measures are strongly correlated at both the mean ($r = 0.86$) and the median ($r = 0.88$). This high level of correspondence provides further strong validation for our imputed estimates, this time at a subnational scale.

Descriptive analysis

Having described the process to build the GEOWEALTH database and our external validation procedures, we conclude by presenting some key patterns in the dataset that can inform future research. We do so by focusing on two separate indicators of wealth inequality, one that captures wealth concentration within commuting zones and one that measures wealth differences between commuting zones.

To examine how commuting zones have been changing with respect to their wealth levels

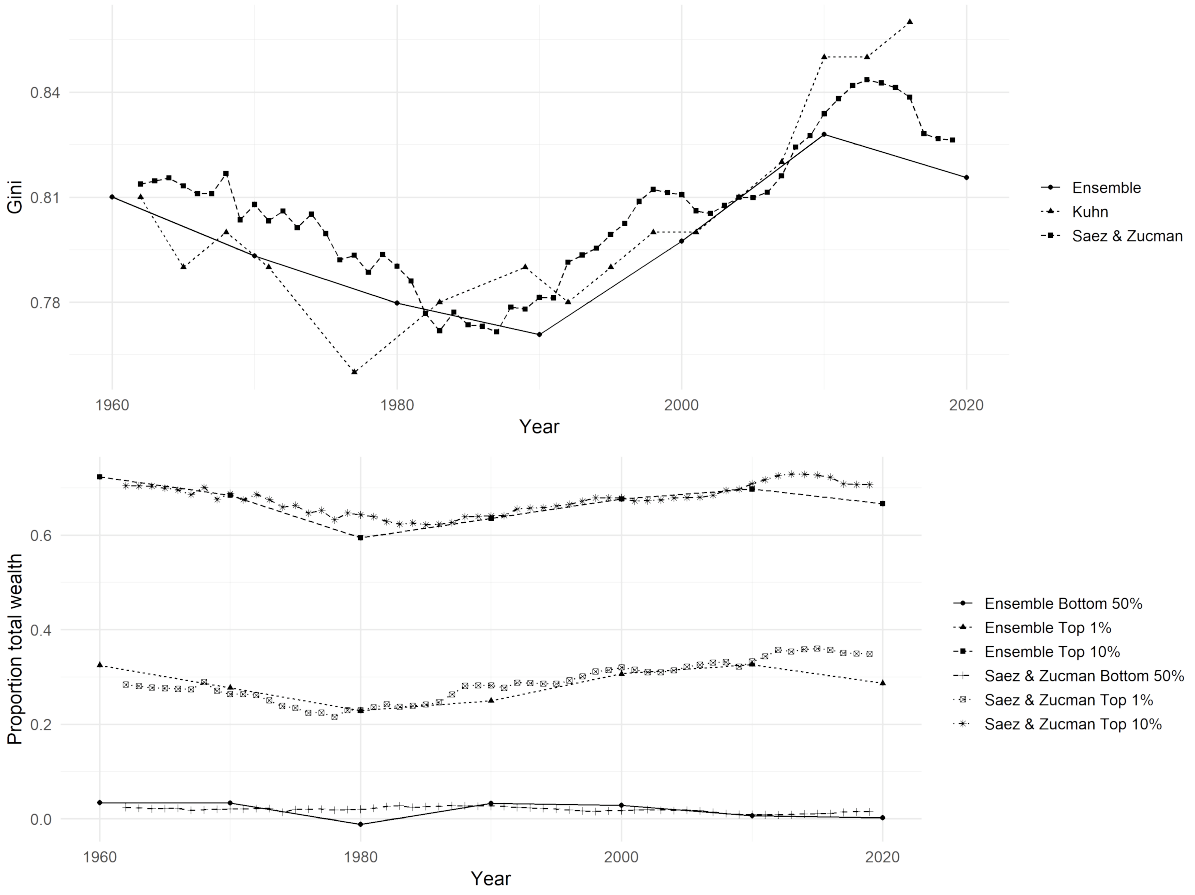

Figure 8: US wealth inequality, comparison of estimates

Note: The top panel compares estimates of national wealth inequality generated using our ensemble model to those that emerge from the distriburtional national accounts method in Saez and Zucman (Saez and Zucman, 2020), as well as the harmonized version of the SCF generated by Kuhn et al (Kuhn et al., 2020), a series that ends in 2016.

over the past 60 years, Figure 10 maps average wealth in 1960 and 2020. Given the large increases in average wealth over the study period, we standardize average wealth into period-specific z-scores, which have a mean of zero and a standard deviation of one. This metric provides a relative sense of the deviation of wealthier and poorer commuting zones from the average commuting zone in each period.

In 1960, commuting zones in traditional industrial regions of country exhibit the highest average levels of wealth. High wealth level are evident across the Northeast, the greater Chicago region, and in the Sunbelt in Southern California and Florida. These patterns closely track well known patterns of early- to mid-twentieth century industrialization and urbanization (Lindert and Williamson, 2017).

While patterns of average wealth in 2020 bear some resemblance to those in 1960, several important differences are evident. Specifically, the advantages of many once wealthy manufacturing regions have regressed toward the mean. This is particularly notable for the metropolitan areas around the Great Lakes such as Buffalo, Cleveland, Chicago, and Milwaukee. In their place, Pacific cities such as Seattle, Los Angeles, San Francisco, interior regions like Denver, and the major Texan metropolises have decidedly improved their relative wealth positions. Although the South as a whole continues to lag the rest of the country in terms of average wealth, Savannah (GA), Raleigh (NC), and Nashville (TN) are examples of Southern commuting zones that have seen substantial growth in average wealth levels.

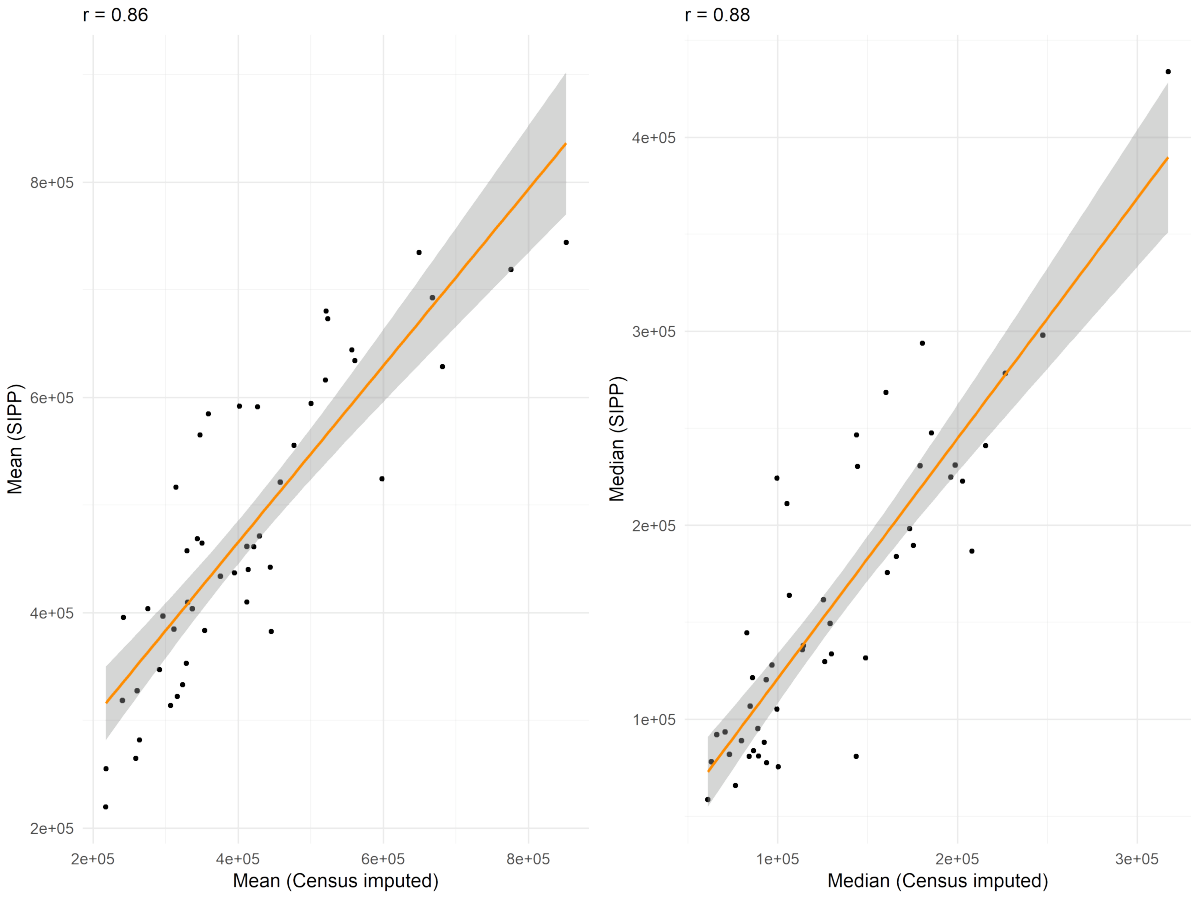

Figure 9: US state-level comparisons, mean and median wealth from imputed estimates and SIPP (2020)

Note: Panels compare state-level mean and median wealth, using estimates generated from our ensemble model and those obtained from the 2020 Survey of Income and Program Participation (SIPP).

As noted above, the changing geography of wealth over this period is also characterized by an intensification of inequality between regions. This is evident in Figure 11, where we plot the trajectories of relative wealth for commuting zones across each decade. Most clearly, this figure exhibits a pattern of fanning out, implying rising levels of inter-regional wealth inequality since 1960. This means that the average wealth gaps between the wealthiest commuting zones and the average commuting zone are significantly larger in 2020 than they were in 1960. For example, Boston and Chicago, which were among the top five wealthiest commuting zones in 1960, had average wealth levels that were approximately 3 to 4 standard deviations above the average. In 2020, however, the average wealth levels of San Jose and San Francisco - the two wealthiest commuting zones today - are 5 to 6 standard deviations above the mean. Preliminary investigation of the GEOWEALTH database therefore reveals that the wealthiest regions have been pulling away from the rest of the country since 1960.

Finally, we turn our attention to the changing dynamics of wealth inequality within regions over time. Figure 12 maps the Gini coefficients for wealth inequality within commuting zones in 1960 and 2020, revealing patterns of change and stability. In 1960, intra-regional wealth was high throughout the South, low in the Midwest and Northern Plains regions, and more mixed along the coasts. The main change to this pattern up to 2020 has, however, been the dramatic rise in inequality in the Midwest and Plains regions. While the South persists as a region that broadly exhibits high inequality, central and formerly manufacturing-dependent Midwestern

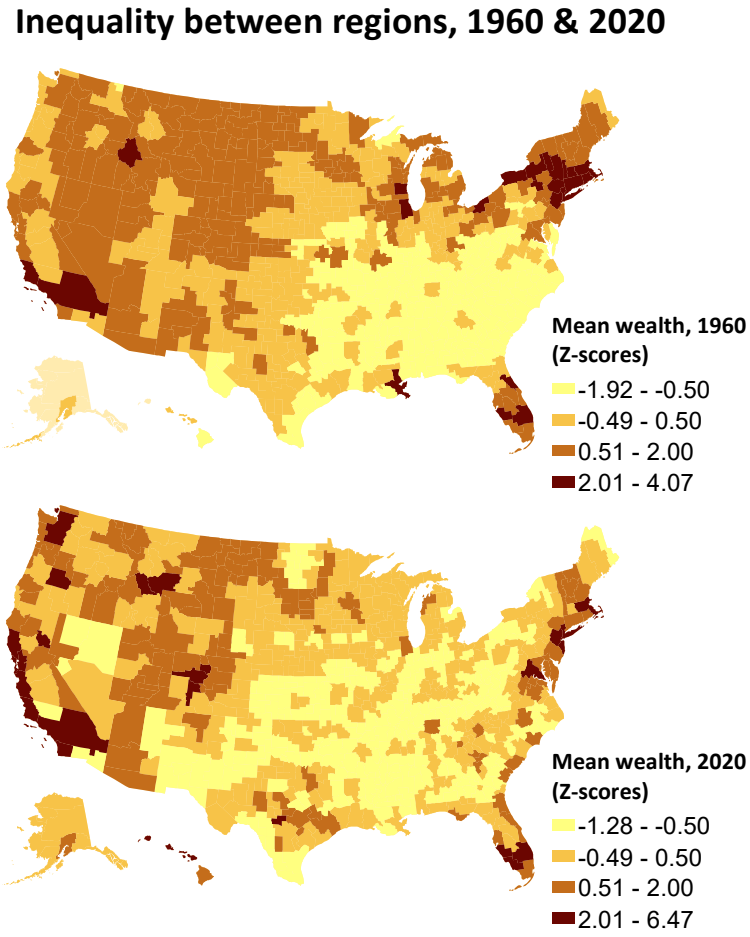

Figure 10: Average wealth levels across commuting zones, 1960 & 2020
Note: For 1990-vintage commuting zones (CZs) as in Tolbert and Sizer (Tolbert and Sizer, 1996), this figure maps mean wealth levels in 1960 and 2020. Local average wealth levels are converted to z-scores such that the distribution has a mean of zero and a standard deviation equal to one.

regions have seen a substantial worsening of inequality over this period. This convergence in terms of inequality levels between the South and the Midwest is consistent with findings from for other studies of inequality such as intergenerational mobility and earnings levels (Kemeny and Storper, 2022; Connor and Storper, 2020a).

We thus present the GEOWEALTH database as a new source of information that can considerably expand the frontier of research into long-term inequality and economic prosperity. Over recent years, our ability to study long-term patterns of spatial inequality have been greatly enhanced through other related efforts. Recent efforts have focused on generating related databases that track leading indicators of inequality such as urbanization, patenting, incomes, and intergenerational mobility within consistent spatial units over long periods of time (Connor and Storper, 2020a; Leyk et al., 2020; Uhl et al., 2023, 2021; Petralia et al., 2016). In the same vein, the GEOWEALTH database provides a first important step toward understanding the long-term spatial dynamics of wealth inequality.

Usage Notes

Subnational-scale GEOWEALTH data are available on an open access basis through a Creative Commons Attribution 4.0 International (CC BY 4.0) license. The data are hosted by the Inter-

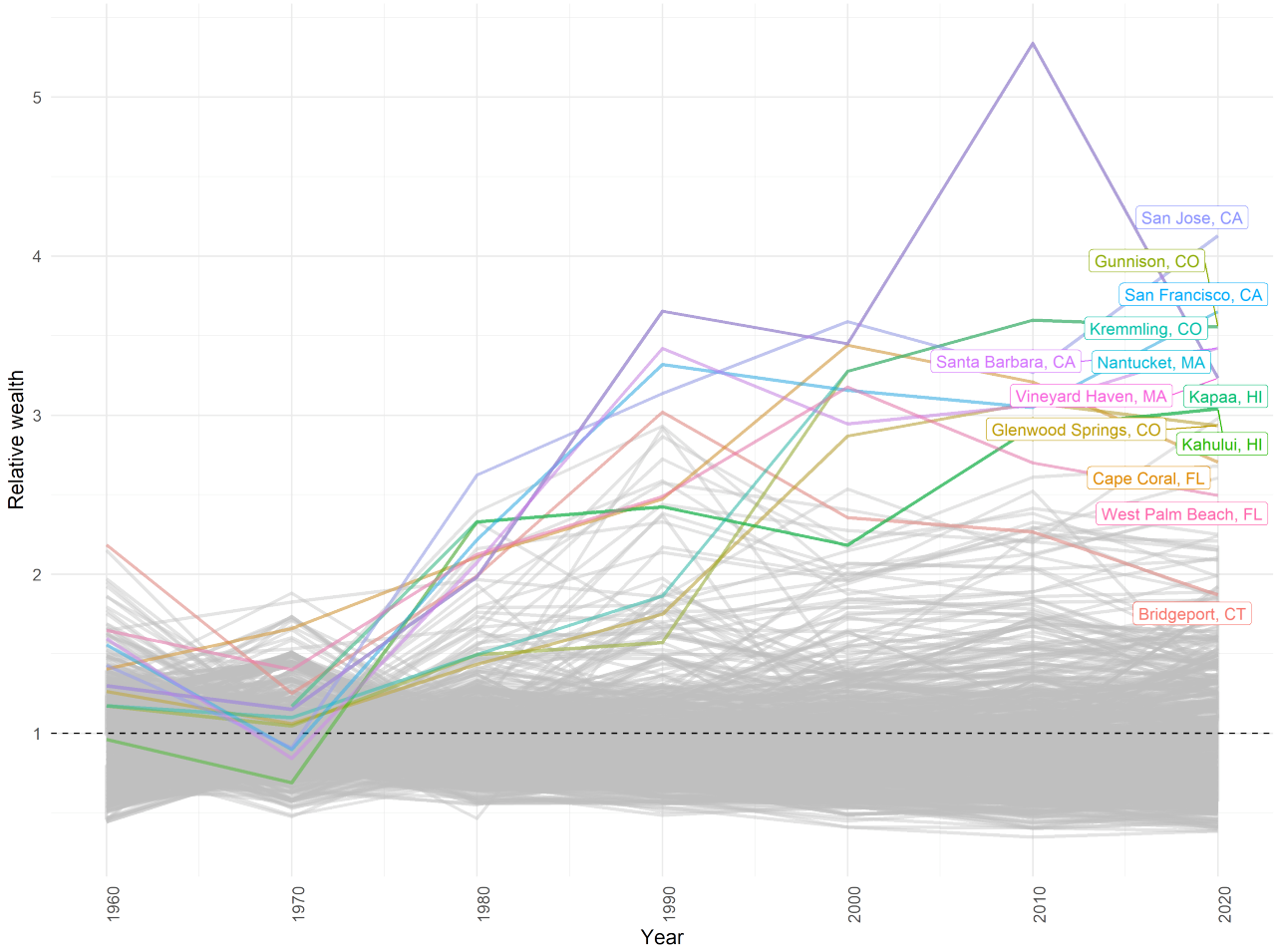

Figure 11: Relative wealth over time, U.S. Commuting Zones

Note: Each line in this figure represents a particular U.S. Commuting Zone (CZ), defined according to 1990 boundaries as per (Tolbert and Sizer, 1996). In each observed year, the Y-axis measures the ratio of average wealth for a CZ to the all-locations average wealth.

university Consortium for Political and Social Research (ICPSR) at Project #192306. Any use of this 1960–2020 compilation of data should be accompanied by a citation of this paper, in addition to a proper use of DOI-reference (doi.org/10.3886/E192306V1) and citation of the actual data.

Data are organized into a series of individual comma-separated files (csv), with each file corresponding to a particular spatial unit of observation: state; 1990-vintage commuting zone; metropolitan area; PUMA; and division. Coverage over time is dependent on spatial units, per Table 5. The data at ICPSR include a brief metadata file in pdf format.

Primary variables included in each dataset include locational identifiers (codes and names of places); the number of Census households surveyed in that location; measures of central tendency (means and medians) for wealth; measures of spread for wealth (standard deviation); ratios of wealth at specific percentiles (for instance the ratio of wealth at the 90th/50th percentiles); as well as selectec key demographic features of locations, derived from the Census. For wealth estimates, we also include upper and lower bounds, based on the bootstrapping procedure described in the text.

We view this dataset as a necessary first step towards the study of spatial wealth disparities. While one would ideally want directly observed, geocoded data on household wealth in the United States and how it has changed, such data are unlikely to become available in the near

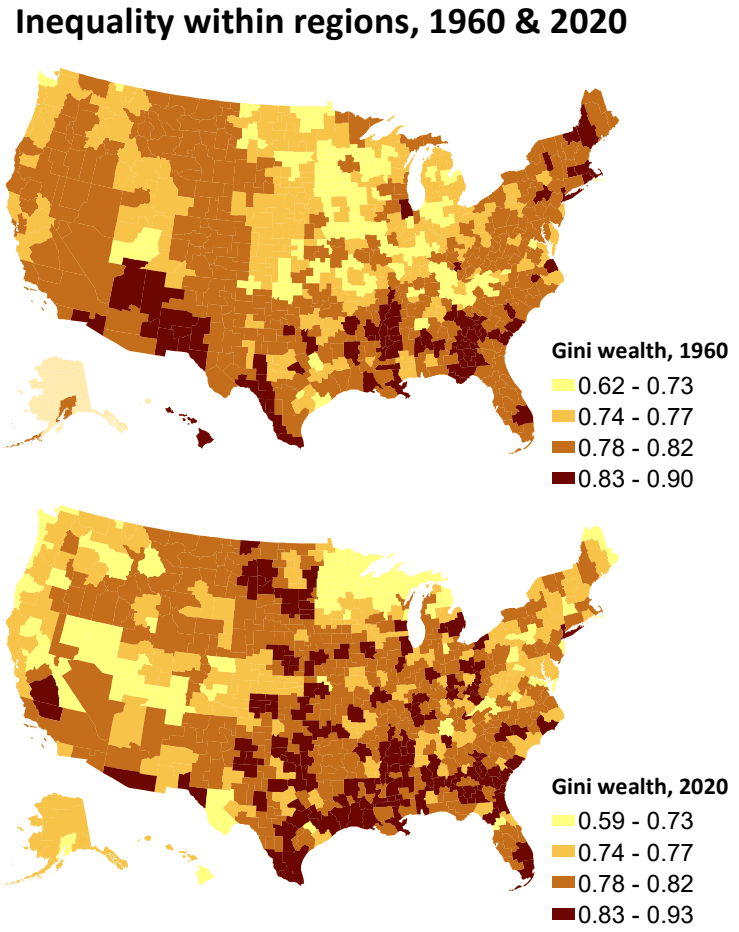

Figure 12: Gini coefficients for wealth distribution within commuting zones, 1960 & 2020
Note: For 1990-vintage commuting zones (CZs) as in Tolbert and Sizer (Tolbert and Sizer, 1996), this figure maps Gini coefficients tracking within-CZ wealth inequality in 1960 and 2020.

future. These data could be used as a foundation to explore a wide range of questions related to the causes and consequences of geographic variation in wealth and wealth inequality, which heretofore have been impossible to explore. Just as there is a growing literature on community and spatial effects of differences in income and poverty, these data provide a basis to explore related questions around wealth.

Potential users of these data should be aware of limits on coverage of the extremely wealthy. One longstanding strand of research (i.e., Piketty and Saez, 2006; Atkinson and Piketty, 2007) uses administrative data to explore the evolution of top incomes – commonly focusing on individuals at or above the top 1 or 0.1 percentile. While the SCF offers improved coverage of middle class assets like housing that do not generate flows of taxable income (Kuhn et al., 2020), only more recent SCF permit more careful modeling of top incomes (Bricker et al., 2016), and even then there may be limits on analysts' ability to explore very top incomes using our data. Indeed, when we look at temporal trends in the share of wealth accruing to the top 0.1% and top 0.01% of wealth holders, we do not see the same large increase in recent decades as is evident in other data, such as (Saez and Zucman, 2020, 2016), although our estimates of the top 0.1% are close to that provided by the Fed using distributional financial accounts (DFA).[3] For these

[3]Federal Reserve Board data on wealth shares using DFA is available here: https://www.federalreserve.gov/releases/z1/dataviz/dfa/distribute/table/.

Dataset Name	Spatial Units	Unique Units	Observations	Years
puma_wealth_inequality.csv	Public-use micro areas	7428	8641	1960–2020
cz_wealth_inequality.csv	Commuting Zones	741	5174	1960–2020
metarea_wealth_inequality.csv	Metropolitan areas	564	1226	1960–2020
state_wealth_inequality.csv	States	51	353	1960–2020
division_wealth_inequality.csv	Divisions	9	70	1960–2020

Table 5: Datasets and coverage

comparisons, see Figures 13 and 14.

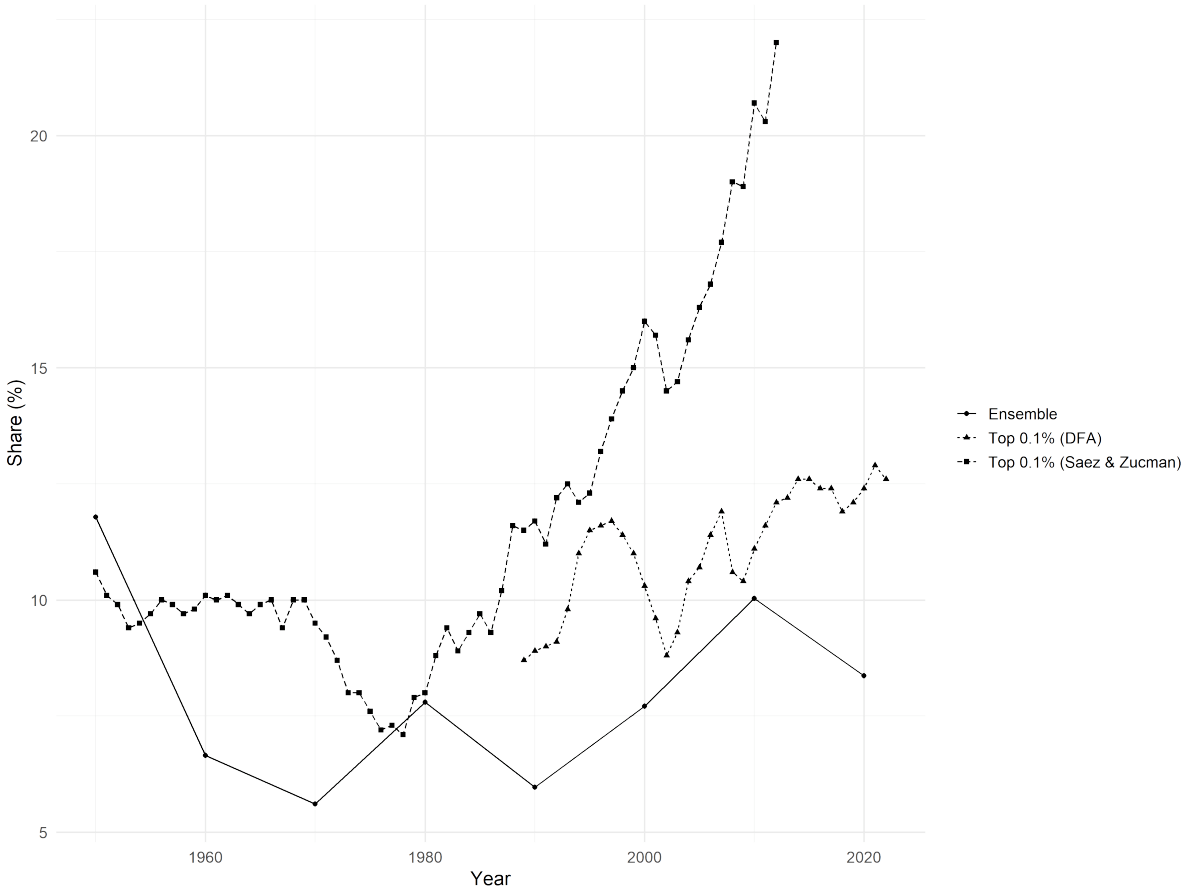

Figure 13: Comparison of national-level estimates of the wealth share of the top 0.1%
Note: For measures of the share of total national wealth held by the top 0.1% of the distribution, this figure compares estimates generated using the modeling techniques described in this paper ('Ensemble') to existing data, specifically the Federal Reserve's Distributional Financial Accounts, as well as Saez & Zucman (2016) (Saez and Zucman, 2016).

Code availability

All replication code is available at github.com/jhsuss/wealth-inequality.

References

Acolin, A. and Wachter, S. (2017). Opportunity and housing access. *Cityscape*, 19(1):135–150.

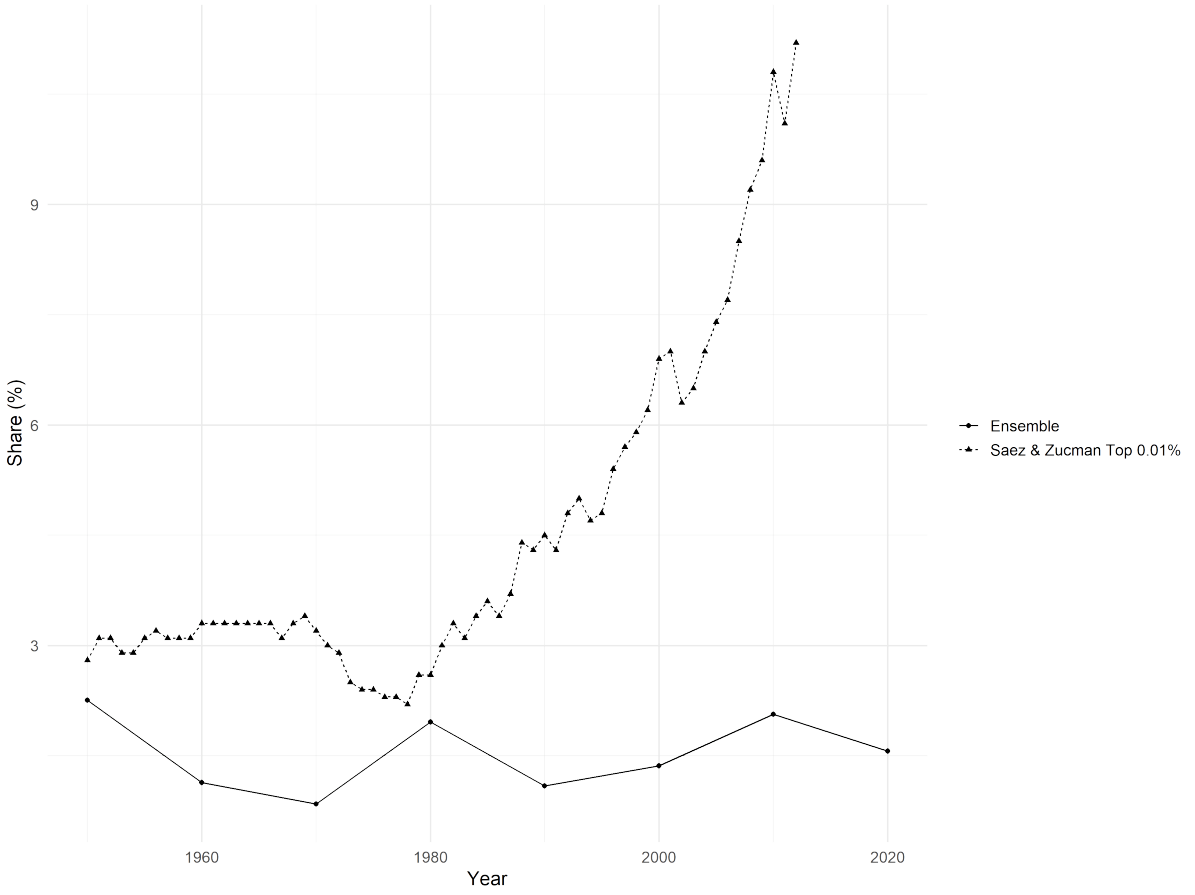

Figure 14: Comparison of national-level estimates of the wealth share of the top 0.01%
Note: For measures of the share of total national wealth held by the top 0.01% of the distribution, this figure compares estimates generated using the modeling techniques described in this paper ('Ensemble') to existing data, specifically Saez & Zucman (2016) (Saez and Zucman, 2016).

Atkinson, A. B. and Piketty, T. (2007). *Top incomes over the twentieth century: a contrast between continental European and English-speaking countries*. Oxford University Press.

Baumgärtner, S., Drupp, M. A., Meya, J. N., Munz, J. M., and Quaas, M. F. (2017). Income inequality and willingness to pay for environmental public goods. *Journal of Environmental Economics and Management*, 85:35–61.

Breiman, L. (1996a). Bagging predictors. *Machine learning*, 24:123–140.

Breiman, L. (1996b). Stacked regressions. *Machine learning*, 24(1):49–64.

Bricker, J., Henriques, A., Krimmel, J., and Sabelhaus, J. (2016). Measuring income and wealth at the top using administrative and survey data. *Brookings papers on economic activity*, 2016(1):261–331.

Broz, J. L., Frieden, J., and Weymouth, S. (2021). Populism in place: the economic geography of the globalization backlash. *International Organization*, 75(2):464–494.

Burkhauser, R. V., Feng, S., Jenkins, S. P., and Larrimore, J. (2011). Estimating trends in us income inequality using the current population survey: the importance of controlling for censoring. *The Journal of Economic Inequality*, 9:393–415.

Chetty, R. (2021). Improving equality of opportunity: New insights from big data. *Contemporary Economic Policy*, 39(1):7–41.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2017). Mobility report cards: The role of colleges in intergenerational mobility. National bureau of economic research Working Paper 23618.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of urban economics*, 63(2):723–742.

Connor, D. S. and Storper, M. (2020a). The changing geography of social mobility in the united states. *Proceedings of the National Academy of Sciences*, 117(48):30309–30317.

Connor, D. S. and Storper, M. (2020b). The changing geography of social mobility in the United States. *Proceedings of the National Academy of Sciences*, 117(48):30309–30317.

Cooper, D., Dynan, K. E., and Rhodenhiser, H. (2019). Measuring household wealth in the panel study of income dynamics: The role of retirement assets. Federal Reserve Bank of Boston Working Paper.

Côté, S., House, J., and Willer, R. (2015). High economic inequality leads higher-income individuals to be less generous. *Proceedings of the National Academy of Sciences*, 112(52):15838–15843.

Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2019). Income growth and the distributional effects of urban spatial sorting. National Bureau of Economic Research Working Paper 26142.

Cowell, F., Nolan, B., Olivera, J., and Van Kerm, P. (2017). Wealth, top incomes and inequality. In *National Wealth: What is missing, why it matters*, pages 175–206. New York: Oxford University Press.

Cramer, K. J. (2016). *The politics of resentment: Rural consciousness in Wisconsin and the rise of Scott Walker*. University of Chicago Press.

Dorn, D. (2009). *Essays on inequality, spatial interaction, and the demand for skills*. PhD thesis, University of St. Gallen.

Efron, B. (1992). *Bootstrap methods: another look at the jackknife*. Springer.

Fichtenbaum, R. and Shahidi, H. (1988). Truncation bias and the measurement of income inequality. *Journal of Business & Economic Statistics*, 6(3):335–337.

Ganong, P. and Shoag, D. (2017). Why has regional income convergence in the US declined? *Journal of Urban Economics*, 102:76–90.

Goldin, C. and Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.

Gyourko, J., Mayer, C., and Sinai, T. (2013). Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–99.

Hansen, M. N. (2014). Self-made wealth or family wealth? changes in intergenerational wealth mobility. *Social Forces*, 93(2):457–481.

Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Kemeny, T. and Storper, M. (2022). The changing shape of spatial inequality in the United States. SocArXiv Working Paper DOI: 10.31235/osf.io/wnd8t.

Killewald, A., Pfeffer, F. T., and Schachner, J. N. (2017). Wealth inequality and accumulation. *Annual review of sociology*, 43:379–404.

Kuhn, M., Schularick, M., and Steins, U. I. (2020). Income and wealth inequality in America, 1949–2016. *Journal of Political Economy*, 128(9):3469–3519.

Leyk, S., Uhl, J., Connor, D., Braswell, A., Mietkiewicz, N., Balch, J., and Gutmann, M. (2020). Two centuries of settlement and urban development in the United States. *Science Advances*, 6(23):eaba2937.

Lindert, P. H. and Williamson, J. G. (2017). *Unequal Gains: American Growth and Inequality since 1700*. Princeton University Press.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Manduca, R. A. (2019). The contribution of national income inequality to regional economic divergence. *Social Forces*, 98(2):622–648.

Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2):373–77.

Neckerman, K. M. and Torche, F. (2007). Inequality: Causes and consequences. *Annu. Rev. Sociol.*, 33:335–357.

O'Brien, D. T. (2022). *Urban Informatics: Using Big Data to Understand and Serve Communities*. CRC Press.

Petralia, S., Balland, P. A., and Rigby, D. L. (2016). Unveiling the geography of historical patents in the united states from 1836 to 1975. *Scientific data*, 3(1).

Piketty, T. and Saez, E. (2006). The evolution of top incomes: a historical and international perspective. *American economic review*, 96(2):200–205.

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics*, 133(2):553–609.

Piketty, T. and Zucman, G. (2015). Wealth and inheritance in the long run. In *Handbook of income distribution*, volume 2, pages 1303–1368. Elsevier.

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, 11(1):189–209.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

Ruggles, S., Flood, S., Goeken, R., Schouweiler, M., and Sobek, M. (2022). Ipums usa: Version 12.0 [dataset]. Minneapolis, MN, https://doi.org/10.18128/D010.V12.0.

Saez, E. and Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2):519–578.

Saez, E. and Zucman, G. (2020). The rise of income and wealth inequality in america: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives*, 34(4):3–26.

Sampson, R. (2019). Neighbourhood effects and beyond: Explaining the paradoxes of inequality in the changing american metropolis. *Urban Studies*, 56(1):3–32.

Song, X., Massey, C. G., Rolf, K. A., Ferrie, J. P., Rothbaum, J. L., and Xie, Y. (2020). Long-term decline in intergenerational mobility in the united states since the 1850s. *Proceedings of the National Academy of Sciences*, 117(1):251–258.

Tolbert, C. M. and Sizer, M. (1996). US commuting zones and labor market areas: A 1990 update. United States Department of Agriculture, Staff report.

Uhl, J., Connor, D., Leyk, S., and Braswell, A. (2021). A century of decoupling size and structure of urban spaces in the United States. *Communications earth & environment*, 2(1):1–14.

Uhl, J., Hunter, L. M., Leyk, S., Connor, D. S., Nieves, J. J., Hester, C., Talbot, C., and Gutmann, M. P. (2023). Place-level urban–rural indices for the united states from 1930 to 2018. *Landscape and Urban Planning*, 236(104762).

Yellen, J. L. (2014). Perspectives on inequality and opportunity from the survey of consumer finances. Speech to the Conference on Economic Opportunity and Inequality, Federal Reserve Bank of Boston.

Zhou, Z.-H. (2012). *Ensemble methods: foundations and algorithms*. CRC press.