A procedure for generating locally representative estimates of household wealth that support flexible disaggregation

Robert Habans and Haleigh Tomlin, The Data Center

April, 2024

This paper outlines a procedure for estimating household wealth at a local (sub-state) level using data from federal surveys. The procedure supports relatively flexible breakdowns by race and other population subgroups, and the results retain the skewed distribution of wealth in the underlying survey data. Using the Survey of Income and Program Participation and the American Community Survey, the procedure builds on a multi-level regression and poststratification (MRP) framework, where predictive models are fit to national data and then poststratified to a local population. The results can be used for descriptive summaries of household wealth in metropolitan areas or large localities (as in a companion report on the New Orleans metro area, for which this paper serves as extended technical documentation), as well as other applications.

Acknowledgements

This project was supported in part by the Black Wealth Data Center (BWDC). The BWDC – launched by Bloomberg Philanthropies' Greenwood Initiative – provides public and private sector leaders with data disaggregated by race to help advance racial wealth equity. The project also received generous support from the Urban League of Louisiana, the W.K. Kellogg Foundation, and the Kresge Foundation.

Introduction

Despite widespread interest in the racial wealth gap and the concentration of wealth of the United States, seekers of readymade data on wealth that reflects conditions in local areas will typically come up empty-handed. The national surveys most often used to explore the distribution and composition of wealth (e.g., the Survey of Consumer Finances) do not support state-level estimates, and even where state-level estimates are technically possible (e.g., the Survey of Income and Program Participation), small samples render further disaggregation challenging. Sub-state estimates are even more restrictive, leaving racial wealth inequality and other related topics difficult to examine through a local lens. This paper documents a procedure to estimate wealth at a local level, such as a state, metropolitan area, or other sub-state geography with a reasonably large population. The procedure supports relatively flexible disaggregation by race and other demographic and socioeconomic characteristics. While our immediate objective is to provide detailed estimates for the New Orleans metro area, the procedure is adaptable to other local areas.

The local estimates combine information from multiple publicly available federal data sources. In a similar vein, a handful of well-documented local wealth estimates have been published previously (e.g., in the methodology for the Prosperity Now Scorecard, a technical paper accompanying the Urban Institute's Financial Health and Wealth Dashboard, and a working paper by the Census Bureau), with each approach varying in important ways (Chenevert et al. 2017; Williams, Zhong, and Braga 2023; Prosperity Now 2021). Our approach also has similarities and differences with previous estimates, as well as unique implications for how the estimates may be used to explore wealth inequality.

Our estimation procedure is adapted from a Bayesian multi-level regression and poststratification (MRP) workflow (for an accessible introduction, see Lopez-Martin, Phillips, and Gelman (2022)). A model is first fit to a national data set derived from the Survey of Income and Program Participation (SIPP) to predict household assets and debts (following convention, we define wealth as net worth, or the sum of all assets minus the sum of all debts). Parameter estimates from this SIPP-based model are then used to generalize, or post-stratify, the predictions to reflect the distribution of the local population. The local population distribution is estimated with American Community Survey microdata (ACS-PUMS), which has a much larger sample and includes more detailed geographic identifiers. ACS overlaps with SIPP in many common demographic and socioeconomic variables, which also tend to be associated with wealth. While SIPP does not support estimates at geographies smaller than the state-level, ACS does not include direct information on household wealth. In a sense, MRP overcomes the gaps in each data set by combining information about wealth from SIPP with the larger samples and more detailed geographic information of ACS. However, this effort involves many decisions and tradeoffs in pre-processing the SIPP and ACS data, specifying a predictive model, and post-processing the results into locally representative estimates.

In these decisions, we have prioritized flexibility to produce a variety of breakdowns relevant to examining the racial wealth gap. This priority drives several distinguishing aspects of our workflow. We fit a Bayesian model with varying intercepts, allowing for flexibility in how uncertainty can be quantified and results can be post-processed into local subgroup estimates. In addition to a conventional set of demographic and socioeconomic status predictors (race, age, education, homeownership, etc.), our model also leverages comparable information in both SIPP and ACS on additional factors that are plausibly associated with wealth, like the presence of special types of non-wage income (e.g., public assistance). We include race/ethnicity interaction terms to reduce biases that can result when quantitative models include a homogeneous effect of race/ethnicity. In addition to deploying a conventional MRP workflow to generate estimates for population subgroups, we also develop a modified post-processing workflow to generate estimates of the race-specific distribution of wealth. Finally, due to the skewed distribution of wealth, we estimate percentile ranking in the national distribution of assets and debts, rather than estimating assets and debts directly.

The body of this paper elaborates on these aspects of our approach and describes the results. The first section describes our motivations and summarizes the procedure. The second section describes the data sources and how they are pre-processed for MRP. The third section specifies a multi-level regression model for household assets and debts. The fourth section describes how the model results are post-processed to generate local breakdowns. The fifth section uses graphical checks and examples to illustrate the results. The paper closes with a brief discussion, highlighting limitations, potential applications, and avenues for future enhancement. We note that, while the content of this paper has a technical focus, a companion brief focuses on interpreting estimates for the New Orleans metro for a more general audience. The companion brief may be more interesting to end-users, with this paper functioning as extended documentation.

Background: Why and how we generate local wealth estimates

Motivations

While the main task of this paper is to document a procedure for generating local, disaggregated estimates of wealth, we offer four practical motivations for context. First, locally representative estimates can play a role in motivating action to address the racial wealth gap, wealth inequality, and financial well-being. While these issues are often reported at the national level, locally representative estimates can resonate in a more direct way with audiences whose interest is primarily local.

Second, disaggregated estimates can help to inform interventions to build wealth or to reduce wealth disparities by enhancing the baseline understanding of differences. Beyond simple racial disparities, more detailed estimates for population subgroups may lead to more targeted planning and resource allocation for local policies, programs, or other interventions, making estimates more informative for local efforts to address the racial wealth gap. For example, information on differences in wealth over the lifetime and by socioeconomic status subgroups may help to design more efficient, effective, and equitable wealth-building interventions – or at least to frame such interventions in a more holistic manner. We also believe there may be potential applications for disaggregated estimates to support distributional analysis, such as the assessment of policies where the goal is not putatively related to building wealth but where the costs and benefits may be distributed unevenly with respect to wealth. These potential applications make a case for flexible disaggregation.

These two motivations serve as the most immediate impetus for the project. The need for local estimates emerged from an engagement between The Data Center and the Urban League of Louisiana to provide data and research support for SEE CHANGE, an initiative to build Black and Latino wealth in the Greater New Orleans area. SEE CHANGE is organized around three "levers": income and wages, homeownership, and businesses and entrepreneurship. Because of the lack of locally relevant estimates of the wealth gap, SEE CHANGE lacks a concise, single-sentence problem statement that could serve as a rallying cry to motivate action. Likewise, efforts to plan and design components of SEE CHANGE are limited by a lack of information to inform a multi-pronged initiative. Our immediate interest is to address this information gap.

A third, broader motivation is to contribute to like-minded efforts to generate usable local estimates of wealth from limited available data. While some aspects of our approach are novel, we are aware of existing projects that have applied the same general idea: using a predictive regression to combine information from SIPP and ACS into local estimates of wealth. Each of these approaches, including ours, has limitations. The absence of direct data has no perfect solution, and any approach to overcome this limitation has tradeoffs that may or may not be tenable for a given use case. Nonetheless, the demand is evident. In the spirit of meeting this demand, we argue that transparent documentation of alternative pathways to similar ends can lead to incrementally more reliable, more usable, and more flexible data products – especially when the cost and complexity of collecting new data is prohibitive. The procedure described below is best read as a background reference for those who might take up a similar challenge in the future.

Fourth and finally, we hope that incremental enhancements to the availability of disaggregated local estimates of wealth can lay the groundwork for deeper, truly evidence-based efforts to reduce racial wealth inequality, especially for actions at a local level. For example, the method used here strictly limits the availability of information to estimate a counterfactual for causal inference. By extension, the estimates likely cannot be used for evaluating interventions to grow wealth or to reduce wealth inequality, nor are they suited to the analysis of non-compositional factors that could explain local differences in wealth outcomes. Evaluation is further complicated by the fact that wealth accrues slowly over the lifetime, so exposure to an effective intervention at age 25 might not show measureable differences until many years later, even without limited data availability. As we believe in the importance of these kinds of causal questions, our optimism lies in the possibility of disaggregated data to point the way toward next-best, approximate indicators for intermediate wealth outcomes.

Summary of the procedure

We briefly summarize the procedure here and cover the details in the sections that follow. We use MRP to adjust for differences between a national sample population and a local sample population of interest. In our application, SIPP serves as a nationally representative sample, and our goal is to adjust these estimates to reflect the population of a local target area, defined as a grouping of one or more public-use microdata areas (PUMAs). In our example, the local target area is the New Orleans-Metairie MSA. The model adjusts for differences between the national sample and the local population by including population subgroup characteristics available in both surveys as predictor "levels" in a multi-level regression. The model uses varying intercepts to allow "partial pooling" of information across levels in the predictors, yielding more reliable estimates for local subgroups where SIPP observations are sparse and/or local geographies are unidentified (Gelman and Hill 2006; Lopez-Martin, Phillips, and Gelman 2022).

We estimate assets and debts separately. For each, we fit two models: one to classify households as having any assets (or debts) and one to predict where households are expected to fall in the national distribution of assets (or debts), conditional on having assets (or debts). We use ACS-PUMS to give reliable estimates of how the local population is distributed across each subgroup. These estimates are collected in a table (hereafter referred to as the "poststratification table") that reports the local target area's population estimate for every unique combination of predictor variables (hereafter, "cell"). We use the model results to predict wealth outcomes onto each cell in the poststratification table. Following a conventional MRP workflow, cell-level predictions are aggregated to yield estimates for a target subgroup – e.g., an estimate for the subgroup "Black homeowners aged 55-64 in Louisiana" is a weighted average of more specific cell-level predictions that roll up into the target subgroup. We also use an alternative post-processing workflow, primarily to estimate race-specific distributions of assets and debts. In essence, we use draws from the posterior probability distribution to simulate within-cell variability and constrain predictions to match the state-level distribution of assets estimated directly from the state-level SIPP sample. These procedures are repeated for assets and debts. The predicted rankings can be transformed to their dollar equivalents to calculate net worth.

We use the R statistical computing environment. To fit the multi-level models, we use the `stan_glmer` command in the `rstanarm` package (Goodrich et al. 2024). A list of additional open source R packages, along with a minimal working version of our code, will be made available online in a Github repository. All source data is freely available from the Census Bureau and IPUMS-USA (Ruggles et al. 2023).

Data

Data sources

Our main data source for wealth is SIPP, a nationally representative, longitudinal household survey administered by the Census Bureau. SIPP participants are interviewed monthly, and the survey reports income, employment, household composition, government program participation, financial status, and other household and family characteristics. Most importantly for our purposes, SIPP includes detailed information on assets and debts, asked at the end of the survey year. While the Survey of Consumer Finances is generally considered to be a better source of information on wealth, SIPP has a larger sample size and includes identifiers for state and metropolitan status, though it does not include identifiers for lower levels of geography (e.g., CBSA or county).

Since 2020, the SIPP survey has struggled with sample attrition and nonresponse bias. The 2018 SIPP is the most recent complete sample. In 2019, the government furlough interfered with the next wave of SIPP, and after that, COVID-19 greatly impacted the collection of data. Out of caution for these issues, we adopt the 2018 SIPP for our analysis as the most recent reliable SIPP sample. SIPP is a longitudinal survey that covers 4 years. For this analysis, we use Wave 1 of the 2018 SIPP sample.

We construct household-level asset and debt measures by combining multiple variables in SIPP. We summarize the dataset to be at the household-level, using the householder as the reference for demographic characteristics and rolling up other variables to the household level when necessary. The householder thus functions as our main unit of prediction. The core workflow involves fitting a predictive model to SIPP and then generalizing the predictions to a poststratification table derived from an ACS-PUMS extract accessed through IPUMS-USA (Ruggles et al. 2023). ACS-PUMS includes a detailed geographic identifier – the Public Use Microdata Area (PUMA) – which supports detailed estimates of the population distribution for a target local area. Generally, PUMAs can be aggregated to represent populous counties and metropolitan areas, and we define our target local area as a grouping of one or more PUMAs.[1] For consistency with 2018 SIPP, we also use 2018 ACS. However, the procedure could be adapted to use multiple ACS years.[2]

Dependent variables

Just as there are numerous ways to conceptualize wealth, there are many possible ways to define and estimate it. Like most survey-based estimates, we follow the convention of defining wealth as household net worth, the difference between the sum of all assets and the sum of all debts for all members of a household. SIPP does not provide household net worth directly, so it must be calculated by carefully conditioning on asset and debt ownership for all household members and rolling individual measures up to the household level (see Appendix A).

Since most modeling techniques assume a normal (Gaussian) distribution, one of the fundamental challenges with fitting a predictive model to assets, debts, or net worth is that their distributions are extremely skewed with heavy tails. We considered various transformations of the dependent variables as well as alternative regression models, like quantile regression. Ultimately, we decided on a transformation that converts direct asset and debt values into percentile rankings in the national distribution of assets and debts. Our rank transformation follows previous small-area estimates of skewed wealth and income outcomes (Chenevert et al. 2017; Chetty et al. 2014, 2018). Intuitively, percentile rankings are arguably more relevant to cross-sectional analysis of wealth inequality than dollar values. Moreover, percentile rankings can be post-processed in a flexible way, including cross-walking predicted ranks back to dollar values.

For greater flexibility and to capture heterogeneity in the composition of wealth, we predict assets and debts separately. This also permits the estimation of extended wealth-related measures, like debt-income leverage ratios for population subgroups. Estimates of net worth can be considered extended measures, since they are derived from the dollar value of the predicted percentile rank in assets minus the dollar value of the predicted percentile rank in debts. In an alternative version, we also estimated net worth directly; and while we focus on the separate asset and debt models, Appendix B includes results from the direct net worth model.

Since a non-trivial number of households have either no debts or no assets, predictions actually occur in two stages. First, we use regression to classify householders by whether or not they have household assets. This is analogous to imputing whether each household has non-zero assets or, equivalently, to predicting the share of each cell with non-zero assets. Second, we predict percentile ranking in the national wealth distribution for householders with non-zero assets. Fitted values from this regression are the cell-level expectation of ranking in the national asset distribution, conditional on having any assets. The same process is repeated for debts.

To illustrate this two-stage concept, take the following simplified example. Imagine a cell with a population of 3 households, where one household has no assets and the other two have percentile rankings $Y = \{.25, .75\}$. The first stage aims to classify one household as having no assets, and the second stage aims to predict the expected ranking of assets for the remaining two households at $\hat{Y} = .5$. Below, we refer to these two stages as the *classification model* and the *rank model*.

Predictors

To implement MRP, we recode all predictors in SIPP as categorical variables. For the post-stratification table, we recode ACS-PUMS variables to match the SIPP recodes and generate a local population estimate for

[1] The examples below use PUMAs to define a New Orleans-Metairie MSA.

[2] Multiple years could help to mitigate issues arising from too few observations in the ACS data to support reliable subgroup-level population estimates.

every unique combination of the predictor variables. The predictors and poststratification cell characteristics must match exactly, requiring care to ensure consistent recoding across both data sets.

Identifying predictors to include in the model involves a balance, as we seek to enable flexibility to disaggregate estimates, to include meaningful predictors of wealth that capture reasonable variation, and to navigate data constraints by including measures with common support in ACS-PUMS and SIPP. We briefly summarize our intuition for including predictors here.

Standard demographic and economic measures. The model includes standard measures of demographic and socioeconomic status, such as householder age, householder race, householder educational attainment, household income, poverty status, family structure, and householder disability and language status.

Homeownership characteristics. Homeownership is strongly related with wealth, and for homeowners, home equity makes up a large portion of assets. Home equity is not directly measured in ACS-PUMS, but we include reported owner-occupied home value and a simple indicator of mortgage status (active mortgage, owned "free and clear," not owned), which may be derived from both surveys (see Appendix A). Where mortgage status cannot be identified directly in SIPP, we use k-nearest neighbor imputation.

Special income variables. ACS-PUMS and SIPP include some variables that are strongly associated with wealth, such as public assistance or social security income. We recode these income sources into dichotomous measures (indicating presence or absence) to make them comparable across the two surveys.

Class of worker and employment status. We include intercepts to indicate wage and salary employment, self-employment, and non-employment. A motivation for including self-employment is to support the exploration of wealth outcomes for business owners, and we differentiate between incorporated and unincorporated businesses since the association is stronger between incorporated businesses and assets, both intuitively and empirically. Most business owners, however, are not incorporated. For householders who are not employed, we differentiate between retirement and working-age non-employment. We note that, although class of worker and employment status use similar definitions in both surveys, the variables are reported in different ways, and Appendix A includes additional detail on our attempt to harmonize recoding across the two data sets.

Race and ethnicity interactions. Because a central objective is to show racial disparities in wealth, we interact race/ethnicity with other major demographic and socio-economic status predictors. This allows for more heterogeneity in the effect of race/ethnicity beyond the group average effect that would be captured by a single intercept term. Without interactions, this effect would measure residual differences unexplained by compositional differences measured in the other variables. Including interactions allows estimates to incorporate more heterogeneity in the relationship between race/ethnicity and other characteristics.

Geography. SIPP identifies state and metropolitan status, but ACS-PUMS has geographic granularity to the PUMA level. We include state intercepts to approximate unmeasured differences at the state level and a metro/non-metro flag to approximate differences between small town/rural and urban/suburban areas. However, the metropolitan status indicator has different forms of missingness in the two data sets. ACS-PUMS does not code the metropolitan status of PUMAs when the PUMA includes both metropolitan and non-metroplitan areas, and SIPP suppresses metropolitan status for some states. We use k-nearest neighbor imputation to remove missingness from both data sets, following how Williams, Zhong, and Braga (2023) addressed the same problem.

In some cases, recodes are constrained to achieve the closest translation between SIPP and ACS-PUMS. For example, ACS-PUMS has less flexibility on the type of public assistance income reported whereas SIPP provides detailed breakdowns. To render the surveys comparable, we condition on several sources of public assistance (TANF, SSI, WIC, etc.) to match the "welfare income" variable in IPUMS. Similar discrepancies exist with respect to the universes available for the variables in each sample. For example, similar variables may be reported in either survey for "households" or "family households," for ages 15 and up or 16 and up, or at the person- or household-level. Appendix A includes more detail about how each variable is recoded to ensure that the levels in the model match the post-stratification table.

Table 1 summarizes the recoded predictors and their corresponding levels. These predictors form the basis of the modeling and poststratification steps described in the next section.

Table 1: Model predictor variables and their respective levels

Variable	Levels
State	State FIPS codes (N = 51)
Metropolitan status	Metro, nonmetro
Household income	No income or loss, Less than \$5,000, \$5,000-15,000, \$15,000-\$20,000, \$20,000-\$25,000, \$25,000-\$30,000, \$30,000-\$35,000, \$35,000-\$45,000, \$45,000-\$55,000, \$55,000-\$65,000, \$65,000-\$75,000, \$75,000-\$90,000, \$95,000-\$105,000, \$105,000-\$125,000, \$125,000-\$150,000, \$150,000-\$200,000, \$200,000-\$500,000, \$500,000 or more
Age	Under 15, 15-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+
Educational attainment	Less than HS, HS equivalent, some college, Bachelor's degree, Advanced degree
Race/ethnicity	White, non-Hispanic; Black, non-Hispanic; Hispanic; Asian, non-Hispanic; Other race, non-Hispanic
Tenure	Owned free and clear, owned with mortgage or loan, not owned
Home value	Not a homeowner, Less than \$50,000, \$50,000 to \$99,999, \$100,000 to \$299,999, \$300,000 to \$499,999, \$500,000 to \$749,999, \$750,000 to \$999,999, \$1,000,000 or more
Class of worker	Self employed, incorporated; self employed, not incorporated; wage and salary; not employed over 64 years; not employed 64 years and under
Household type	Married with children present, married without children, single with children present, single without children
Sex	Female (0), male (1)
Disability status	Does not have a disability (0), has a disability (1)
Citizenship status	Not a US citizen (0), US citizen (1)
Speaks english at home	Speaks a language other than english at home (0), speaks english at home (1)
Public assistance income receipt	Household did not receive public assistance income (0), household received public assistance income (1)
Social security income receipt	Household did not receive social security income (0), household received social security income (1)
Poverty status	Household income was above poverty level (0), household income was below poverty level (1)

Model specification

To describe the SIPP-based models, we refer to assets as the dependent variables in this section, but debts are estimated in the same way. We predict percentile rankings in the national distribution of assets (the *rank model* for short). Because some households are reported in SIPP as having zero-dollar or no assets, we also use a *classification model* to address these households. Both models use varying intercepts.

Every predictor is a categorical variable, so all individual respondent-level predictors are included in the cell-level poststratification table. For the purposes of this explanation, cell- and respondent-level predictors are thus interchangeable and indexed with i.[3] To introduce how predictors are included in the model, we begin with a simplified example. Let the dependent variable y indicate a household's percentile ranking in the asset distribution. For illustration, we specify a simplified rank model with a reduced set of predictors, where the relationship between the predictors and the dependent variable are modeled with varying intercepts for race/ethnicity, state, race/ethnicity $\times$ state interactions, and a fixed intercept for dichotomous gender as defined in the data:

$$y_{[i]} = \alpha^{race}_{j1[i]} + \alpha^{state}_{j2[i]} + \beta^{race,state}_{j1[i],j2[i]} + \gamma^{male} male_i$$

The effect of race and ethnicity α^{race} has levels $j1 = 1...5$, the state effect α^{state} has levels $j2 = 1...50$, and their interaction has $j1 \times j2$ levels. Gender is captured with a binary dummy. The full model takes the same form but includes additional predictors. To accommodate the full set of predictors, the rank model may be written more generally as:

$$y_{[i]} = \sum_{j=1}^{J} \alpha^{j}_{j[i]} + \sum_{k=1}^{K} \beta^{k}_{k[i]} + \sum_{l=1}^{L} \gamma_l l_i$$

The rank model includes α varying intercepts for every level of predictors indexed by j. Referencing table 1, j are ordinal and categorical variables with more than two levels. Interactions between a subset of predictors included in j are also given varying intercepts β and indexed with k. For example, if the model includes race/ethnicity $\times$ state interactions, varying intercepts are included for race/ethnicity and state in α^j and for their interactions in β^k. In the full model, we interact race/ethnicity with other predictors. Where predictors have only two levels (e.g., gender), we include a dummy variable, and these fixed intercepts γ are indexed by l. The use of fixed intercepts for dichotomous variables is motivated by computational efficiency (Lopez-Martin, Phillips, and Gelman 2022).

The model above is fit to a subset of SIPP data with non-zero assets, so $\hat{y}_i$ gives a cell's expected percentile rank in the asset distribution conditional on having any assets. We also fit a classification model to identify households with non-zero assets. We use a logistic regression of the same general form as the rank model, such that $\hat{\theta}_i$ gives the expected share of households with any assets for a cell.

$$\theta_i = logit^{-1}\Big(\sum_{j=1}^{J} \alpha^{j}_{j[i]} + \sum_{k=1}^{K} \beta^{k}_{k[i]} + \sum_{l=1}^{L} \gamma_l l_i \Big)$$

Intuitively, the classification model plays a role in poststratification by ensuring that each cell is given an appropriate weight when "rolling up" cell-level estimates into larger subgroups, i.e., by removing members of each cell who have no assets. It also removes the mass of observations at the lowest rank (i.e., where assets and debts are zero or NA) that would otherwise appear in the data used to fit the rank model. Ninety-five percent of households in the US report having non-zero assets, while 75 percent of households in the US report having non-zero debts in the 2018 SIPP.

[3] The main difference is the weight applied to cell- or respondent-level data. For example, an unweighted regression fit to the SIPP data applies a weight proportional to the number of observed respondents that fall in i; a weighted regression or conventional estimate applies a survey weight to observed respondents that fall in i; and MRP poststratifies fitted predictions by applying weights for every i derived from a separate data set (i.e., ACS-PUMS).

Predictors from table 1 are included in these models as follows. The classification models include slightly fewer predictors due to warnings about singularity. The classification model also excludes the home value indicator since it is conditional on having assets. Predictors in *italics* are included in the rank model but excluded from the classification model.

- Predictors with varying intercepts (α^j): state, household income, age, race/ethnicity, educational attainment, tenure, household type
- Interactions with varying intercepts (β^k): race/ethicity and state, race/ethnicity and educational attainment, race/ethnicity and age, race/ethnicity and household income
- Dummy variable predictors (l): sex, metropolitan status, disability status, class of worker, public assistance income receipt, social security income receipt, poverty status, citizenship, *speaks english at home*, *home value*

The variables included in the above table and final models were selected based on iteratively running classic (non-Bayesian) multi-level models with combinations of predictor variables and interactions and predicting the results on the SIPP data. We considered the mean absolute error (MAE), correlation coefficient, and computational intensity when selecting the final model specifications. The final choice of variables included in the model are based on balancing predictive accuracy, computational efficiency, intuition about the relationships between variables, and flexibility to support disaggregation. We also ran models with net worth percent rank as the outcome using the model specification described above. These models performed similarly, but we opted to focus on modeling assets and debts separately, which enabled us to generate more flexible estimates such as the debt leverage ratio (debts divided by assets).

Post-processing to generate local wealth estimates

In this section, we describe two workflows to post-process the model results into local subgroup estimates. The more straightforward way to estimate subgroup breakdowns follows a slightly modified version of a conventional MRP workflow. Our goal is to generate an estimate for target subgroups that may be defined at different levels of detail, such as "Louisiana households," "Black households in Louisiana," or "Black homeowners aged 55-64 in Louisiana." For any target subgroup g that can be defined as a combination of cells i, we first assign the population of every cell as having non-zero assets (or debts) using the cell portion with any assets $\hat{\theta}_i$ estimated from the classification model, N_i as given by the local poststratification table (in this case, estimated directly from ACS-PUMS sample weights), and $\hat{Y}_i$ estimated from the rank model (using `rstanarm::posterior_epred` or an equivalent). Then, we calculate a weighted average of cell-level predictions that "roll up" into the target subgroup, including only the portion of each cell expected to have non-zero assets.

$$\hat{Y}_g = \sum_{i \in I_g} \hat{\theta}_i N_i \hat{Y}_i \Big/ \sum_{i \in I_g} \hat{\theta}_i N_i$$

$\hat{Y}_g$ is thus interpreted as a subgroup-level estimate of the expected ranking in the national asset distribution, *conditional on having any assets*. Predicted percentile rankings may be crosswalked back to dollar values for reporting. If net worth is the desired measure for subgroup-level reporting, assets and debts are predicted separately, crosswalked to dollars, and reported as the difference in dollars between expected assets and expected debts.

This MRP workflow may be adequate for some applications, but it has drawbacks that can limit its ability to support certain local disaggregations. First, the number of cells is large, and fitted values may not conform to defined percentiles (i.e., $0 < \hat{Y}_g < 1$) that can be crosswalked back to dollar values. Second, the necessity of separating households with and without assets requires awkward qualifications like "conditional on having any assets," limiting the way that estimates can be presented as representative of the whole population. Finally, since wealth tends to be skewed within each cell, cell-level predictions will be artificially pooled toward the center of the distribution. A population distribution cannot be recovered from the conventional workflow, only a distribution or cell or subgroup means. This limits certain ways of summarizing estimates, such as *race-specific distributions* of assets and debts at a state or sub-state level. To generate such estimates, we develop a modified post-processing workflow. The modified workflow has the effect of introducing additional within-cell randomness and constraining the estimates to a state-level percentile distribution of assets and debts derived directly from SIPP. We note that this modified workflow is more demanding on computing resources.

In the modified workflow, we expand the poststratification table to include a number of observations equal to the population of householders $\sum N_i$ for the target state. For this entire simulated state population, we use $\hat{\theta}_i$ to classify individuals as having assets and, conditional on having assets, we predict $\hat{Y}_i$. In effect, for each household in cell i, $\hat{Y}_i$ is randomly imputed using N_i draws from the posterior (using `rstanarm::posterior_predict` or an equivalent). This yields a data set with imputed *national* percentile rankings that correspond exactly with the entire local target population, incorporating sample weights from ACS and prediction error from the SIPP-based regression model. These imputed ranks are then *ranked again* to constrain the predicted distribution (and any extreme values) to a baseline asset distribution derived from the weights in the SIPP state sample. Re-ranking prohibits the resulting simulated population data from having undefined percentiles and forces the simulated population's asset ranks to have the correct uniform distribution. This ensures that the simulated population's distribution of assets converges to the "ground-truth" distribution of assets from the state-level SIPP sample.

The alternative workflow propagates uncertainty from the rank and classification models (Gelman, Hill, and Vehtari 2020) into simulated variation in relative ranks within each cell and between members of different cells. This has the effect of allowing random variation in individual relative ranks while preserving the expected relative ranks of each cell.

To further illustrate, we pose an example. Cell A in the ACS data has a weight of 500. Using the conventional MRP workflow, the expected percentile rank is .40, which is the mean of the predictive distribution from this model. Cell B has a weight of 450 and a predicted percentile rank of .60. There are two problems with using the conventional workflow to describe the true distribution. First, there is no reason to believe the distribution of *expected* percentile rankings of cell A, cell B, and all other cells matches the *true* distribution of percentile rankings, which is uniform by definition. Second, there is uncertainty in the percentile rank of each cell's members relative to the members of other cells. In the case of our example, we should not believe that *all* members of cell A have a lower percentile rank than all members of cell B without accounting for the error around their predictions.

Using the alternative workflow, each cell weight is expanded into N_i household-level predictions that reflect variation in the cell population, since the predictions instead come from a probability distribution. The average percentile rank of cell A's expanded predictions still equals .40, but the household-level predictions for this cell may now range from .20 to .60. Similarly, the average percentile rank of cell B is still .60, but the household-level predictions may range from .40 to .80. When the entire populations of these cells are expanded, we can now account for the probability that members of cell B are ranked lower than members of cell A, since a portion of cell A's household-level predictions are now ranked higher than some of cell B's household-level predictions.

Once we incorporate the uncertainty in the rank order of our predictions, we rerank our expanded predictions to constrain their percentile ranks between 0 and 1, with relative rank orders of the cell means preserved. The re-ranked percentiles now correspond with the state, so the mean for cell A might shift from .40 in the US distribution to .45 in a less wealthy state's distribution, but a few simulated members of cell A will have more assets than simulated members of cell B. The re-ranked distribution is correct by definition, so the only issue at stake involves who sorts into which position in the state distribution. To ensure that the re-ranked simulation does not unduely distort how population subgroups are sorted, we also compare the resulting point estimates with conventional SIPP estimates, checking for plausibility.

Again, the re-ranked percentile rankings crosswalk back to their corresponding dollar values. However, as the percentile rankings now correspond with state-level rankings, the dollar crosswalk uses percentile ranking-dollar estimates derived from the target state's SIPP sample. Because the state sample is relatively small with sparse dollar values, we use smoothed percentile ranking-dollar crosswalks, which we estimate with

LOESS regression. This mostly affects predictions in the tail of the distribution, where small differences in ranking have large differences in dollars in the state-level SIPP sample.

To report sub-state estimates using the modified workflow, the re-ranked state sample is filtered to a target set of PUMAs. Our examples use Louisiana for the reranking step and then filter to the reranked simulations to PUMAs in the New Orleans-Metairie MSA. State percentile rankings or, more usefully, their dollar equivalents then may be summarized flexibly and reported at the local level.

To evaluate the robustness of the results and to illustrate how they may be used to present disaggregated point estimates of wealth, the next section offers graphical examples. In the accompanying brief, the authors explore racial wealth disparities in the New Orleans metro in greater detail, and portions of that analysis are included here only as examples.

Assessment of results

Model Fit

Bayesian hierarchical models are known for their long run times. These models are quite computationally intense even without adding an abundance of random effects with many levels and interaction terms. Each model used for our predictions took days to run. We used the default priors for the models, but ran into some issues with divergent transitions and thus increased our sampler acceptance probability (`adapt delta = .999`) and ran for more iterations. Adding more iterations and increasing the adapt delta can dramatically increase the model run time. Our final asset classification and rank models gave a warning of 2 divergent transitions each. It is sometimes acceptable to assess the model fit and allow a small number of divergent transitions (Team 2024). After examining the model's predictions for various breakdowns of the national SIPP data, we determined that our model was predicting well on our variables of interest and the 2 divergent transitions could be safely ignored.

Graphical model checking (MRP workflow)

For the purpose of illustration, we explore our models' predictive performance using only the asset outcome models. Demonstrations of the debt and net worth outcome models can be found in Appendix B. First, we test the classification model results by predicting the expected percentage (weighted average) of householders who have assets by state with the first method described in the section above (figure 1). We compare the conventional SIPP survey estimate, the MRP estimate based on SIPP, and the MRP estimate based on ACS. For assessment purposes, the ACS state estimates in this section were computed using one percent of the ACS-PUMS sample (.01 percent of the full US population) for demonstration. Otherwise, we use the full state sample from ACS to generate estimates for the local target area.

Figure 1 shows that results from the MRP workflow are comparable to the conventional SIPP survey estimates, but the standard errors on the SIPP and ACS MRP estimates are much smaller. This is especially evident for low-population states with small samples.

Based on the classification model displayed above, we use the asset rank model to predict the expected percent rank value by state, for those who were predicted to have non-zero assets. Figure 2 shows the mean percent rank estimates by state from conventional SIPP survey estimates, MRP estimates based on SIPP, and MRP estimates based on ACS.

Based on the results in figure 2, the average asset percent rank MRP estimates tend to be close and within the confidence interval of conventional state SIPP estimates. Figure 3 presents similar information in a scatterplot. It shows that, while our MRP prediction on ACS tends to produce a slight over-estimate compared to the SIPP survey's average asset percent rank by state, the MRP and conventional estimates are generally correlated at the state level.

Next, we examine sub-state estimates, our driving motivation for doing the MRP workflow. We look within the New Orleans metro, the focus area of our accompanying brief, to further disaggregate our estimates. In

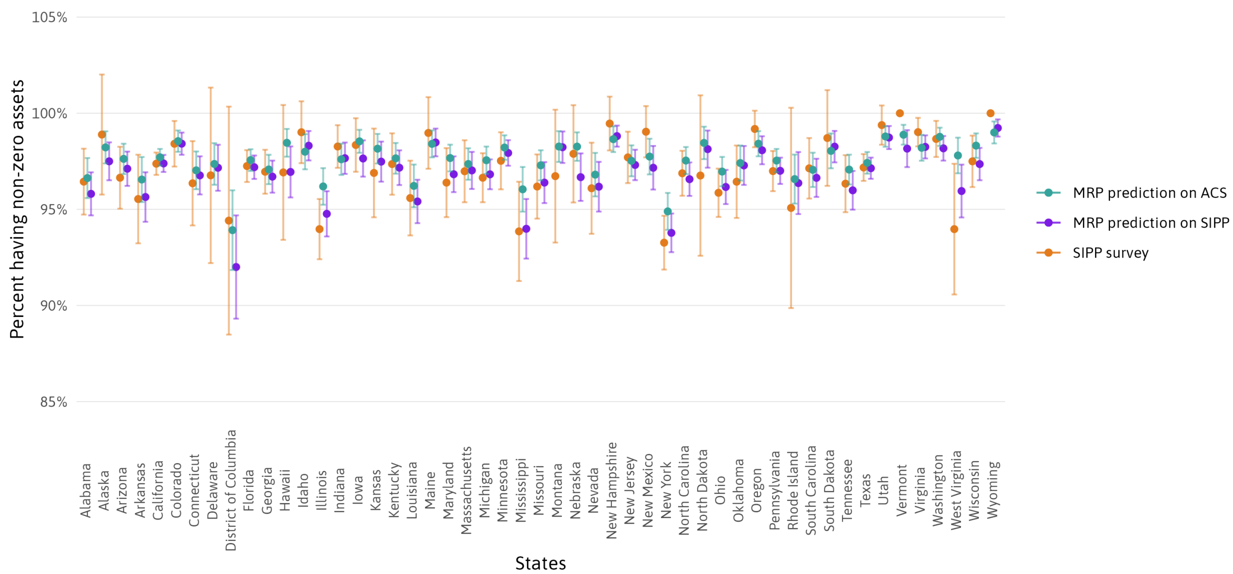

Figure 1: Asset classification model state estimates

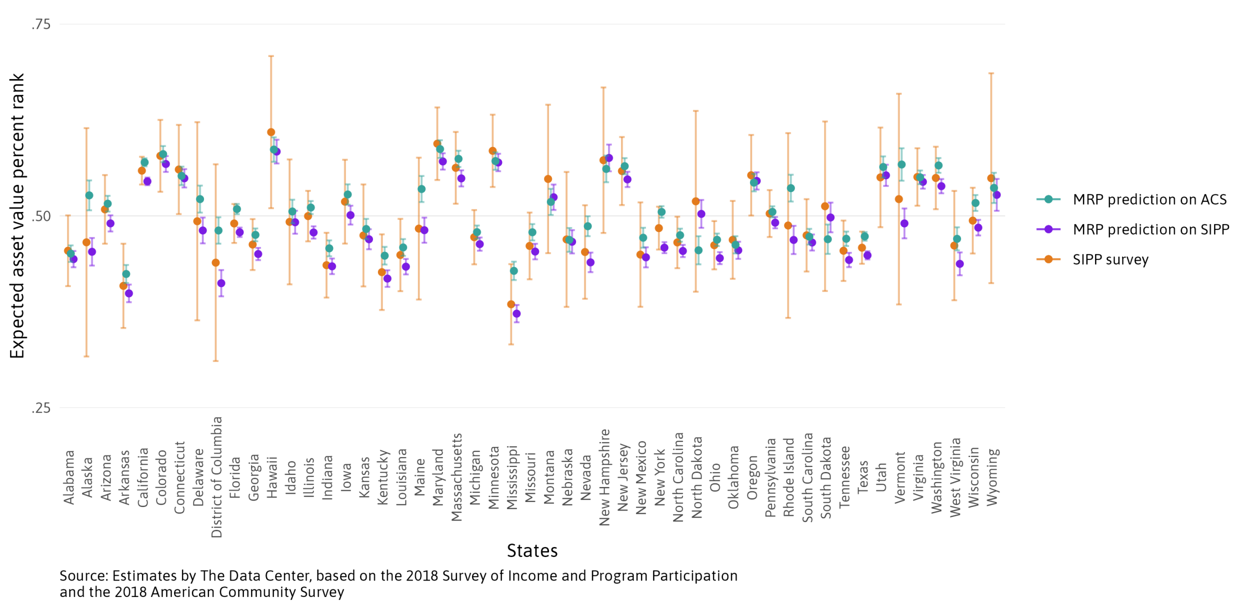

Figure 2: Asset rank model state estimates

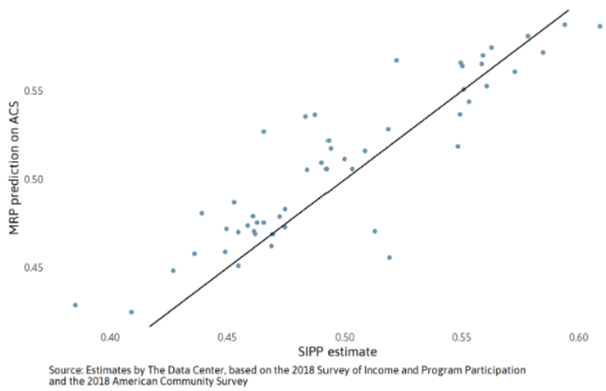

Figure 3: Expected asset percent rank estimates from SIPP compared with ACS predictions by state

the brief, we disaggregate our estimates in at most 3 ways: education by race and age, and tenure by race and age. Figure 4 shows the asset percent rank predictions by race/ethnicity in the New Orleans metro, using the MRP workflow. These are interpreted as expected asset ranks by race/ethnicity, conditional on having non-zero assets, as predicted by the classification model.

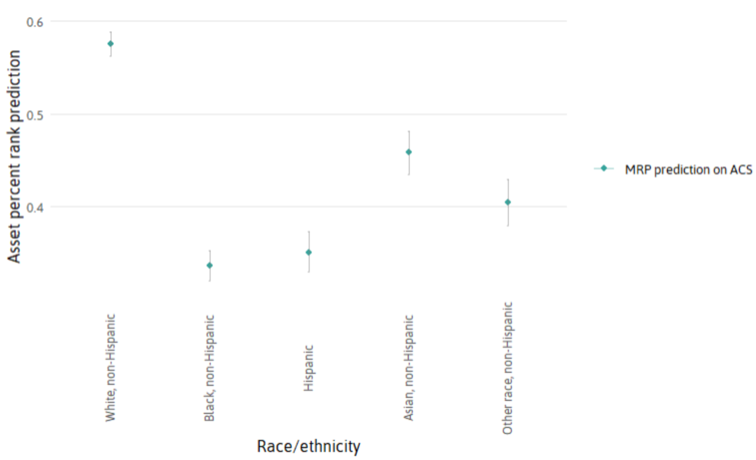

Figure 4: Expected asset rank by race and ethnicity in the New Orleans metro area

As we further disaggregate, our prediction estimates are stable. In the accompanied brief, we break apart our estimates by at most three variables. When testing the variability of our predictions for more detailed breakdowns by tenure or educational attainment along with race and age, we find that the standard errors of our subgroup means do not exceed two percentiles in the New Orleans metro.

While our predictions are stable, what is of more concern is the sample population within the breakdowns of interest. Our decision to estimate wealth using MRP rather than a different method was so we could gain the flexibility of breaking down our estimates by various subgroups. The limitation of our breakdowns comes from our local ACS sample rather than our predictions. We avoid showing estimates for subgroups whose number of observations are fewer than 50 for our sub-group breakdowns. Due to this threshold, we show sub-group estimates only for White and Black householders in the New Orleans metro.

Re-ranking percent rank estimates at the state level (modified workflow)

As explained in the previous section about post-processing our predictions, the standard MRP workflow of generating weighted subgroup means has drawbacks that can limit its ability to support certain local disaggregations. Here, we show results from the modified workflow, which first assigns rank predictions for a state-level simulated population of households and then re-ranks the predictions. The results have the correct population distribution while preserving predicted rank order and incorporating random variation. The re-ranked predictions are reflective of the Louisiana asset distribution rather than the nation's asset distribution. Since the original ranks were assigned based national rankings, they will tend to be lower than our re-ranked estimates that correspond to the state distribution of assets – Louisiana has lower levels of wealth compared to the nation.

Since our Louisiana data was predicted based on the national asset values, our local sample is less likely to assign predictions near the top of the asset distribution. The effect of re-ranking smooths the distribution of ranks so that they are more evenly spaced out. In the expanded data, the spread of predictions for a given cell is varied. Because of the variation in each cell's prediction, the smoothed out distribution has the effect of mixing the population represented in each cell to represent a varied Louisiana distribution of asset ranks.

Using this modified workflow allows us to generate estimates and graphics that are not feasible when estimating group-weighted means. For example, having a household population data set of asset and debt predictions allows us to estimate median (rather than mean) net worth by race, as well as the distribution of net worth values by sub-group. This gives us the ability to visualize the racial wealth gap in more dynamic ways, such as the race/ethnicity-specific distributions shown in figure 5.

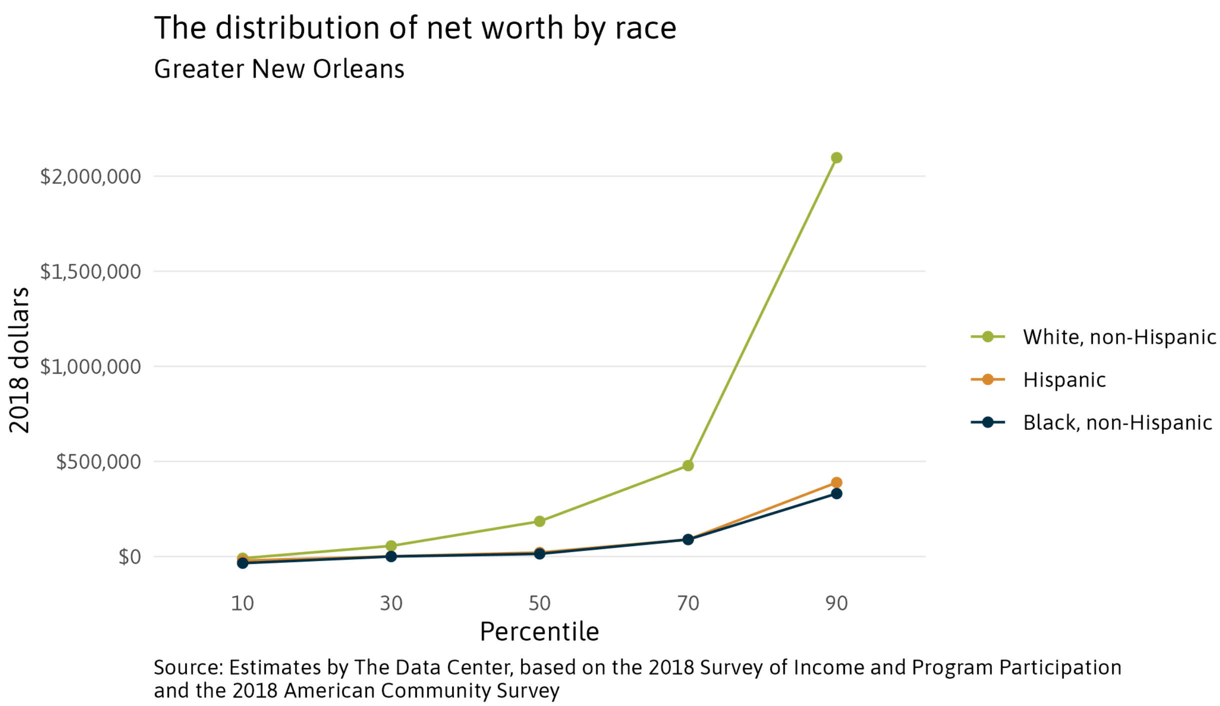

Figure 5: Example of race/ethnicity-specific distributions on net worth

Additional examples of estimates from the modified workflow are available in the companion publication that focuses on wealth in the New Orleans-Metairie MSA.

Discussion

In closing, we discuss potential applications of the local estimates, as well as limitations and avenues for improvement. The procedure described above prioritizes flexibility when generating a range of breakdowns for a local target area. As such, the most immediate use case involves descriptive summaries. Our descriptive summaries were tailored for the New Orleans-Metairie MSA, but the procedure is adaptable to other geographies, with the condition that the level of subgroup disaggregation is sensitive to the number of observations in ACS-PUMS and, by extension, the demographic makeup of the target local area. In our examples, we conservatively opted not to report three-way breakdowns (race/ethnicity-age-homeownership and race/ethnicity-age-education) for Hispanic and Asian households, since many of the subgroup cells have few observations (less than 50) in ACS-PUMS. This limitation could be circumvented with a multi-year ACS-PUMS sample, but we also urge the use of ground-truthing and local stakeholder feedback to guide the use and interpretation of estimates. ACS and SIPP's method of suppressing metropolitan status also may have consequences worth considering on a state-by-state basis.

Ultimately, these kinds of technical decisions are best made in dialog with local context and informed by stakeholders. In line with best practices for equitable data initiatives, making disaggregated estimates more available and more accessible to broader audiences may contribute to shared understanding of the racial wealth gap and motivate local action to address it, especially when grounded in local context. Descriptive summaries also can support a baseline of information available to design and allocate resources to wealth-building interventions. Local breakdowns could also fill gaps as an input in certain kinds of distributional analysis of policies, including policies that are not nominally focused on wealth but that might have implications for racial and wealth equity.

There are several aspects of the model and post-processing workflow that could be improved. The model includes many predictors, and constraints on computing resources and time limited the thoroughness of our ability to test alternative specifications. Both the way that these predictors are coded and the model itself could be optimized for predictive performance, e.g., using cross-validation and regularization techniques from the growing literature on MRP (Ornstein 2023). We have used default priors in our Bayesian model, but more informative priors also might help to improve computational and predictive performance. In addition, fitting a complex Bayesian model and post-processing the results as described here are intensive in their demand on computational resources. In future iterations, we may explore ways to enhance these core technical aspects of the workflow.

More generally, it is important to stress the limitations of local estimates in the context of their use case. While our primary aim is to support more flexible disaggregation of wealth, the estimates are limited in their ability to explain *why* wealth differs from place to place. Geographic differences in wealth levels or wealth inequality can be broken down into two sources of difference. *Composition effects* exist when differences between the outcomes in two places are explained by differences in the characteristics of their populations – age, race, income, homeownership, etc. *Context effects* exist when differences are explained by unique characteristics of the place, many of which may be difficult to measure, such as conditions in the local economy and place-based opportunity structures, institutionally or spatially embedded drivers of racial inequity, local housing market issues, or the state and local policy environment.

Like other local estimates of wealth, the procedure used in this paper is largely determined by composition effects. Indeed, MRP is a tool to adjust for differences in composition, and while the addition of a state intercept captures some residual state-level context effect, this effect is also partially pooled across the effect of other predictors. Other information could be added to the model to better capture context effects, but this may blur the lines between statistical inference and model-based prediction. Nonetheless, there may be an opportunity to supplement the model with state-level predictors, derived either from SIPP or from other data sets, to better reflect context effects.

There already exists a literature on how various factors contribute the racial wealth gap, typically based on decomposition methods and the Survey of Consumer Finances (e.g., see Derenoncourt et al. (2024); Sabelhaus and Thompson (2022); Aliprantis and Carroll (2019); Barsky et al. (2002)). Such studies could be framed generally as efforts to conduct inference about how differences in composition explain differences in key wealth outcomes, e.g., the extent to which income, homeownership, retirement, or inheritance disparities explain the racial wealth gap. In this paper, we essentially leverage information on differences in composition to predict differences in key wealth outcomes at a local level. Both approaches leverage similar information on compositional differences in wealth outcomes to answer different questions. In future work, we may seek to merge these two approaches more directly, drawing from the decomposition literature.

Our aim was to provide local wealth estimates that can be broken down in a flexible way. In concluding, we also stress that contributions to the tiny literature on producing local wealth estimates can have indirect value beyond the direct use of new estimates. They add to the small but growing toolbox for making information for the wealth gap more available, more accessible, and more actionable. In future work, we hope to improve the procedure described above and extend its application to different use cases. Ultimately, local estimates that support a wider range of applications have a role to play in bolstering and animating efforts to address the racial wealth gap, and we hope that the procedure described here contributes to these efforts.

Appendix A: Data preparation

Dependent variables

We predict household assets and debts as reported in SIPP. Wealth is extremely skewed, and addressing this skew presents a fundamental challenge when fitting a regression to assets.

In the SIPP, household asset and debt values are only presented for individuals that have some sort of asset or debt, respectively. Anyone who does not have any assets to their name would have an 'NA' rather than $0 as their reported total assets. In the cases where the asset value is reported to be $0, this means that the individual or household has some sort of asset but it is valued at $0.

We combine those who report $0 in assets and those who do not report any assets together, to create an indicator of whether or not a respondent has "non-zero assets."

Asset and debt questions are only asked once in the SIPP, at the end of the year. This is an individual-level variable. We create an indicator for whether a household has any assets, debts, or net worth by rolling up the individual responses to determine whether anyone in the household has assets, debts or net worth. Note that 97% of households in the US report having non-zero assets, while 75% of households in the US report having non-zero debts in the 2018 SIPP.

Asset and debt dollar values are reported both at the individual and household level. We use the household-level variable conditional on whether the household has any assets or debts.

These variables are heavily right-skewed, which presents a fundamental challenge when fitting a regression. We chose to convert asset and debt values to percent ranks to force the underlying data to be uniform rather than skewed. When we instead model percent rank, we predict the position or order of households in the asset and debt distributions. We can easily crosswalk back to the dollar values from percent rank values later to make the predicted values more comprehensible as dollars.

Predictor Variables

The MRP process requires that all predictor variables be categorical factor variables. Perfectly matching variable definitions between SIPP and ACS required careful coding and researcher decisions to account for missingness or nonresponse in either sample. We explain the nuances of each variable and describe these decisions here.

Recoding was necessary to prepare the SIPP data for a household-level analysis. Some variables in both samples were already presented at the household level, but in other cases, we had to take person-level variables and roll them up to household-level variables. Since the SIPP is a longitudinal survey with detailed information about changes in income and employment, many variables are provided monthly and need to be rolled up to reflect the full year.

Matching predictor variable definitions between SIPP and ACS

Sex, age, educational attainment, race/ethnicity, citizenship, language spoken at home These demographic variables are easily matched based on the SIPP and ACS variable definitions. We report demographic variables for the householder.

Household income In ACS, household income is reported for the past 12 months previous to the month interviewed. In SIPP, the household income variable is reported for each month, so we we sum this over the survey year to match ACS as closely as possible.

Disability Status ACS has more detailed disability status variables, while SIPP has a variable indicating that a respondent has at least 1 of 6 core disability measures: Hearing, Seeing, Cognitive, Ambulatory, Self-care, or Independent Living. ACS provides variables for each of these 6 measures that can be combined to match SIPP's indicator.

Class of Worker This variable is meant to capture the employment status of the householder, as well as the type of work the householder is involved in, if employed. We also want to capture non-employment and self-employment as major indicators of where someone might fall in the wealth distribution compared to the rest of the population.

We create the following levels for Class of Worker:

- Self employed (incorporated)
- Self employed (not incorporated)
- Works for wages
- Not Employed under age 65
- Not Employed over age 65

ACS defines workers based on the workplace in which they spend the most of their time during the reference period, for the population over 16 years old. SIPP, on the other hand, provides comprehensive employment/income data for those older than 15 years old and allows respondents to list up to 7 businesses or jobs per month. To match these two, we select only the first job listed from SIPP and capture this information only for the householder, and condition on the householder being over 16.

SIPP measures employment by week in each month, and Class of Worker is conditional on holding a job during the reference month. ACS is conducted on a rolling schedule and asks about employment "in the last week." To best match this, we randomly select a month in SIPP for each houeshold to calculate the class of worker variable.

The only missingness from this variable creation are for "Unpaid family workers" in the ACS data. We call these people "not employed."

Household Structure/Family Type The household structure and family type variables provided by both surveys do not match well. Instead of trying to match the recoded variables from either survey, we create our own household type variable that is easy to generate from both surveys.

Our variable levels are:

- married householder with children under 18 present in household
- married without children under 18
- single householder with children present
- single householder without children present

There will be both family and nonfamily households present in many of these levels, but it should capture almost everybody.

We create a "child flag" in both surveys which counts the number of people within a household who are under 18, then sum this variable to create a "children present" flag for a household.

Both surveys have marital status conditional on being age 15+. There are very few edge cases where the only person present in the household is under 18, and that is considered "single with children," or two teenagers married would be considered "married with children" under this definition. Because of this, in the cases where a householder is under 18, I check the household size. If the householder is under 18 and is the only person in the household, they are coded as "single householder without children present." If the householder is under 18, married, and there are only 2 people in the household, they are coded as "married householder without children.” Similarly, “married with children” requires over 2 people present in the household, while “single with children present” requires over 1 person present in the household.

The very small edge case that we are unable to fully prevent (if there is any presence in the sample), is if someone under the age of 18 is a single householder and has other roommates under the age of 18. These people will be categorized as “single with children” rather than “single without children” (most other roommates in the sample will be categorized under “single without children”). If someone under the age of 18 lives alone they will be categorized as “single without children.”

Public assistance income, social security income, poverty status In SIPP, all types of supplemental income are reported at the individual level. We roll these up to the household level. ACS reports welfare income for each individual over the past 12 months at the time of the interview. To match this, we sum public assistance over the survey year in SIPP.

The ACS public assistance variable includes Supplemental Security Income (SSI), Aid to Families with Dependent Children (AFDC), and General Assistance (GA).

We translate this to SIPP by adding the income associated with SSI, TANF (which replaced AFDC), and GA. These are reported individually and must be rolled up to the household-level in both SIPP and ACS.

Household poverty ratio is reported similarly in both surveys.

Tenure, homevalue Homevalue is reported the same way in both surveys.

When matching the SIPP tenure definition to ACS, we generate some missing data. ACS provides 3 responses to the tenure variable: Not owned, owned free and clear, or owned with a mortgage or loan. To create this in SIPP, we need to combine an indicator for whether the respondent owns their home with another variable reporting whether there is debt owned against the home. The latter indicator has missingness, even when the home is owned. To account for this missingness, we use k-NN to impute whether or not the homeowner owned their primary residence free and clear or with a mortgage. Our imputation was 98% accurate on our testing dataset.

State and metropolitan status State is reported the same way in both surveys.

Metropolitan status has missingness in both surveys, but for different rea.sons. In SIPP, there is a level for metro status which is “not identified.” This is because SIPP identifies metro/nonmetro status only when either both the metro and non-metro populations both exceed 250,000 or the state is entirely metropolitan. SIPP codes areas as “not identified” when it doesn’t meet one of these criteria or it is uncoded.

In ACS, metropolitan status is sometimes coded as “indeterminable (mixed)” when a county group or PUMA is located only partially within a metropolitan area boundary. Since these are two different reasons for missingness, we use k-NN to impute metropolitan status in both surveys when it is not coded as either metropolitan or non-metropolitan. In SIPP, k-NN predicted metro status with 80% accuracy with $K = 5$.

Appendix B: Supplemental Figures

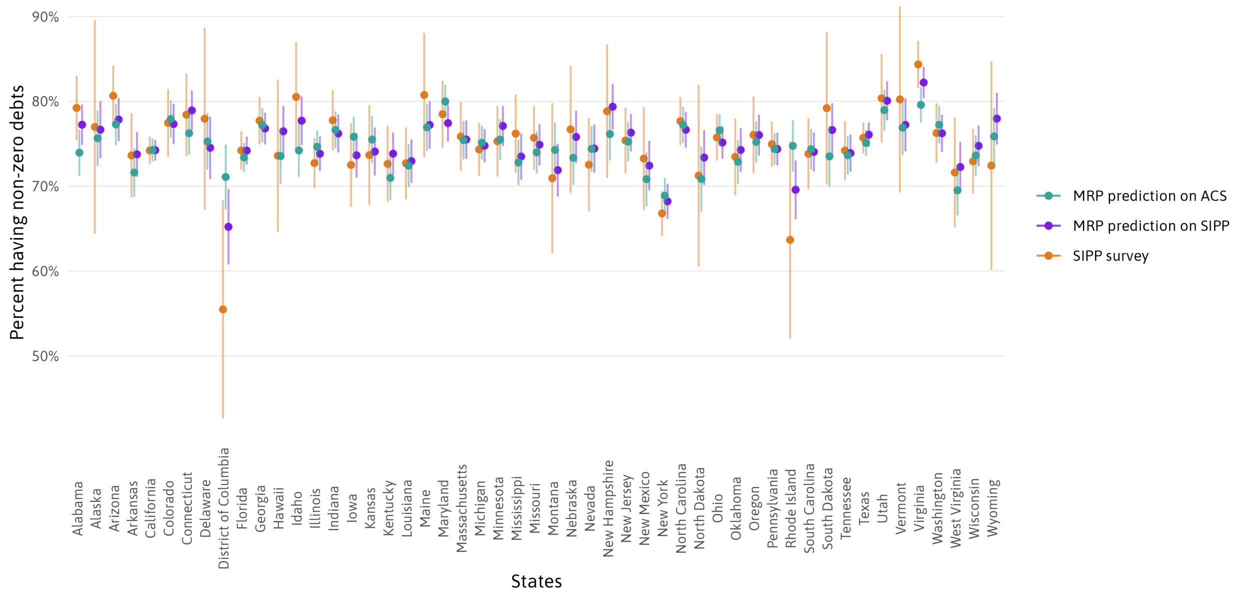

Figure 6: Debt classification model state estimates

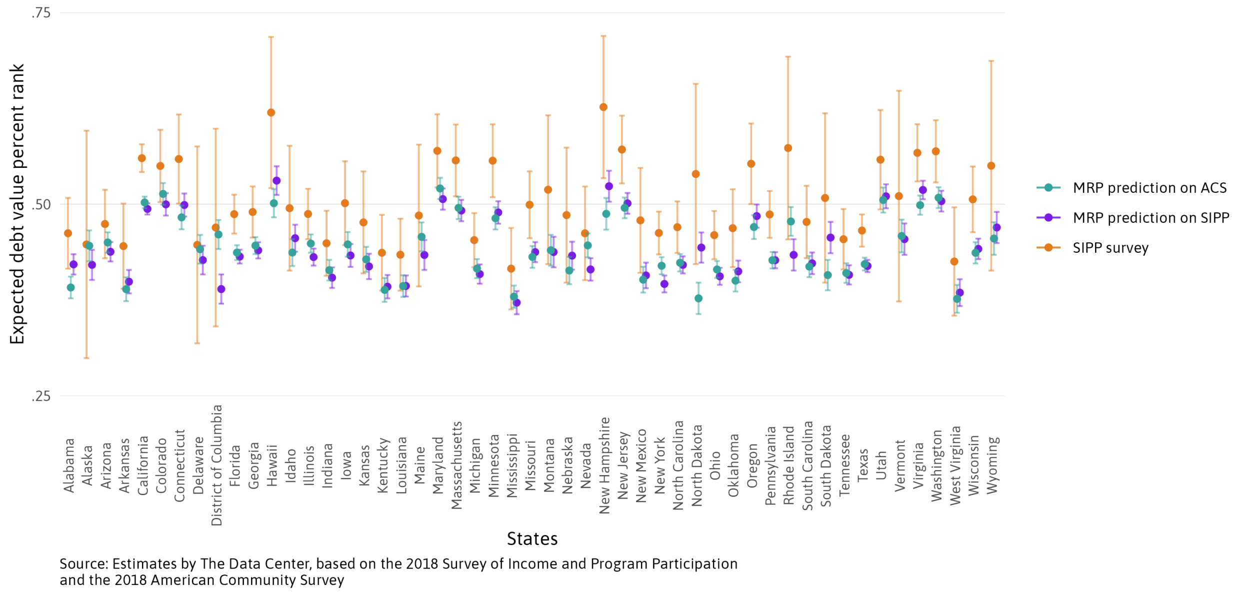

Figure 7: Debt rank model state estimates

References

Aliprantis, Dionissi, and Daniel R Carroll. 2019. "What Is Behind the Persistence of the Racial Wealth Gap?" *Economic Commentary*, no. 2019-03.

Barsky, Robert, John Bound, Kerwin Ko'Charles, and Joseph P Lupton. 2002. "Accounting for the Black–White Wealth Gap: A Nonparametric Approach." *Journal of the American Statistical Association* 97 (459): 663–73.

Chenevert, Rebecca, Alfred Gottschalck, Mark Klee, and Xingyou Zhang. 2017. "Where the Wealth Is: The Geographic Distribution of Wealth in the United States." US Census Bureau.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." National Bureau of Economic Research.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129 (4): 1553–623.

Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. 2024. "Wealth of Two Nations: The US Racial Wealth Gap, 1860–2020." *The Quarterly Journal of Economics* 139 (2): 693–750.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge university press.

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and Other Stories*. Cambridge University Press.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2024. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm.

Lopez-Martin, Juan, Justin H. Phillips, and Andrew Gelman. 2022. "Multilevel Regression and Poststratification Case Studies." https://bookdown.org/jl5522/MRP-case-studies/.

Ornstein, Joseph T. 2023. "Getting the Most Out of Surveys: Multilevel Regression and Poststratification." In *Causality in Policy Studies: A Pluralist Toolbox*, 99–122. Springer International Publishing Cham.

Prosperity Now. 2021. "Technical Documentation: Generating Household Wealth and Financial Access Estimates for Local Geographies." Prosperity Now. https://prosperitynow.org/files/PDFs/local_data_technical_documentation.pdf.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie

Richards, Renae Rogers, and Megan Schouweiler. 2023. "IPUMS-USA: Version 14.0 (Data Set)." Minnesota Population Center, University of Minnesota.

Sabelhaus, John, and Jeffrey P Thompson. 2022. "Racial Wealth Disparities: Reconsidering the Roles of Human Capital and Inheritance." *Federal Reserve Bank of Boston Research Department Working Papers*, no. 22-3.

Team, Stan Development. 2024. "Stan Modeling Language Users Guide and Reference Manual Version 2.34." https://mc-stan.org.

Williams, Aaron R., Mingli Zhong, and Breno Braga. 2023. "Using Machine Learning to Estimate Racially Disaggregated Wealth Data at the Local Level." Urban Institute.