# Financial Health and Wealth Dashboard: A Local Picture of Residents' Financial Well-Being

## Technical Appendix (October 2022)

*Mingli Zhong, Aaron R. Williams, Alexander Carther, Breno Braga, and Signe-Mary McKernan*

This dashboard contains information derived from a nationally representative panel of deidentified, consumer-level records from a major credit bureau. The credit bureau data are from February and August 2021. The February credit records cover 5 million consumers, which is a 2 percent sample of all consumers. In August 2021, we doubled the sample size to 10 million consumers, which is a 4 percent sample of all consumers.

We also incorporated estimates from the Census Bureau's American Community Survey (ACS). Where possible, we used ACS one-year estimates from 2019 (the most recent year of data), but for areas with smaller populations, we used ACS five-year estimates (2015–19).

This dashboard also contains information imputed by machine learning models. We used the 2018 Survey of Income and Program Participation (SIPP) to train the machine learning models. Based on the models, we estimated the net worth and emergency savings for the 2019 one-year ACS.

# Metric Definitions

The following are definitions of all the metrics in the dashboard and their data sources.

- *Residents with delinquent debt*: Share of residents with a credit record who have debt that is at least 60 days past due, including delinquent debt, derogatory debt, and debt in collections. Delinquent debt is debt more than 60 days past due but less than 180 days late. Derogatory debt is debt more than 180 days late. Debt in collections is debt that has been turned over by the lender to a collection agency. (Source: August 2021 credit bureau data)
- *Student loan holders with delinquent student loan debts*: Share of residents with student loans who have delinquent student loan debt. Delinquent student loan debt is debt that is at least 60 days past due. The student loan delinquency rate has been lower since March 2020, which could be related to the pause in federal student loan repayment authorized in the Coronavirus Aid, Relief, and Economic Security Act[1] during the COVID-19 pandemic. (Source: August 2021 credit bureau data)
- *Low-income households with housing-cost burden*: Low-income households are defined as households with annual incomes below $50,000. Renters and homeowners who pay more than 30 percent of their income on housing are considered "cost burdened." For renters, housing costs include rent and utilities. For homeowners, housing costs include mortgage payments or similar property debts, real estate taxes, home insurance costs, and utilities. (Source: 2019 ACS data)

- *Households with at least $2,000 in emergency savings (estimated)*: Share of households with nonretirement savings greater than or equal to $2,000. Savings data at the local level are not available, so we used machine learning to estimate these values. We consider the following nonretirement savings as sources of emergency savings: government securities, checking accounts, savings accounts, money market accounts, certificates of deposit, municipal and corporate bonds, stocks, and mutual funds. We chose $2,000 as a measure of emergency savings because other studies have used the same amount to gauge a family's ability to meet unexpected financial needs (Deevy et al. 2021). (Source: estimated using machine learning based on 2018 SIPP and 2019 ACS data)
- *Median credit score*: Median VantageScore credit score among residents with a credit record. (Source: August 2021 credit bureau data)
- *Mortgage holders who had a foreclosure in the past two years*: Share of residents with a mortgage who have a foreclosure on their record from the past two years. Only residents with a mortgage are considered in this metric. The mortgage delinquency rate has been lower since March 2020, which could be related to mortgage forbearance for homeowners authorized in the Coronavirus Aid, Relief, and Economic Security Act[2] during the COVID-19 pandemic. (Source: February 2021 credit bureau data)
- *Residents with health insurance coverage*: Share of people with health insurance. (Source: 2019 ACS data)
- *Median net worth (estimated)*: 50th (median) percentile amounts of household net worth. Net worth is the sum of asset values minus the sum of liabilities for a household. (Source: estimated using machine learning based on 2018 SIPP and 2019 ACS data)
- *Homeownership rate*: Share of households who are homeowners. (Source: 2019 ACS data)
- *Median home value among homeowners*: 50th percentile values of owner-occupied housing units. Only households with positive home values (not zero or negative values) are considered when calculating the median home value. (Source: 2019 ACS data)

In the dashboard, we present data for all the metrics above at the geographic level of Public Use Microdata Areas (PUMAs) defined by the US Census Bureau.[3] The map breaks are determined using the Jenks Natural Breaks method.[4]

We also show data for five of these metrics at the national and state levels and the city level for select cities, where data are available. Those five metrics are *residents with delinquent debt, households* *with at least $2,000 in emergency savings (estimated), median credit score, mortgage holders who had a foreclosure in the past two years, and median net worth (estimated*). A list of the select cities can be found in the “Select Cities” section.

In addition, we also disaggregate data by racial and ethnic communities for these five metrics for the nation, states, and select cities. Specifically, for *households with at least $2,000 in emergency savings (estimated)* and *median net worth (estimated)*, we report values for Hispanic households, non-Hispanic Asian American and Pacific Islander (AAPI) households, non-Hispanic Black households, non-Hispanic other race(s) households, and non-Hispanic white households.[5] We also report values for households of color by aggregating Hispanic, non-Hispanic AAPI, non-Hispanic Black, and non-Hispanic other races,

For *residents with delinquent debt*, *median credit score*, and *mortgage holders who had a foreclosure in the past two years*, we report values for communities that are majority Hispanic, majority non-Hispanic Asian American Pacific Islander (AAPI), majority non-Hispanic Black, majority non-Hispanic other races, and majority non-Hispanic white, when data are available. We also report values for communities of color when we aggregate Hispanic, non-Hispanic AAPI, non-Hispanic Black, and non-Hispanic other races. We define a community as residents with the same zip code.

Users can search for cities or zip codes to locate a PUMA and see these metrics. Among more than 41,000 zip codes in the US, about 200 zip codes are not linked to PUMAs in the dashboard. These 200 zip codes either cover public places or are areas with no recorded population.

# Data Sources and Methodology

## Credit Bureau Data

We used credit bureau data to generate credit metrics, including *median credit score*. The credit metrics are derived from a random sample of deidentified, consumer-level records from a major credit bureau. The credit bureau data are from February and August 2021. The February credit records cover 5 million consumers, which is a 2 percent sample of all consumers. In August 2021, we doubled the sample size to 10 million consumers, which is a 4 percent sample of all consumers.

The credit bureau data do not include information about race, so metrics for race and ethnicity are based on the racial makeup of zip codes within the geographic area (nation, state, city). Specifically, we determined which communities are majority white based on credit records for people who live in zip codes where most residents are white (more than 50 percent of the population is white). And we determined which communities are majority communities of color based on credit records for people who live in zip codes where most residents are people of color (more than 50 percent of the population is AAPI, American Indian or Alaska Native, Black, Hispanic, another race other than white, or multiracial). Using the same definition, we also created statistics for other races and ethnicities when data were available. Data for specific racial groups or communities of color are not reported because no zip codes are predominantly made up of one racial or ethnic group. The ACS data include information on race, so ACS metrics for white people and people of color are calculated directly for those populations.

## American Community Survey (ACS)

The ACS is a household survey of demographic, social, economic, and housing characteristics administered by the US Census Bureau on a rolling basis. Survey responses are pooled from all twelve months in a year to create an annual dataset. The survey does not include information about liquid assets or net worth, but it does contain many important variables that can predict wealth that overlap with the SIPP including demographics, income, and home value.

We generated four metrics directly from the 2019 ACS: *homeownership rates*, *median home values*, *low-income households that are housing-cost burdened*, and *residents with health insurance*.

For *homeownership rates*, we calculated the share of households who have a home relative to all the households for each PUMA. For the median home value, we limited the sample to homeowners and computed the median home value for households with a home within each PUMA.

For *low-income households that are housing-cost burdened*, we calculated the share of *low-income households with housing-cost burden* relative to all low-income households. We defined households with annual household incomes below $50,000 as low-income households, and we considered renters and homeowners. If any housing-related costs are more than 30 percent of a household's income, that household is considered housing-cost burdened. For renters, housing costs are rent, utilities (water, electricity, gas), and fuels (oil, coal, kerosene, wood, etc.). For homeowners, housing costs are the sum of payments for mortgages, deeds of trust, contracts to purchase, or similar debts on the property (including payments for the first mortgage, second mortgage, home equity loans, and other junior mortgages); real estate taxes; fire, hazard, and flood insurance on the property; utilities (electricity, gas, and water and sewer); and fuels (oil, coal, kerosene, wood, etc.).

For *residents with health insurance*, we calculated the share of *residents with health insurance* relative to all residents. This metric was calculated at the individual level, not at the household level.

The US Census Bureau receives more than 2 million responses to the survey each year, which makes it a useful data source for small area estimation. We used a Public Use Microdata Sample (PUMS) accessed through IPUMS, which contains 1,217,716 households (Ruggles et al. 2021). PUMS data provide Public Use Microdata Areas (PUMAs). PUMAs are Census-generated, nonoverlapping, and state-partitioned geographic areas that contain 100,000 or more people each. PUMAs are drawn based on population, so some are larger than entire counties while others in densely populated areas like New York or Chicago are much smaller. In these cases, PUMAs can be thought of as large neighborhoods.

## Machine Learning Models

We developed a machine learning approach to estimate median net worth and households with emergency savings above $2,000. Although net worth and emergency savings can be measured at the state level, data at the local levels are not widely available. We used machine learning to estimate household assets at the household level and then aggregated results to the PUMA level. We used 2018 SIPP data to train the machine learning models. Based on the models, we estimated the net worth and emergency savings for households in the 2019 one-year ACS. Emergency savings, which is the sum of liquid assets in checking accounts, stocks, bonds, and other liquid savings accounts, can indicate households' resilience and ability to bounce back from financial shocks. Households with liquid assets above $2,000 are considered to have sufficient emergency savings. Net worth, which is calculated as total assets minus total debt, can provide an overview of households' economic well-being and their ability to pursue new opportunities.

We used detailed wealth data and socioeconomic information from SIPP to train machine learning models. The models' predictors include socioeconomic and financial variables from SIPP. The models' outcomes are liquid assets and net worth. Taking the models as given, we then estimated the liquid assets[6] and net worth for all households in the ACS. The ACS has more granular geographic information than SIPP but lacks detailed asset outcomes. The ACS sample covers all PUMAs. Given the estimated wealth measures for the ACS sample, we finally provided summary statistics for net worth and liquid assets for PUMAs and cities.

Our general approach was to leverage the detailed variables in the SIPP to impute values in the ACS, which contains enough households for small area estimation. The approach of using a smaller, more detailed survey to impute variables on a larger, less detailed dataset was also employed by Blumenstock and colleagues (2015) to impute wealth in administrative cell phone metadata in Rwanda. We estimated models on the SIPP, which has detailed information about household demographics, income, and wealth, and then predicted wealth information using the ACS, which has detailed information about demographics and income. After imputing household microdata on the ACS, we summarized liquid assets as the share of households in a PUMA with more than $2,000 in liquid assets, and we summarized net worth as median net worth for households in a PUMA. We also summarized this information at the national, state, and city levels.

Data are reported at the national, state, city, and PUMA levels for all 50 states and Washington, DC. Imputed asset metrics (median net worth and emergency savings) are reported when the sample size is less than 50 or the coefficient of variation is more than 0.4. In addition, for net worth, we did not suppress values when the standard error was less than $5,000 and when the standard error was less than $15,000 and the CV was less than 0.5. Credit metrics are reported when they are based on at least 50 people.

### SURVEY OF INCOME AND PROGRAM PARTICIPATION

SIPP is a household survey of demographic, economic, and government program participation information administered by the US Census Bureau. The survey is nationally representative and is representative of some, but not all, states. The SIPP is an important source of information about assets, and its questionnaire asks detailed questions about retirement accounts, interest-earning assets, other income-generating assets, and other assets.

We constructed liquid assets and net worth from about a dozen individual questions about assets. Liquid assets include transaction accounts and interest-earning accounts such as checking accounts, savings accounts, certificates of deposit, money market accounts, government securities, municipal and corporate bonds, mutual funds, and stocks. Net worth is the sum of asset values minus the sum of liabilities for a household. We included 26,548 households as our core sample from the 2018 SIPP, with one designated "head" person in each household used to determine demographics.

In the dashboard, we reported all metrics for the following 107 US cities (see next page) when data were available. We selected these 107 cities because they contained at least two best-matching PUMAs. See the IPUMS website for more information about cities and their best-matching PUMAs.[7]

Akron, OH
Albuquerque, NM
Amarillo, TX
Anaheim, CA
Anchorage, AK
Arlington, TX
Arlington, VA
Atlanta, GA
Aurora, CO
Austin, TX
Bakersfield, CA
Baltimore, MD
Birmingham, AL
Boise, ID
Boston, MA
Buffalo, NY
Chandler, AZ
Charlotte, NC
Chattanooga, TN
Chesapeake, VA
Chicago, IL
Chula Vista, CA
Cincinnati, OH
Cleveland, OH
Colorado Springs, CO
Columbus, OH
Corpus Christi, TX
Dallas, TX
Denver, CO
Detroit, MI
Durham, NC
El Paso, TX
Fayetteville, NC
Fontana, CA
Fort Wayne, IN
Fort Worth, TX
Fresno, CA
Garland, TX
Gilbert, AZ
Glendale, AZ
Greensboro, NC
Hialeah, FL
Hollywood, FL
Honolulu, HI
Houston, TX
Indianapolis, IN
Irving, TX
Jacksonville, FL
Jersey City, NJ
Kansas City, MO
Laredo, TX
Las Vegas, NV
Lexington-Fayette, KY
Lincoln, NE
Long Beach, CA
Los Angeles, CA
Louisville, KY
Lubbock, TX
Memphis, TN
Mesa, AZ
Miami, FL
Milwaukee, WI
Minneapolis, MN
Mobile, AL
Modesto, CA
Nashville-Davidson, TN
New Orleans, LA
New York, NY
Newark, NJ
Oakland, CA
Oklahoma City, OK
Omaha, NE
Orlando, FL
Overland Park, KS
Paradise, NV
Philadelphia, PA
Phoenix, AZ
Pittsburgh, PA
Plano, TX
Portland, OR
Raleigh, NC
Riverside, CA
Rochester, NY
Sacramento, CA
San Antonio, TX
San Diego, CA
San Francisco, CA
San Jose, CA
Santa Ana, CA
Scottsdale, AZ
Seattle, WA
Shreveport, LA
Spokane, WA
St. Louis, MO
St. Paul, MN
St. Petersburg, FL
Stockton, CA
Sunrise Manor, NV
Tampa, FL
Tempe, AZ
Toledo, OH
Tucson, AZ
Tulsa, OK
Virginia Beach, VA
Washington, DC
Wichita, KS
Winston-Salem, NC

# Notes

1 Coronavirus Aid, Relief, and Economic Security (CARES) Act, S. B. 3548, 116th Cong., 2nd Sess. (March 19, 2020), https://www.congress.gov/bill/116th-congress/senate-bill/3548/text.

2 CARES Act. S. B. 3548, 116th Cong., 2nd Sess. (March 19, 2020).

3 "Public Use Microdata Areas (PUMAS)," US Census Bureau, last updated July 11, 2022, https://www.census.gov/programs-surveys/geography/guidance/geo-areas/pumas.html.

4 "Data Classification Methods," ArcGIS Pro, Esri, accessed July 27, 2022, http://pro.arcgis.com/en/pro-app/help/mapping/layer-properties/data-classification-methods.htm.

5 The race or ethnicity of the household is based on the race or ethnicity of the head of the household.

6 We first calculated the total amount of household liquid assets. Then we identified households with liquid assets above $2,000 as having sufficient emergency savings.

7 Information about cities and their best-matching PUMAs can be found in the spreadsheet called "PUMA Match Summary by Large Place (>75,000 Population)" on the IPUMS website (accessed July 27, 2022): https://usa.ipums.org/usa-action/variables/CITY#comparability_section.

# References

Blumenstock, Joshua, Gabriel Cadamuro, and Robert On. 2015. "Predicting Poverty and Wealth from Mobile Phone Metadata." *Science* 350 (6264): 1073–76.

Deevy, Martha, Jialu Liu Streeter, Andrea Hasler, and Annamaria Lusardi. 2021. *Financial Resilience in America*. Washington, DC: FINRA Investor Education Foundation.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. IPUMS USA: Version 11.0 [dataset]. Minneapolis: IPUMS. https://doi.org/10.18128/D010.V11.0

# Acknowledgments

This feature was funded by a grant from JPMorgan Chase. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

We thank Serena Lei, Christina Baird, and Alice Feng for designing the feature; Oriya Cohen, Courtney Jones, Catherine Harvey, Elizabeth Forney, Alana Kasindorf, and Liza Hagerman for collaboration in developing the feature; and Cary Lou and Vincent Pancini for technical assistance.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

For more information on this project, see our Financial Health and Wealth Dashboard at https://apps.urban.org/features/financial-health-wealth-dashboard.

# Data Citation

Mingli Zhong, Aaron R. Williams, Alexander Carther, Breno Braga, and Signe-Mary McKernan. 2022. "Financial Health and Wealth Dashboard: A Local Picture of Residents' Financial Well-Being." Accessible from https://apps.urban.org/features/financial-health-wealth-dashboard.