

A more equal union? A New Source on Income Distributions for U.S. Communities Defined by Race, Place, and Ethnicity

Illenin Kondo, Kevin Rinz, John Voorheis, Abigail Wozniak

Opportunity and Inclusive Growth Institute, Federal Reserve Bank of Minneapolis
Center for Economic Studies, U.S. Census Bureau

Any opinions and conclusions expressed herein are those of the authors and do not reflect the views of the U.S. Census Bureau, the Federal Reserve Bank of Minneapolis, or the Federal Reserve System or their staffs. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. Disclosure Review Board release authorization numbers CBDRB-FY23-CES014-016 and CBDRB-FY23-CES014-019.

A need for granular income distributions in the US

- Statistically credible measures of income distributions often cannot be tabulated using public microdata sources.
  - for many demographic groups
  - for administrative units below federal
  - for the intersection of these
- Result is that researchers and policymakers cannot answer many important but straightforward questions.
- Ongoing demand from institutional leadership and the public to provide this kind of information.

Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK OF MINNEAPOLIS

A data source for demographically granular income distributions has many potential applications

- Deepening frontier research into income patterns and processes
  - Which detailed demographic groups contribute most to overall patterns of secular and cyclical inequality?
  - How are income inequality and income mobility related locally?
  - How do income dynamics shape the income distribution? Does this differ by group or place?
- Answering basic descriptive questions
  - Have all age, race-ethnicity, and gender groups in America experienced the same changes in income inequality?
  - Do states with higher income inequality have higher or lower income mobility?
  - Do groups that have the most persistent income levels also have the highest earnings and vice versa?
- A lot of overlap in these!

Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK of MINNEAPOLIS


Income
Distributions and
Dynamics in
America

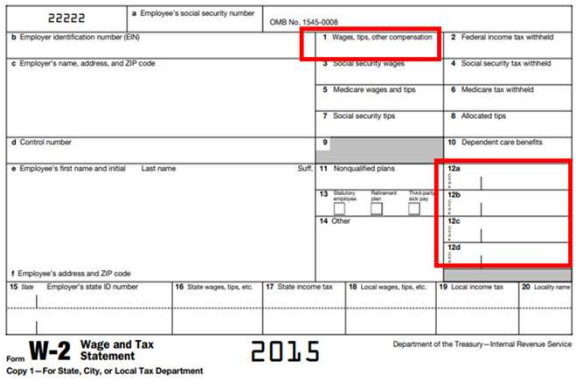
Income Distributions and Dynamics in America

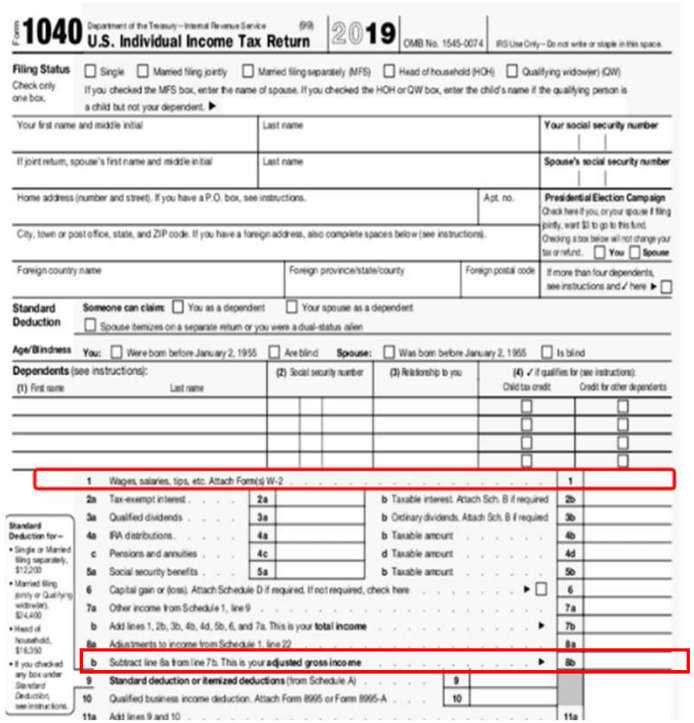
- A gold-standard source for statistics on income distributions and dynamics for US subpopulations and sub-national geographies
- Over 6M statistics built on 20+ years (1998-2019) of IRS administrative earnings data combined with US Census Bureau demographic information
- Census disclosure complete. Accessible this month in soft launch phase.

Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK of MINNEAPOLIS

Building IDDA: Income concepts

- Tax filing unit (tuid) level **[IRS - 1040]**
  - Total wages, salaries and tips
  - Adjusted gross income
  - Aggregated over filers at same mafid
  - 1998-2019
- Individual level **[IRS - W2]**
  - Total wages and salaries
  - Total compensation
  - 2005-2019
- Linkages [Census]
  - Race/ethnicity, gender, foreign born, age group
  - Census **housing unit ID**





Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK OF MINNEAPOLIS

Building IDDA: Tabulation samples

- Universe of tuid-level 1040 information
  - Unit is excluded from years in which there is no filing
  - Aggregate to a "household unit" by totaling income sources of all tuids with same mafid → assign individuals to hh unit
  - Exclude households above 99th percentile of household size
  - Exclude individuals with missing Census information (groups)
- Universe of individual level W-2 information
  - Same exclusions for non-filing or missing linked group information
  - No aggregation
  - Subsample of prime-aged workers 25 to 54
- **Three tabulation samples:** Household level 1040, Individual level W-2, Prime aged individual level W-2

Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK OF MINNEAPOLIS

Building IDDA: Tabulation levels

- Geography
  - US
  - State
  - Native homelands – available as supplement
- Demographic
  - Age (16-24, 25-54, 65+)
  - Sex (M,F)
  - Place of birth (US + US citizens abroad, foreign-born)
  - Race/ethnicity (AIAN*, Asian*, Black*, Hispanic, NHOPI*, White*) where * indicates non-Hispanic
  - From Census numident except for bestrace
- Two-way interactions between demographics and within geography

Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK of MINNEAPOLIS

IDDA statistics tabulated within four-dimensional cells

- **Source sample and income concept *k*** (1040 HH, W-2 individual, PA W-2 individual x AGI, total or wage + salary, total)
- **Year *t*** (1998 to 2019)
- **Geographic level *l*** (US, state)
- **Demographic group *g*** (age bins, sex, race/ethnicity, birth location, or an intersection)

$$\rightarrow S_{ktlg}$$

Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK of MINNEAPOLIS

Statistics modules in IDDA

- **Income distribution - within**
  - Within *ktlg*
  - Income percentiles
    - p10, p25, p50, p75, p90, p95, p98 + deep tails US-wide
  - Top income shares
    - p90, p95, p98 + deep tails US-wide
- **Income distribution - across**
  - Across *g* within *ktl*
  - Top income shares
    - p90, p95, p98 + deep tails US-wide
  - Share of top income earners
    - p90, p95, p98 + deep tails US-wide

- **Income mobility matrix**
  - Across *g* within *ktl*
  - P(final year bin | initial year bin)
    - p0-p25, p25-p50, p50-p75, p75-p90, p90+
  - 1 and 5 year horizons

- **Conditional income growth**
  - Across *g* within *ktl*
  - P(income growth bin | initial year income level bin)
    - See documentation for bins
  - 1 and 5 year horizons

ty &
Growth
INSTITUTE
FEDERAL RESERVE BANK OF MINNEAPOLIS

Quantifying statistical concepts and coverage

- Statistics *S* in IDDA are defined by four dimensions
- Suppression applied to defined statistics to preserve confidentiality per ~standard Census procedures → defined S_{ktlg} may not be available
- Decisions about interactions to disclose made to manage request size → set of available *ktlg* interactions not exhaustive
- One way to explore IDDA is to quantify defined and available S_{ktlg}

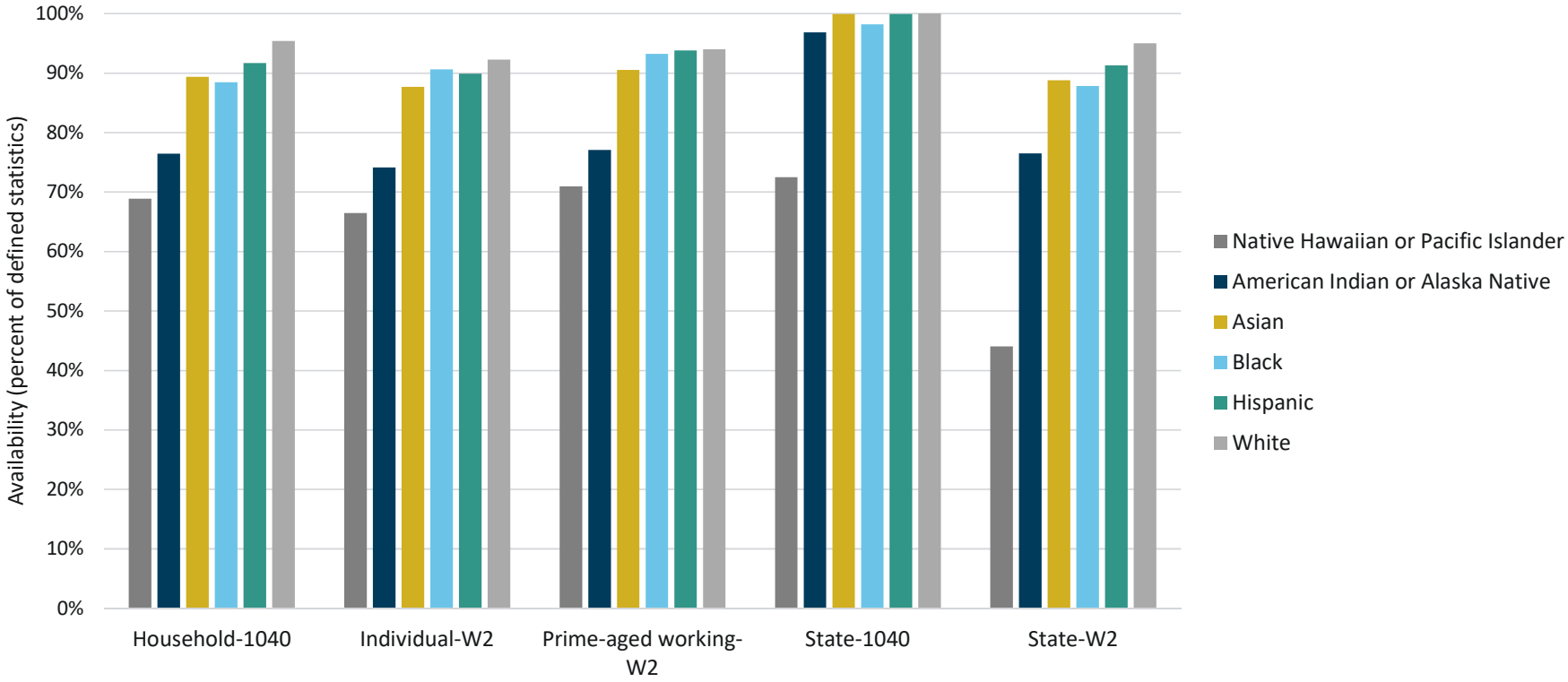
Opportunity & Inclusive Growth INSTITUTE
FEDERAL RESERVE BANK OF MINNEAPOLIS

Defined and available statistics

	U.S.-level				State-level			
	1998-2004		2005-2019		1998-2004		2005-2019	
Household-1040	**Defined**	**Available**	**Defined**	**Available**	**Defined**	**Available**	**Defined**	**Available**
Income Levels	49056	86%	105120	88%	187068	97%	400860	98%
Mobility	22050	100%	37800	100%	514080	97%	881280	97%
Risk	18900	100%	32400	100%	514080	97%	881280	98%
Individual-W2								
Income Levels	0		121680	87%	-		1054170	85%
Mobility	0		129360	97%	-		499392	95%
Risk	0		110880	>99%	-		499392	99%
Prime-aged working-W2								
Income Levels	-		80370	89%	-		-	
Mobility	-		156672	76%	-		-	
Risk	-		72000	100%	-		-	

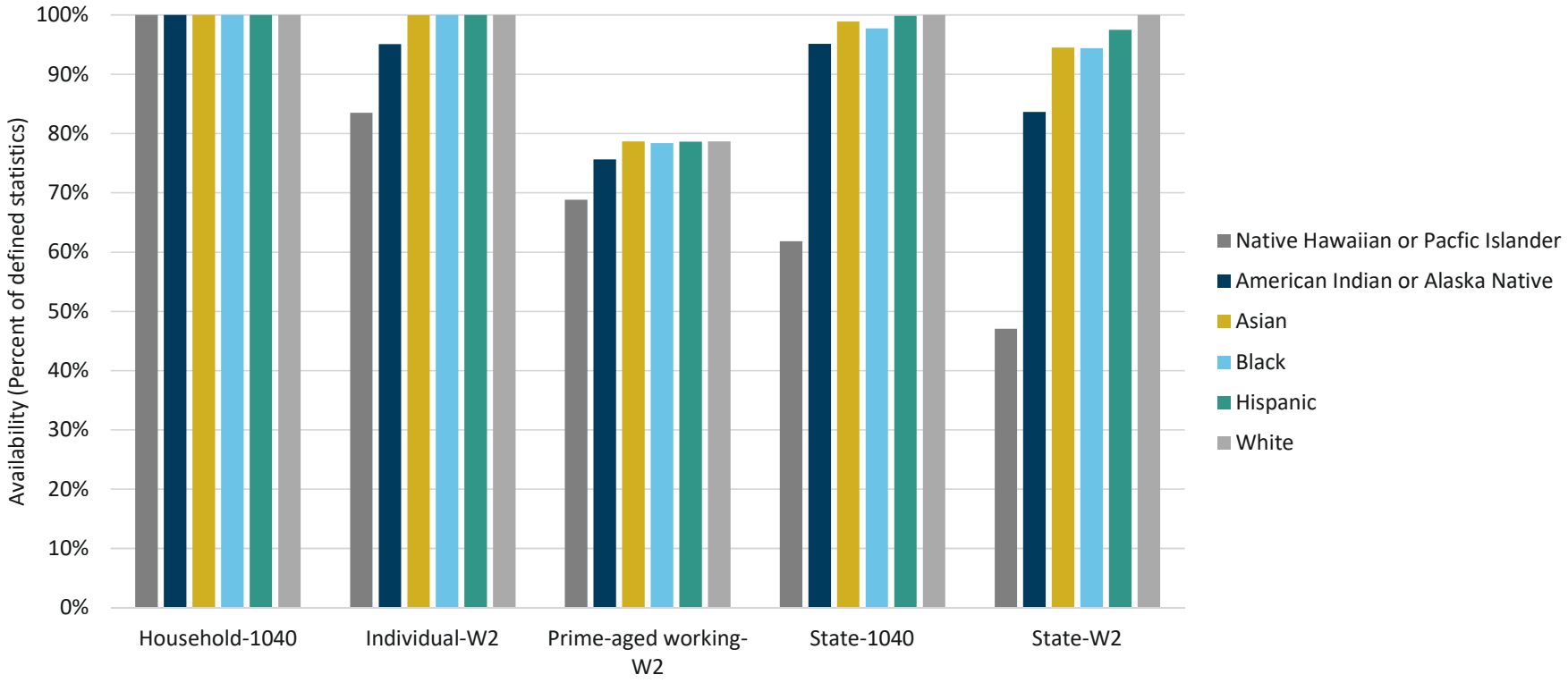
Notes: Any set of statistics where age-by-race breakdowns are available has lower coverage, particularly in the income levels (Age-by-race reported for U.S. and State W-2 Income Levels, U.S. W-2 Mobility and Risk, U.S. 1040 Income Levels). Mobility coverage is low in the prime-aged working sample because individuals can “age in” or “age out” of the sample over time, leading to a number of statistics that are defined for small populations.

Availability by Race/Ethnicity: Income Levels



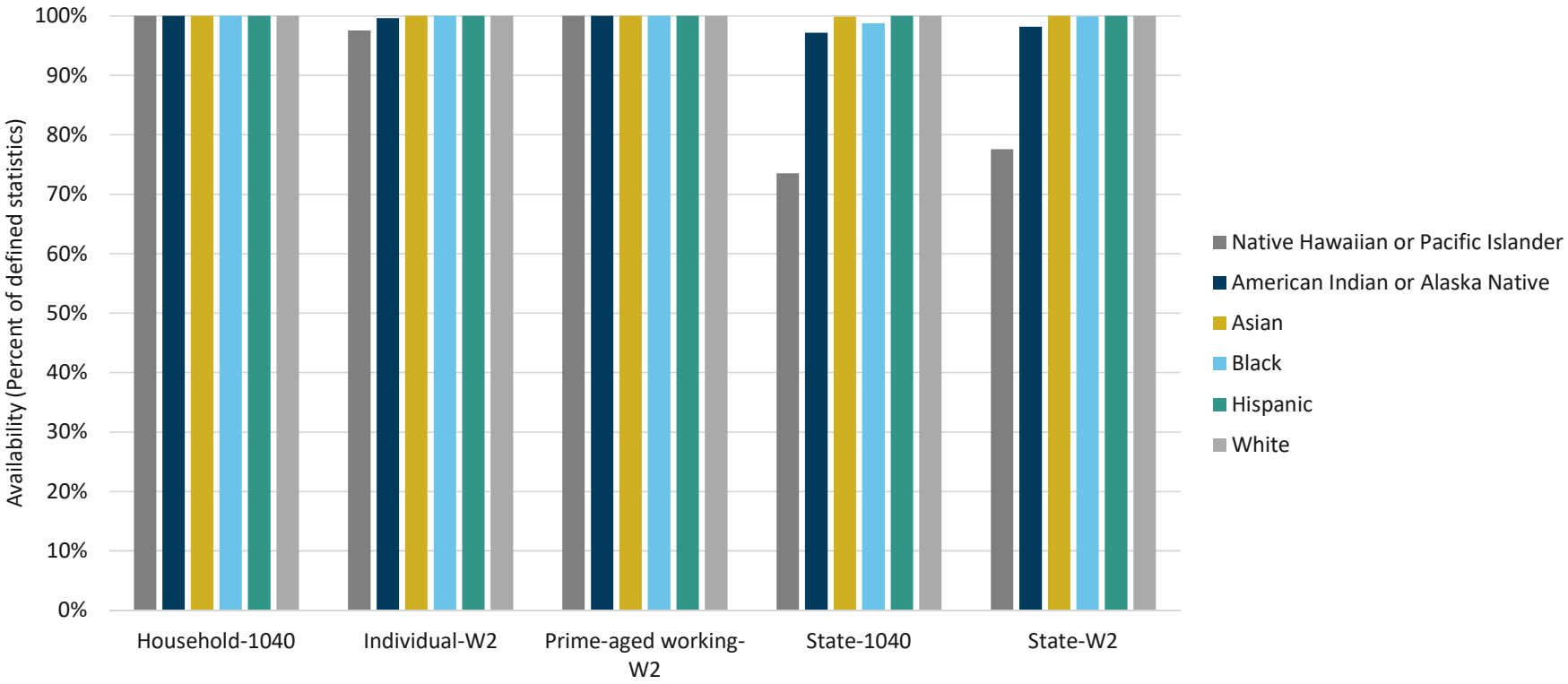
Notes: Availability is calculated as the percent of all defined statistics pertaining to a race or ethnicity group, including race by sex (male and female) and race by age (16-24, 25-34, 35-44, 45-54, 55-64, 65+) where reported.

Availability by Race/Ethnicity: Income Mobility



Notes: Availability is calculated as the percent of all defined statistics pertaining to a race or ethnicity group, including race by sex (male and female) and race by age (16-24, 25-34, 35-44, 45-54, 55-64, 65+) where reported.

Availability by Race/Ethnicity: Income Risk



Notes: Availability is calculated as the percent of all defined statistics pertaining to a race or ethnicity group, including race by sex (male and female) and race by age (16-24, 25-34, 35-44, 45-54, 55-64, 65+) where reported.

Potential structures for CRIW paper

- Data source overview with sample applications on 2-3 topics
- Detailed descriptive paper on one topic
- Something else?

Topic Options

Top incomes by race and ethnicity: levels, inequality, dynamics

Relative inequality at the state level: changes over time and the business cycle

Income distributions, dynamics, and geographic mobility on Native lands*

Comparison of group-level inequality patterns using public CPS data versus IDDA*

Role of non-wage income in inequality by race and ethnicity*

**** Uses data supplement***

Website Preview

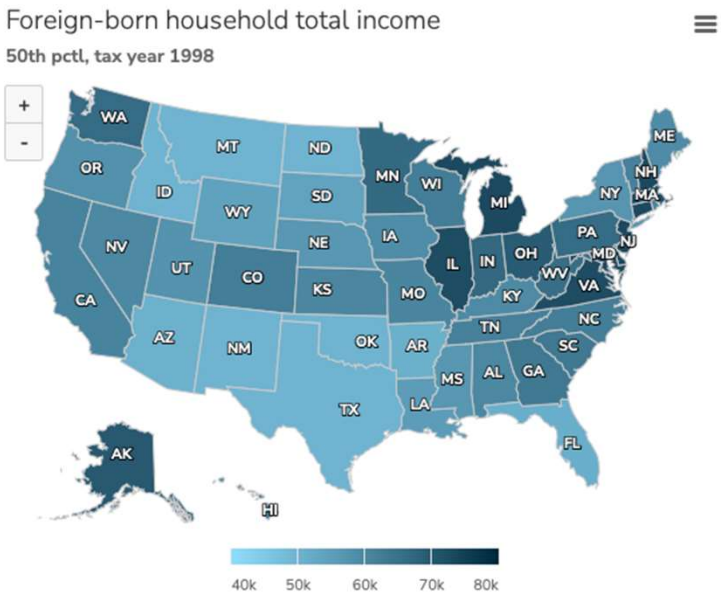


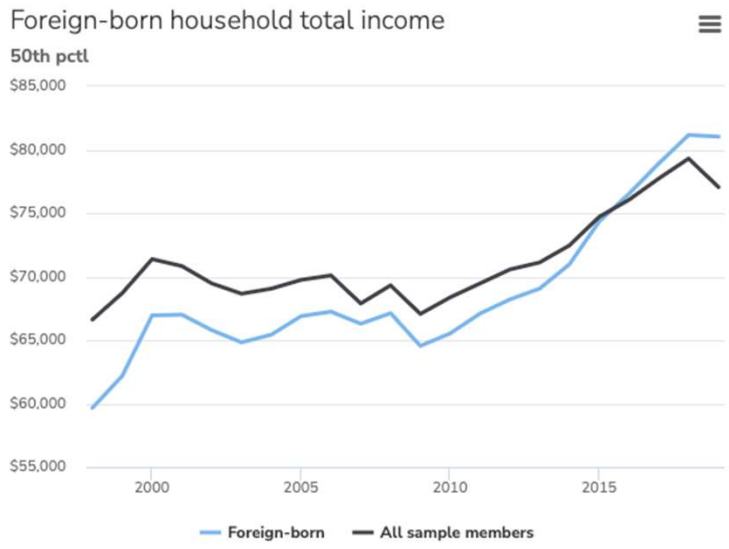
Race and ethnicity | Sex | Country of birth | Age

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

How do household total income in 1998 for people who are Foreign-born compare across states and over time at the 50th percentile?

Do you want to explore how these values compare to **US-born**? ○Yes ◉No





All values are inflation adjusted so they represent an individual's purchasing power in 2012. Values are inflation adjusted using the Personal Consumption Expenditures: Chain-type Price Index deflator. Individual data come from W-2 tax forms and household data come from 1040 tax forms.


Thank you!
Email: abigailwozniak@gmail.com