

Implementing an Adaptive Approach to Collect Acreage on the American Community Survey

Ariel Binder

US Census Bureau

Any opinions and conclusions expressed herein are those of the author(s) and do not reflect the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed, release authorization numbers CBDRB-FY23-CES019-014.

Background

- Incorporating administrative records and third party data into survey production is an important Census Bureau priority
- The American Community Survey's housing questions have been identified as candidate test cases for these efforts
 - First test case: lot size/acreage
- This presentation:
 - Summarize third party data quality
 - Results from simulated partial replacement
 - Plans for adaptive design implementation

The Acreage Item (ACR): Background Information

- Sample universe:
 - Mobile homes (BLD==1)
 - Single-family detached homes (BLD==2)
 - Single-family attached homes (BLD==3)
- Use cases:
 - HUD: excludes 3s in its construction of Fair Market Rents.
 - Pathing for agricultural sales (AGS) question used by BEA: universe excludes 1s.
 - PUMS file for researchers.

➔ Please answer the following questions about the house, apartment, or mobile home at the address on the mailing label.

1 Which best describes this building?
Include all apartments, flats, etc., even if vacant.

A mobile home

A one-family house detached from any other house

A one-family house attached to one or more houses

A building with 2 apartments

A building with 3 or 4 apartments

A building with 5 to 9 apartments

A building with 10 to 19 apartments

A building with 20 to 49 apartments

A building with 50 or more apartments

Boat, RV, van, etc.

2 About when was this building first built?

2020 or later – Specify year

2010 to 2019

2000 to 2009

1990 to 1999

1980 to 1989

1970 to 1979

1960 to 1969

1950 to 1959

1940 to 1949

1939 or earlier

3 When did PERSON 1 (listed on page 2) move into this house, apartment, or mobile home?

Month Year

A Answer questions 4 – 5 if this is a HOUSE OR A MOBILE HOME; otherwise, SKIP to question 6a.

4 How many acres is this house or mobile home on?

Less than 1 acre → SKIP to question 6a

1 to 9.9 acres

10 or more acres

5 IN THE PAST 12 MONTHS, what were the actual sales of all agricultural products from this property?

None

\$1 to \$999

\$1,000 to \$2,499

\$2,500 to \$4,999

\$5,000 to \$9,999

\$10,000 or more

6 a. How many separate rooms are in this house, apartment, or mobile home? Rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling.

- INCLUDE bedrooms, kitchens, etc.
- EXCLUDE bathrooms, porches, balconies, foyers, halls, or unfinished basements.

Number of rooms

b. How many of these rooms are bedrooms?
Count as bedrooms those rooms you would list if this house, apartment, or mobile home were for sale or rent. If this is an efficiency/studio apartment, print "0".

Number of bedrooms

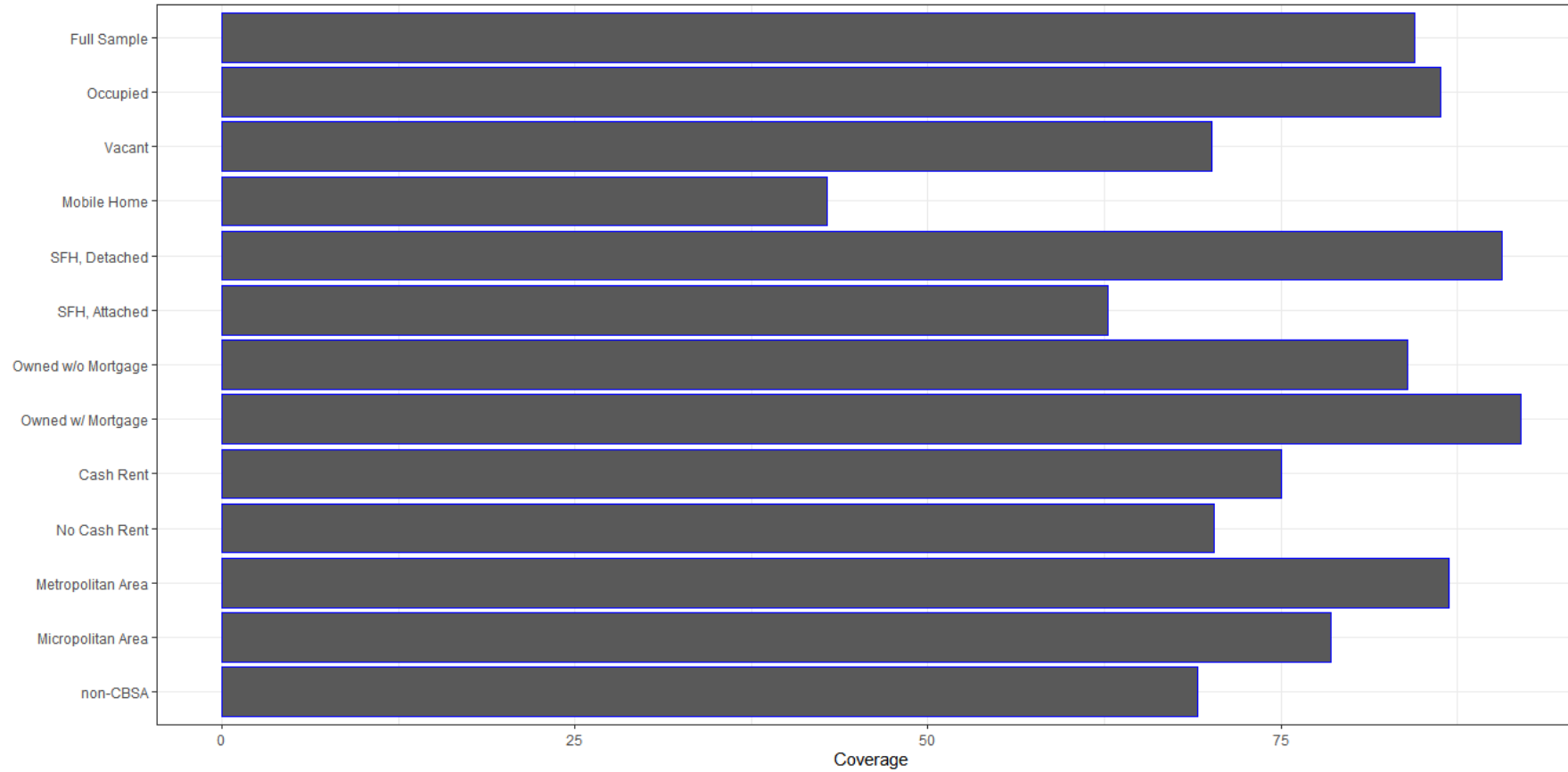
Property Tax Data

- Information on parcels and their characteristics (e.g. lot size) are collected by localities for the purposes of property tax administration
- Several data aggregators collect these county level data and make harmonized national datasets available, Census has a contract with one such aggregator
- Two different ways to link property tax data to the ACS
 - Address-based linkage via MAFIDs
 - Geospatial linkage using parcel boundary shapefiles and GIS methods
 - GS coverage rate > AB coverage rate
 - GS data are **less reliable** than AB data for mobile homes and attached homes, but **as reliable** as AB data for detached homes.
- Assessment data are meant to be a complete panel.
 - However, we have found that a modest share of parcels in one year do not appear in a subsequent or previous year (beyond what we might expect due to new construction / demolition).
- Our approach is to use multiple years of data, and both Geospatial and Address linkages to maximize coverage

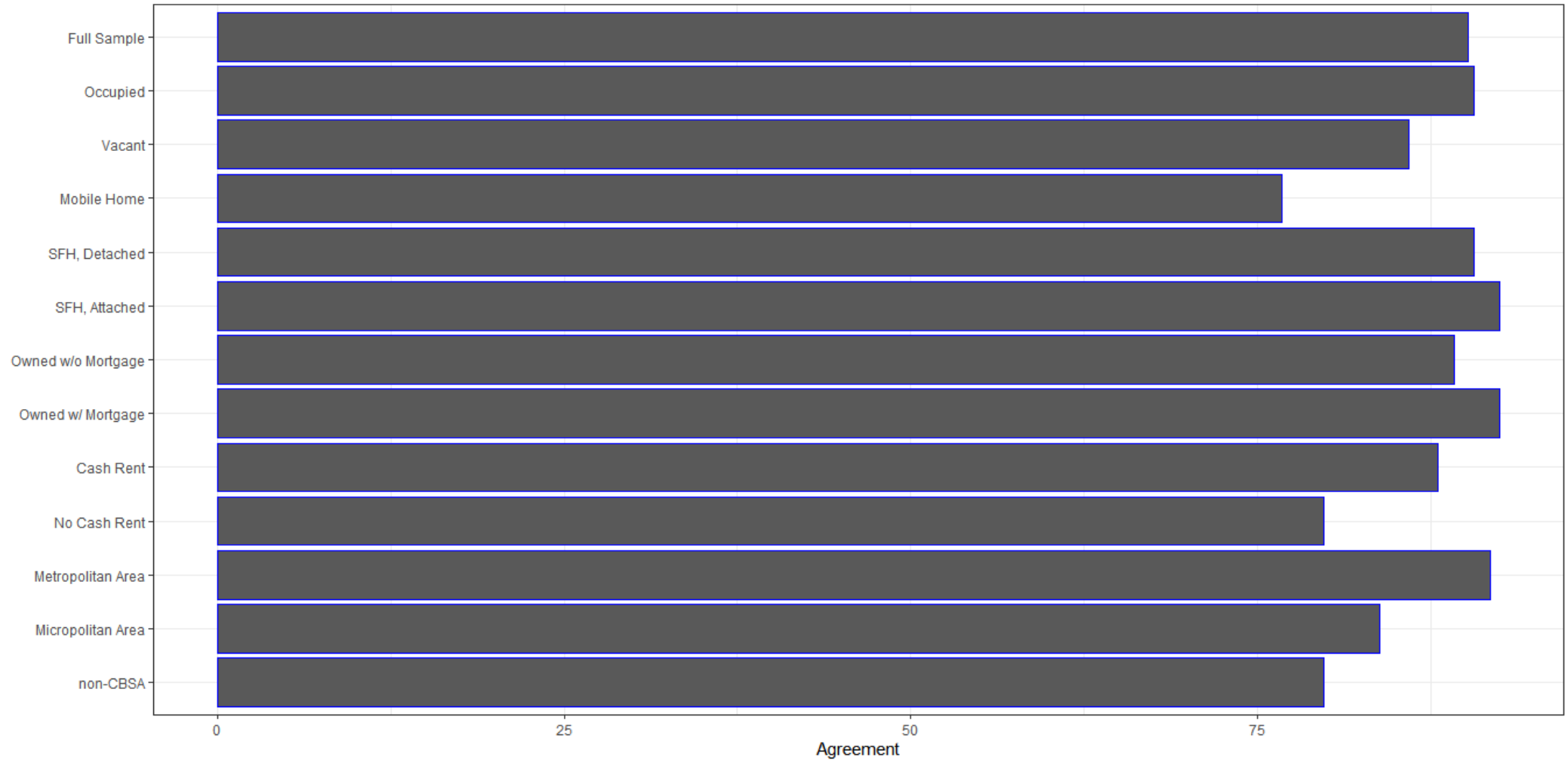
Analyzing Property Tax Data

- How do acreage information in property tax records compare to the ACS?
- Compare 2019 and 2021 property tax data to 2019 ACS
 - Construct composite property tax variable by pooling years and information type (geospatial or address-based) together and merging to ACS. Recode to ACR scale.
- Coverage: for each ACS acreage response, is there a corresponding property tax value?
- Agreement: conditional on property tax coverage, does the property tax value match the ACS value?

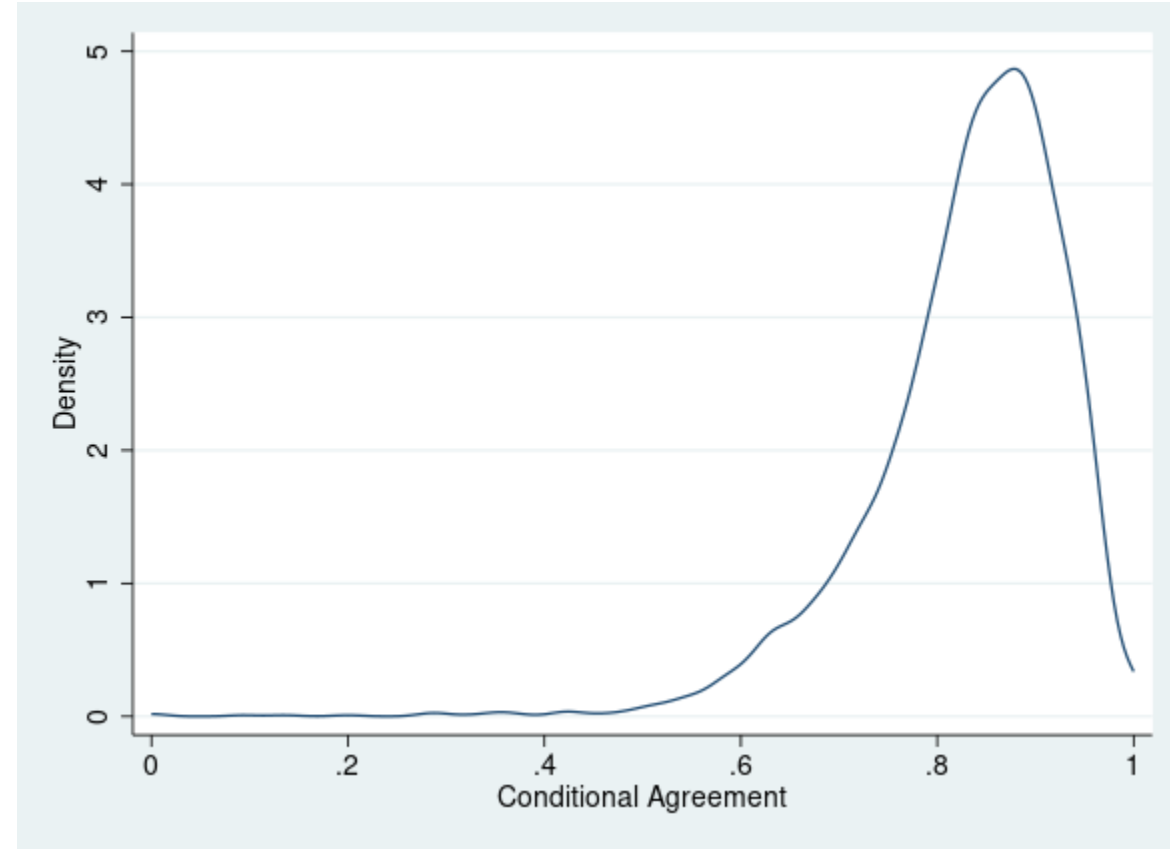
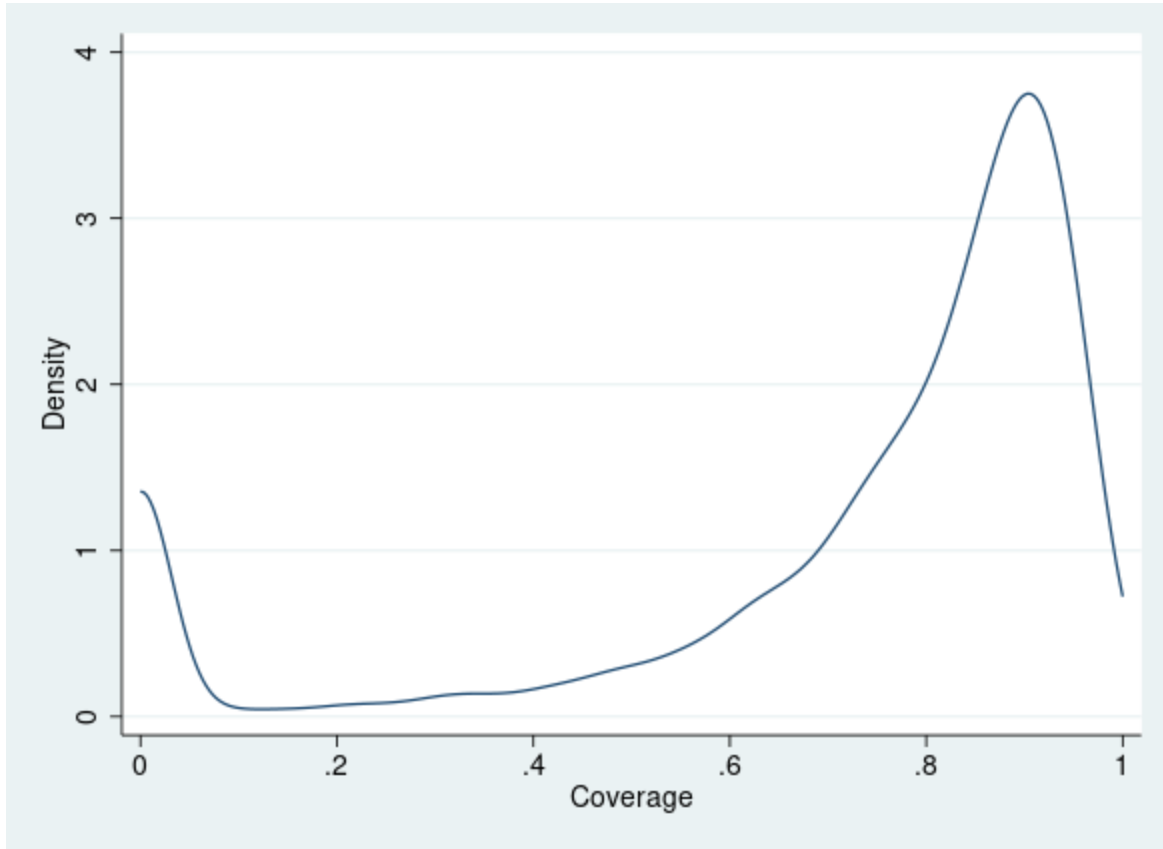
National Coverage



National Agreement



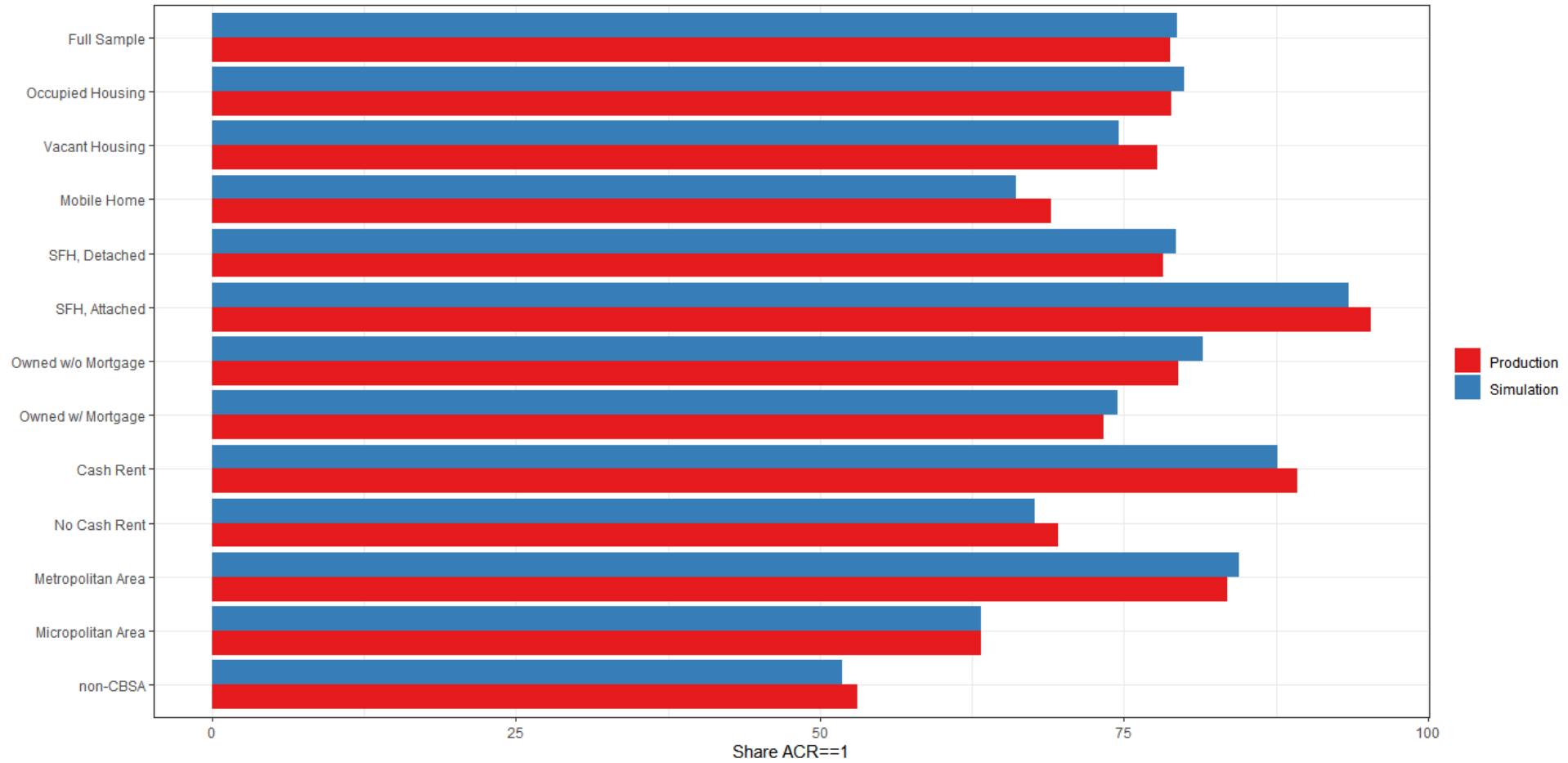
County Coverage/Agreement



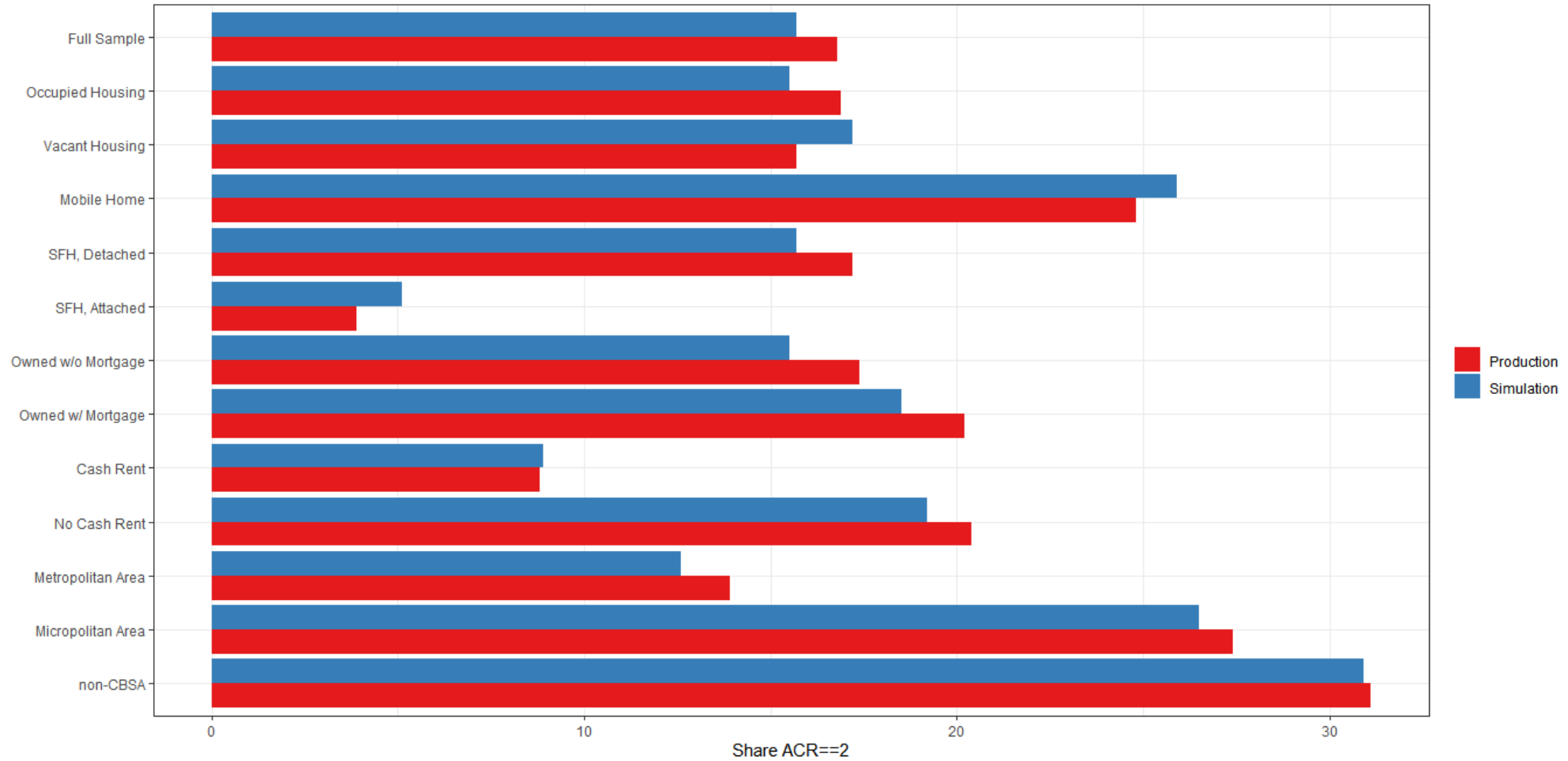
Simulating Partial Replacement

- The property tax data is high quality and fit for use in the survey, although it does not have complete coverage
 - Complete replacement is probably not possible while maintaining quality standards
- We instead simulate a partial replacement of the Acreage question
- Parameters/assumptions:
 - Paper mode is unchanged
 - Web and CAPI do partial replacement: we replace responses with Black Knight data on lot size except for low coverage counties
 - Hot deck imputation (and weighting) uses the Black Knight observations as donors

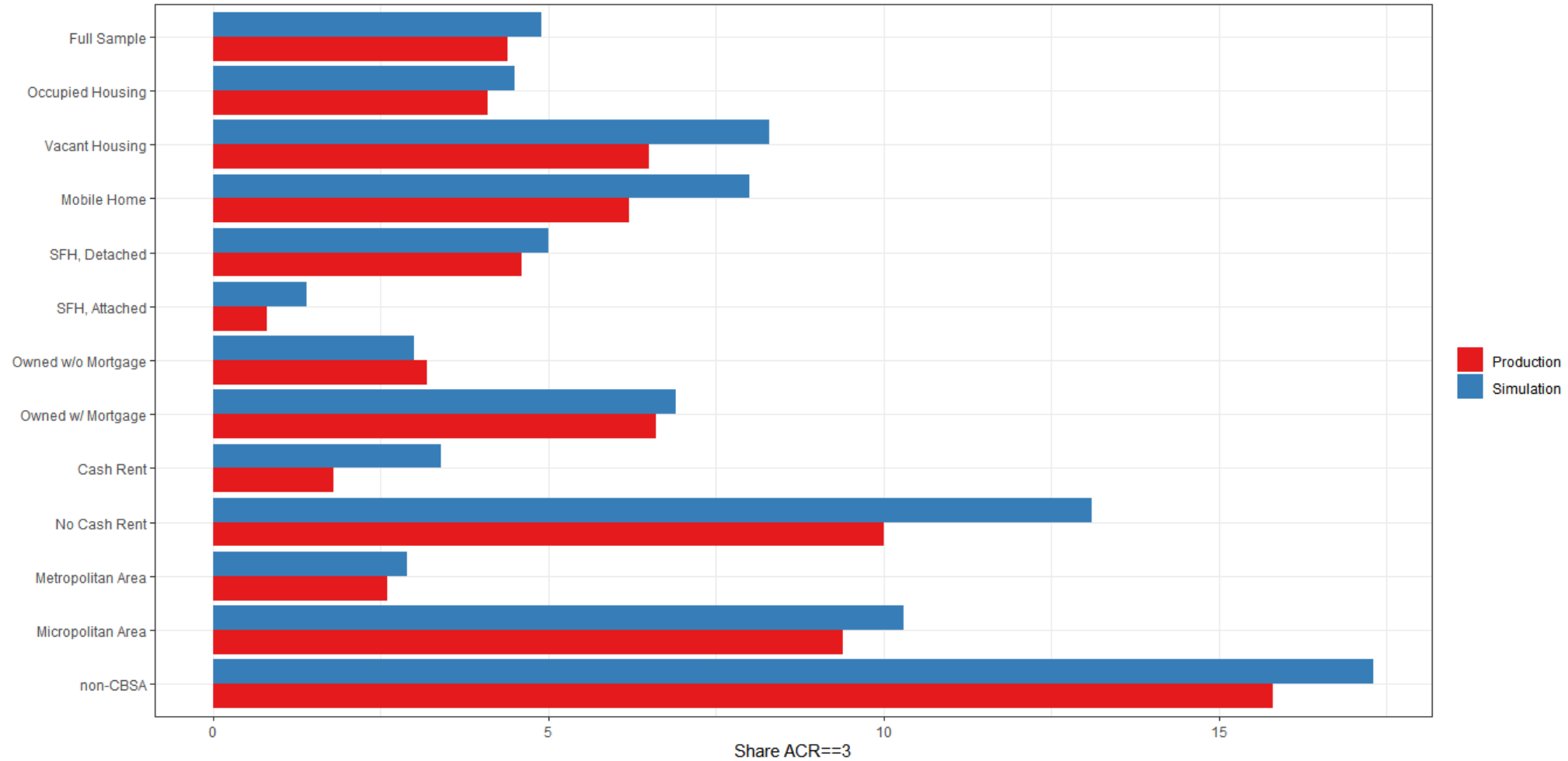
Simulation Results: ACR==1



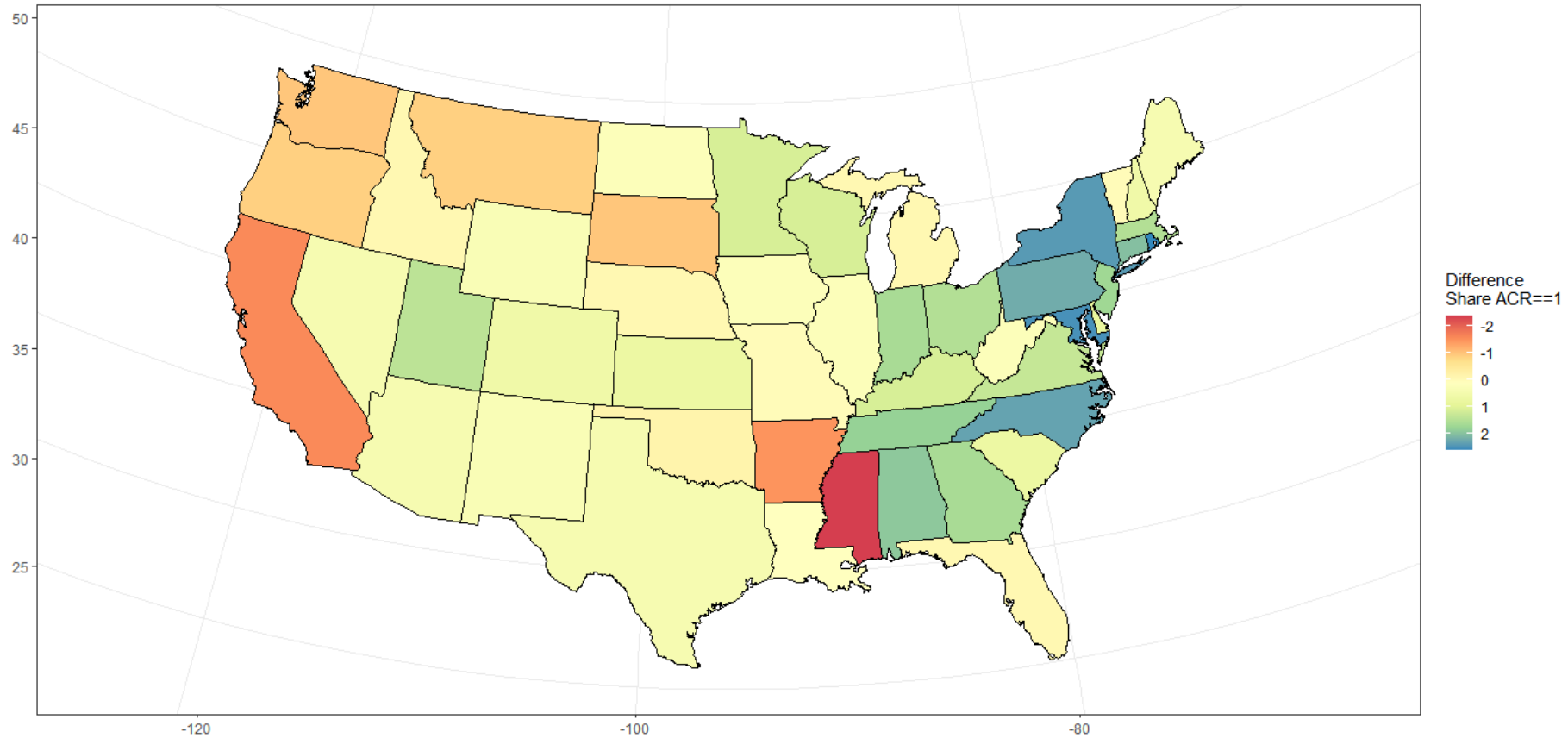
Simulation Results: ACR==2



Simulation Results: ACR==3



Simulation Results



Implementing an Adaptive Design for DY 2024

- ACS sample is drawn twice per survey year (Sept. 2023 and Apr. 2024).
- We will combine property tax deliveries 2023 and 2022 to create **composite adrec variables** that we will pass to the survey instrument.
- Outline:
 1. Construct two composite adrec variables by using delivery 2023 info if it exists and 2022 info if it does not. (Data are missing if a value does not exist for either year.)
 2. Blank out the adrec variables for a list of counties that have unreliable data, pre-determined based on prior research.
 3. Add on a flag variable that will indicate whether the acreage question should be asked of the given CMID, and if not, which adrec value to populate the field.
 4. Pass this composite file to the instrument
 5. For web and CAPI responses, respondents with an adrec value will not be asked the lot size question. Lot size will remain on paper questionnaire.
 6. Post-collection, adrec and respondent provided values will be used as donors in hot deck imputation

Expected Burden Reduction

- We expect this approach to yield substantial burden reduction
- In 2019, nearly 80 percent of responses were from Web or CAPI
- Assuming future coverage and mode share stays constant, this implies the adaptive design approach will yield a **70% burden reduction**

	Less than 1 Acre	1-10 Acres	10+ Acres	Overall
Paper	0.193	0.2429	0.2717	0.2049
CATI/TQA	0.007	0.0083	0.0081	0.0073
CAPI	0.3904	0.3277	0.4015	0.3804
Web	0.4096	0.4211	0.3187	0.4075
Web+CAPI	0.8	0.7488	0.7202	0.7879

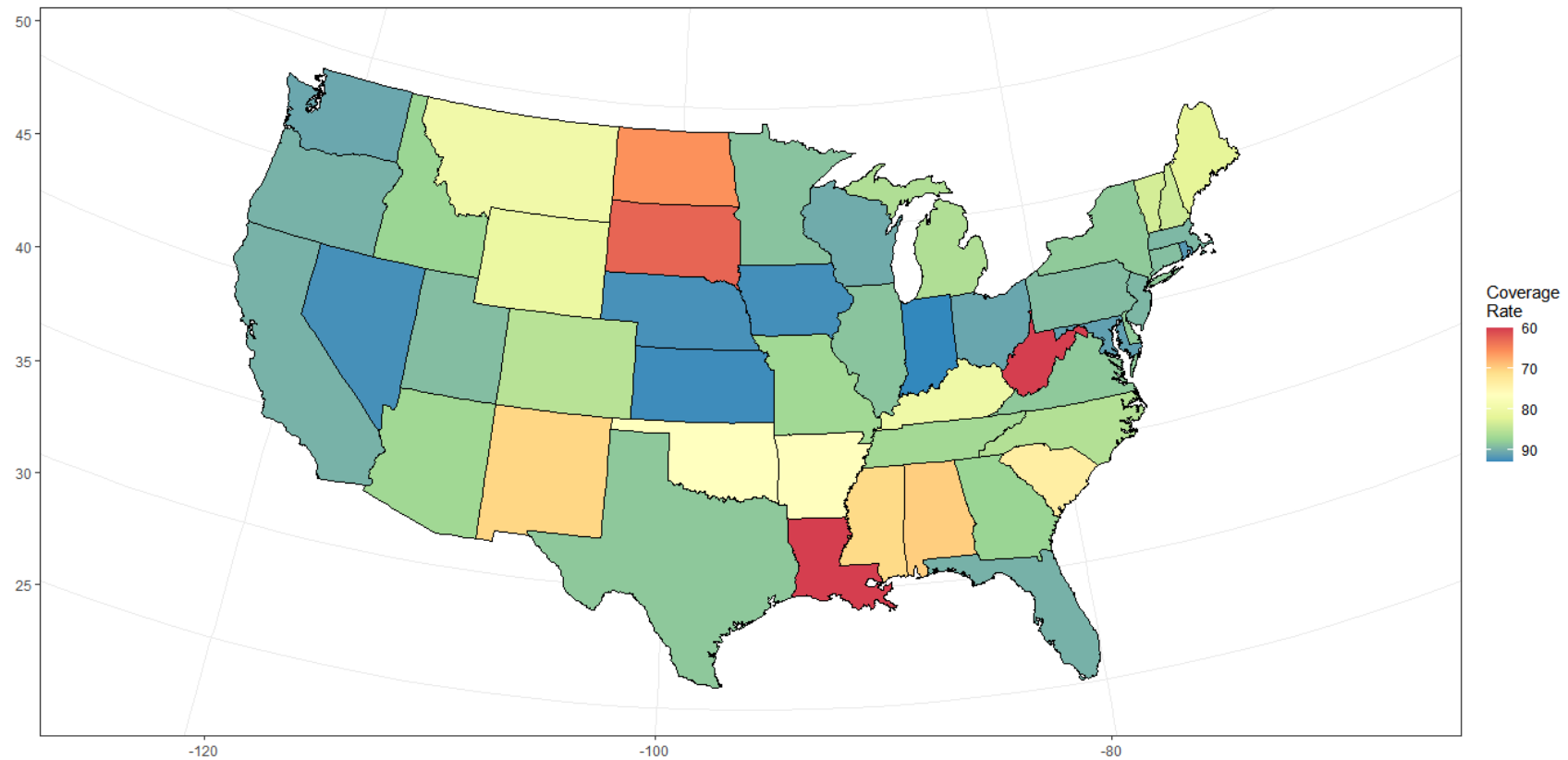
Questions?

Contact information

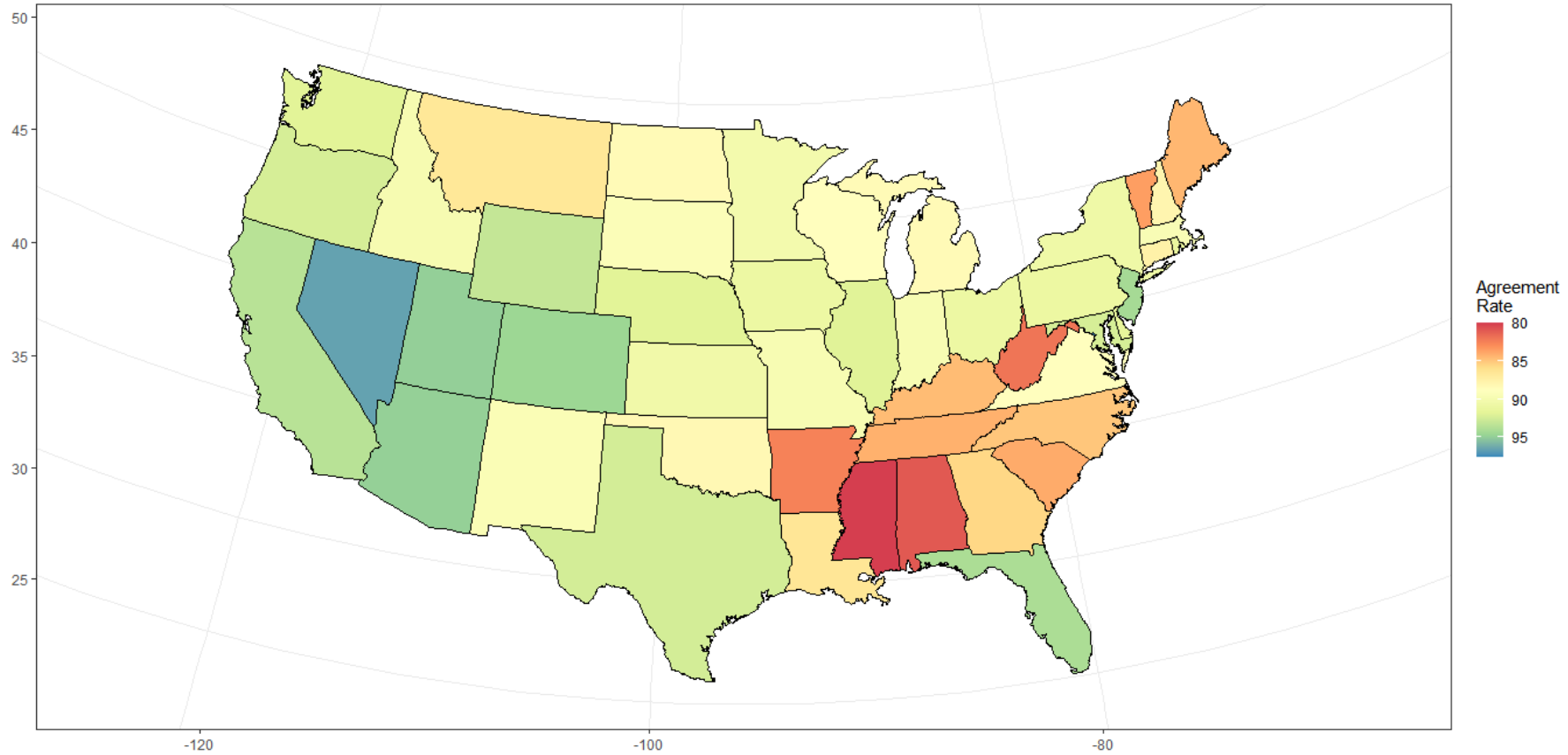
- Email: ariel.j.binder@census.gov

Backup Slides

State Coverage



State Agreement



Simulation Results: Confusion Matrix

		Simulation value			
		missing	<1 Acre	1-10 Acres	10+ Acres
Production value	<1 Acre	0.122	0.633	0.024	0.008
	1-10 Acres	0.024	0.035	0.102	0.006
	10+ Acres	0.009	0.002	0.007	0.027