Challenges of Using Administrative Records to Supplement Response Data in the ACS
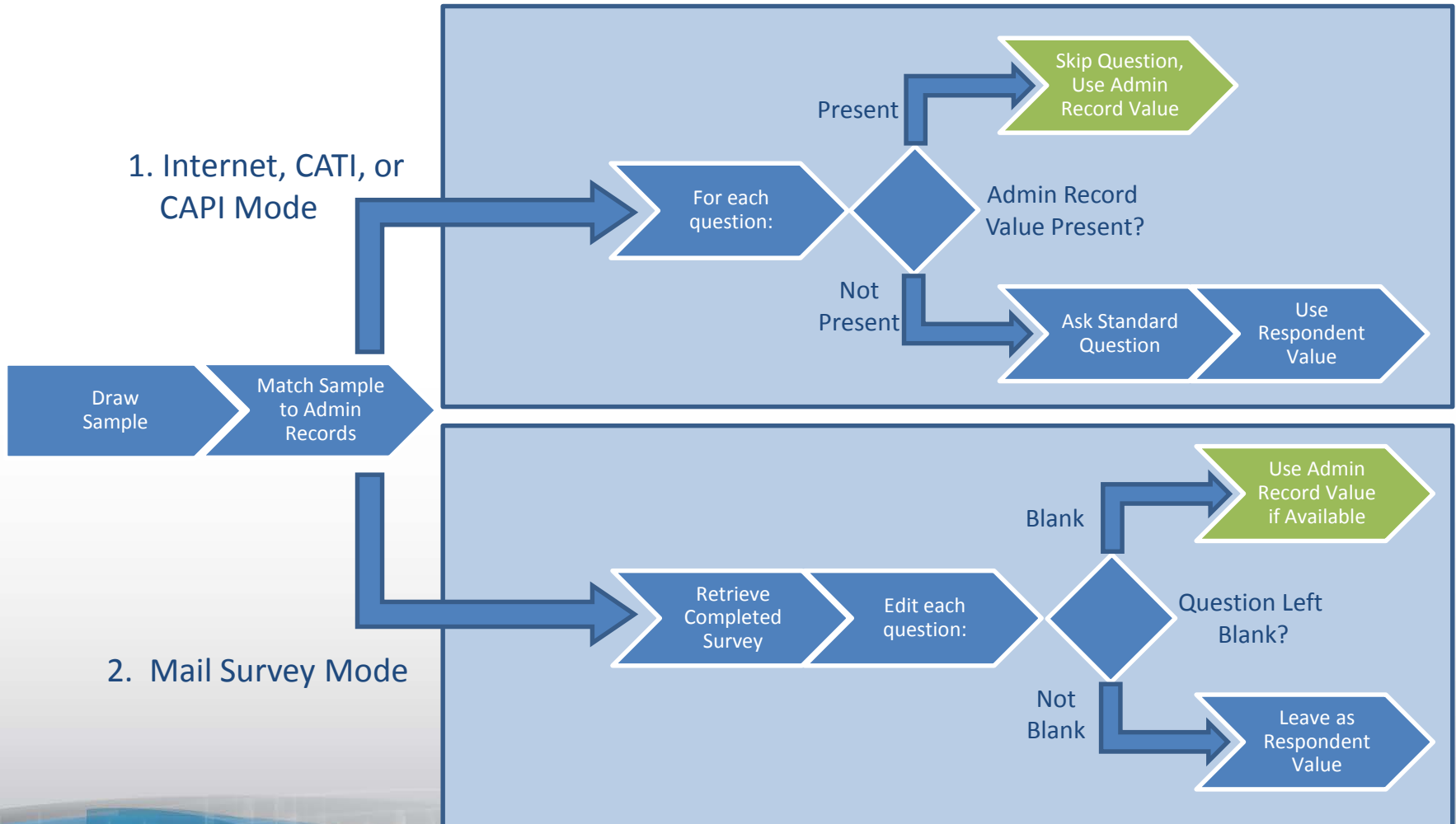
Preliminary Research Findings for the
2018 FedCASIC Workshop
Wednesday, April18, 2018

Sandy Clark and Chase Sawyer

American Community Survey Office

Introduction

- Previous research shows:
  - good match rate
  - promise for some housing items
- ACS perfect platform for simulating the use of Administrative Records (ARs)
  - huge sample
  - spans multiple geographic areas
  - covers many topics
  - high profile
- Satisfy ACS program goals to be reputable, researched, and responsive

Adaptive Design

1. Internet, CATI, or CAPI Mode

Draw Sample → Match Sample to Admin Records

For each question: → Admin Record Value Present?
- Present → Skip Question, Use Admin Record Value
- Not Present → Ask Standard Question → Use Respondent Value

2. Mail Survey Mode

Retrieve Completed Survey → Edit each question: → Question Left Blank?
- Blank → Use Admin Record Value if Available
- Not Blank → Leave as Respondent Value

Simulating this Process

- Used 2015 ACS Responses
- Direct substitution of the four most promising housing items:
    - Year Built
    - Acreage
    - Property Value
    - Real Estate Taxes
- For housing units that have:
    - ACS Internet, CATI, or CAPI return
    - ACS mail return and ACS response is missing

Test vs Control Estimates

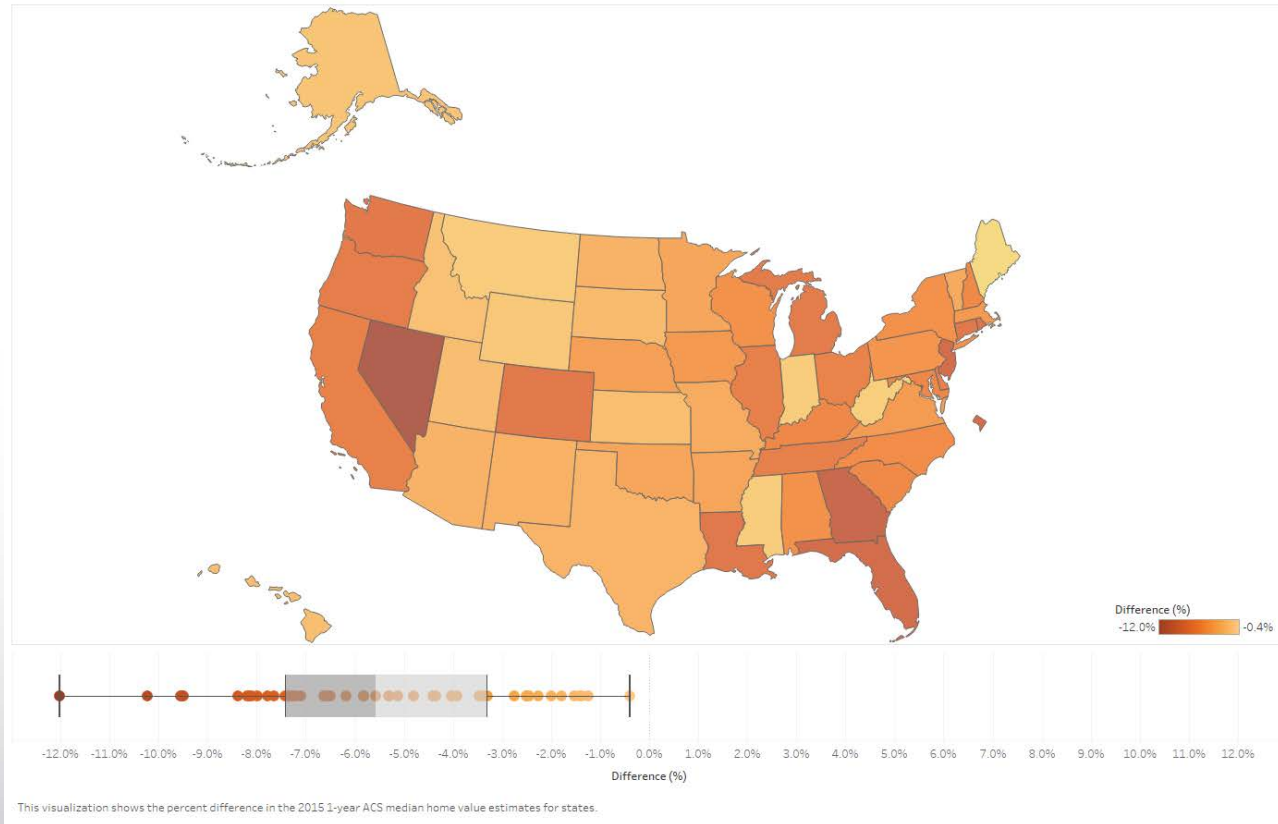## Calculated summary metrics and key measures

- Approx. **79%** of 575 US level estimates statistically different

- Direction of differences varied, but for many items Test (w/AR) was lower than Control

- Test item allocation rates significantly lower than Control by:

    Acreage - 2.4% pts.   **Property Value - 6.8% pts.**
    Year Built - 5.2% pts.   **Real Estate Tax - 12.5% pts.**

- Impacts other survey items besides 4 test items
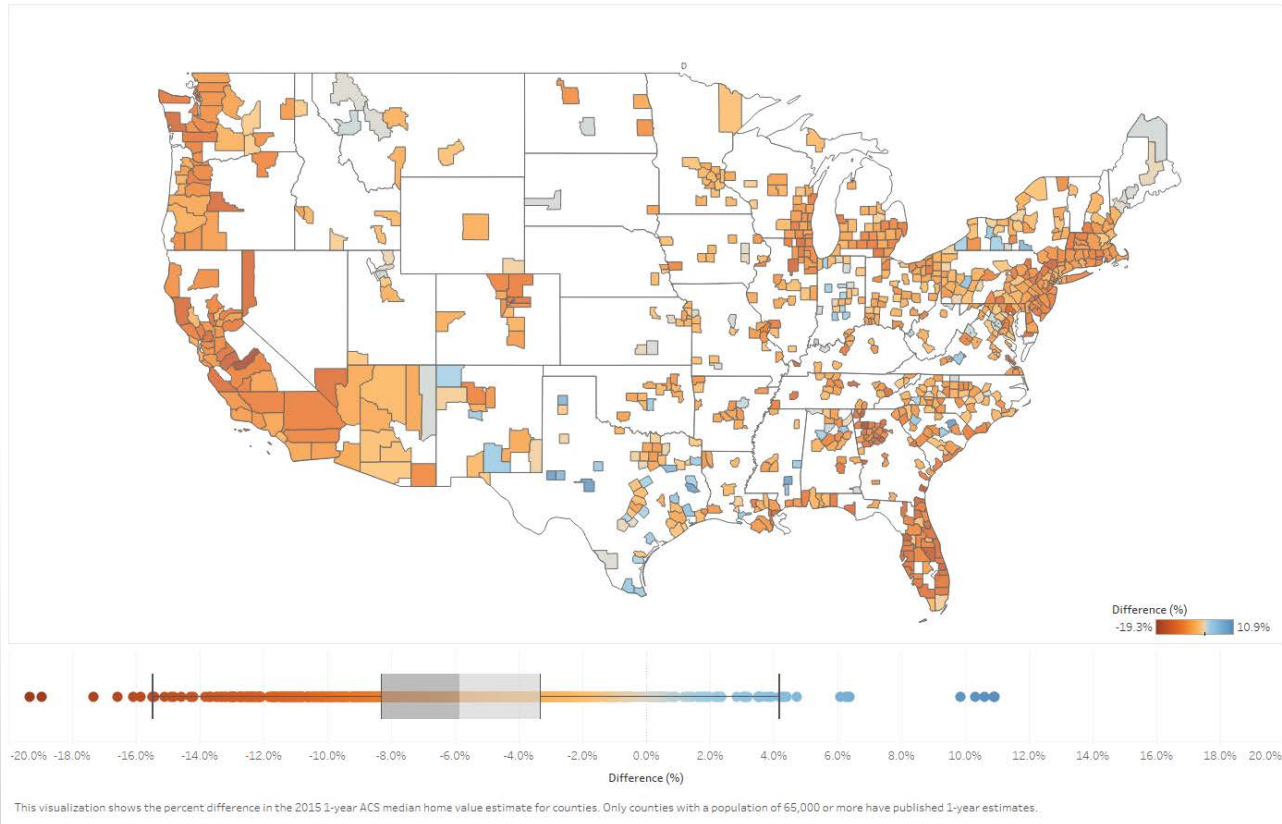
Burden Reduction

Burden defined as receiving the survey question

- Using ARs greatly reduces respondent burden for all 4 items

- Able to skip question for the following percentage of ACS responders:
  - Acreage - 52.0%
  - Property Value - 47.2%.
  - Year Built - 37.5%
  - Real Estate Tax - 53.5%

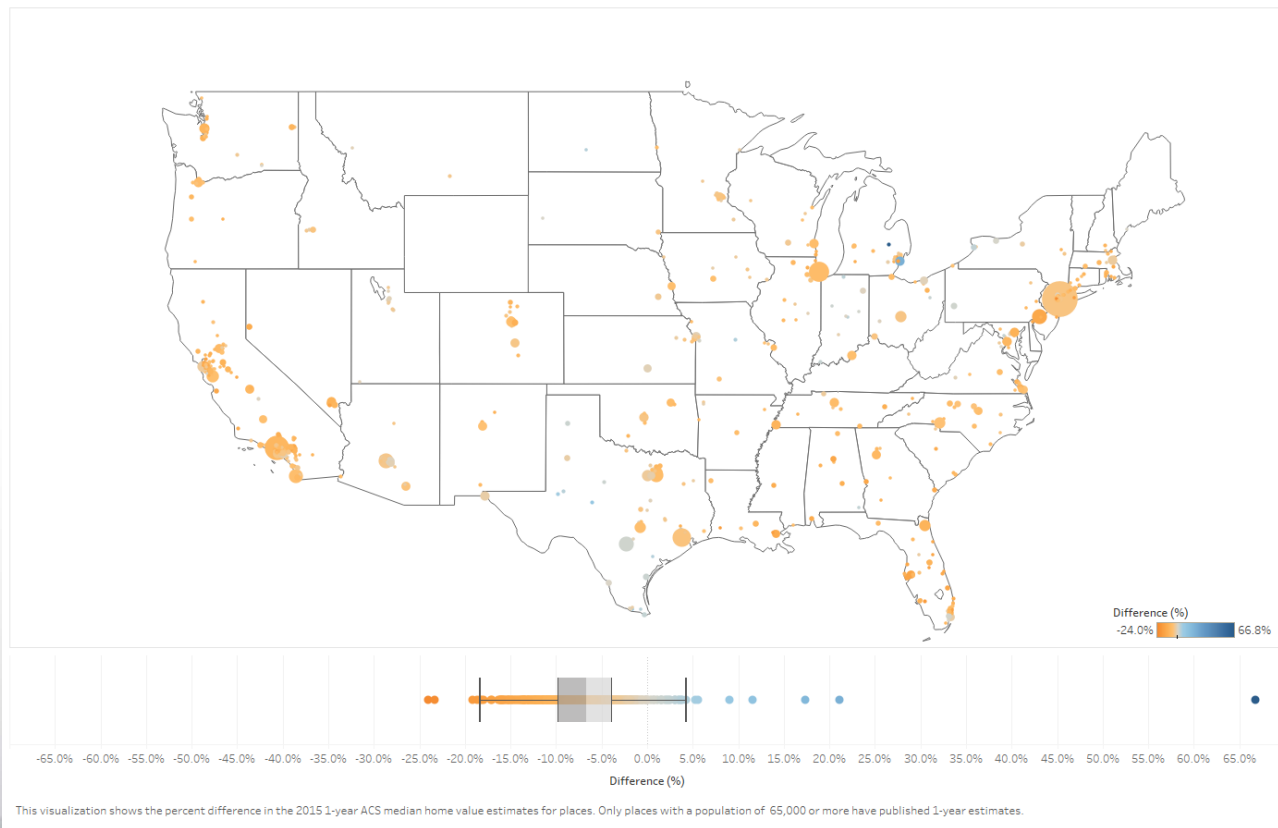# Percent Difference in Median Property Value: Test minus Control

State

This visualization shows the percent difference in the 2015 1-year ACS median home value estimates for states.

Percent Difference in Median Property Value: Test minus Control

County

This visualization shows the percent difference in the 2015 1-year ACS median home value estimate for counties. Only counties with a population of 65,000 or more have published 1-year estimates.

Percent Difference in Median Property Value: Test minus Control

Place

This visualization shows the percent difference in the 2015 1-year ACS median home value estimates for places. Only places with a population of 65,000 or more have published 1-year estimates.
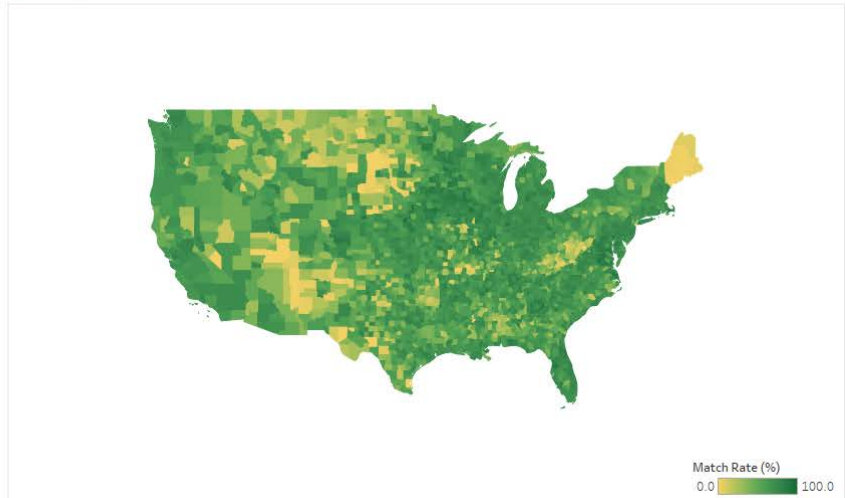
Burden Reduction Visualized
Overall Match Rate

What percentage of ACS households match to an administrative record?
Overall, 65.4% of households that responded to the ACS had a match to an administrative record that contained one of the four items.
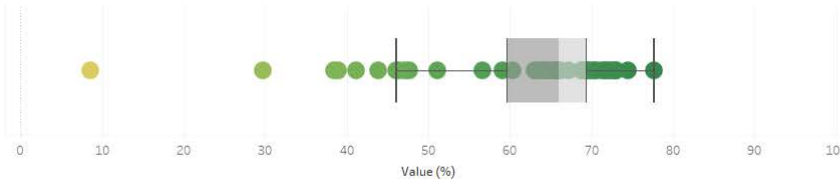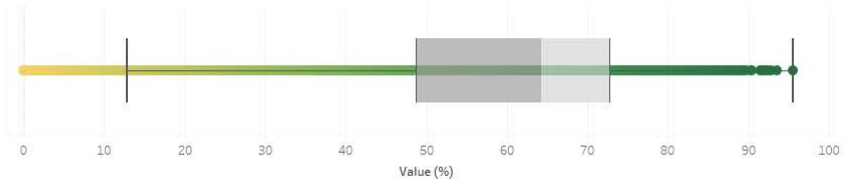
This visualization shows the percentage of households that responded to the ACS that we were able to match to an administrative record containing at least one of the four items.

Burden Reduction Visualized
By State and Variable

What percentage of ACS households have an admin record match in each state?
Sorted by percentage of household with any admin record matches
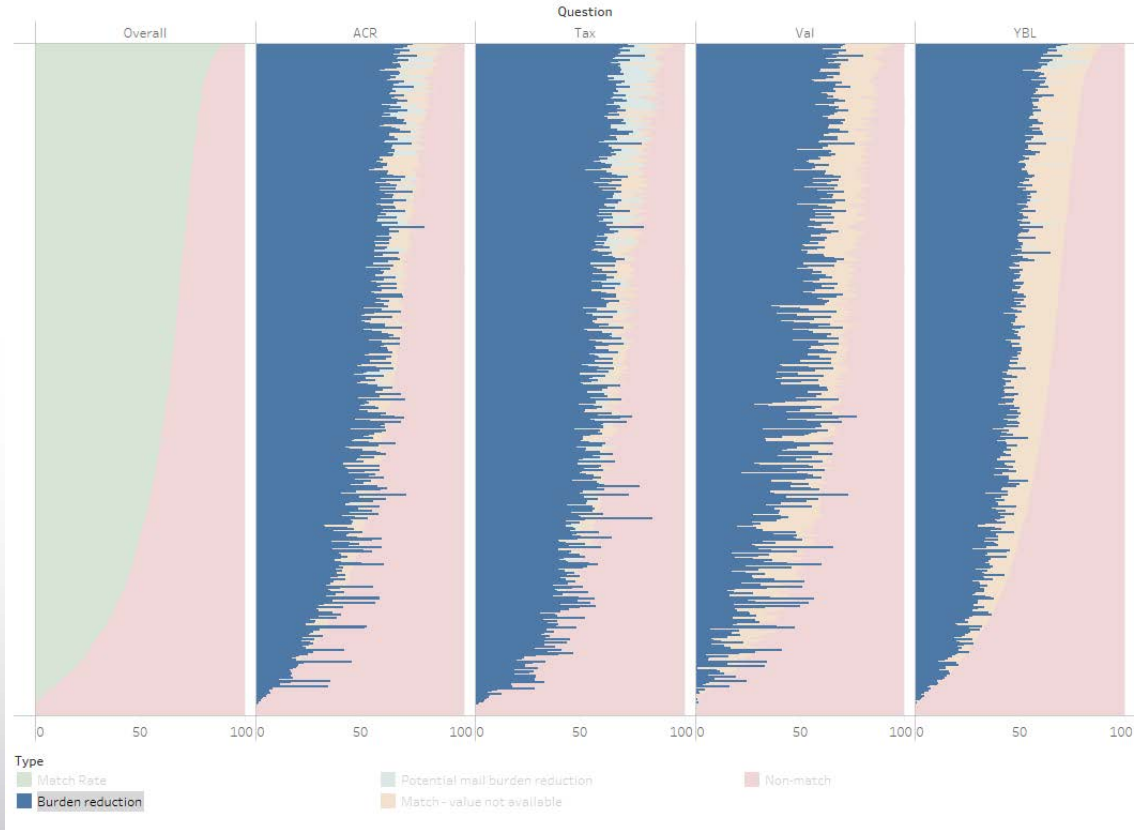
Burden Reduction Visualized
By County and Variable

What percentage of ACS households have an admin record match in each county?
Sorted by percentage of household with any admin record matches

Question

Overall | ACR | Tax | Val | YBL

Type
- Match Rate
- Burden reduction
- Potential mail burden reduction
- Match - value not available
- Non-match

Burden Reduction Visualized
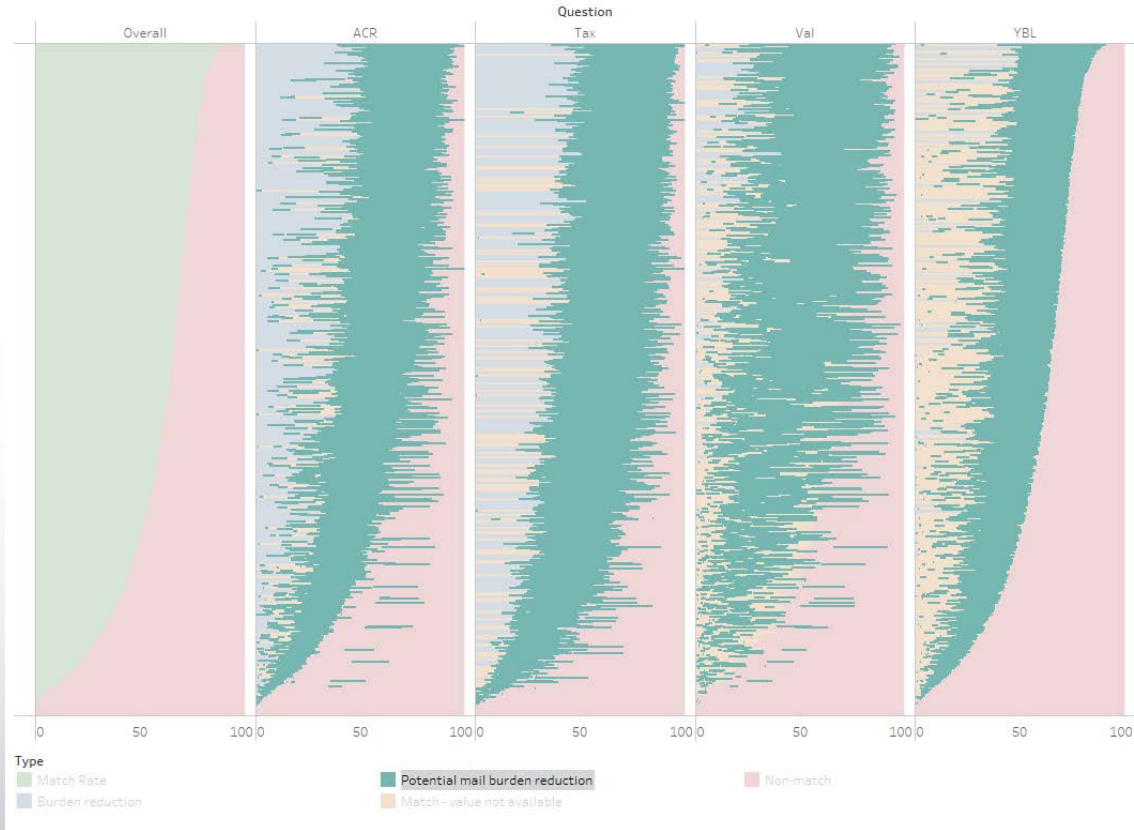## By County and Variable – Burden Reduction

What percentage of ACS households have an admin record match in each county?
Sorted by percentage of household with any admin record matches

Burden Reduction Visualized
By County and Variable – Match, no value available

What percentage of ACS households have an admin record match in each county?
Sorted by percentage of household with any admin record matches

Burden Reduction Visualized
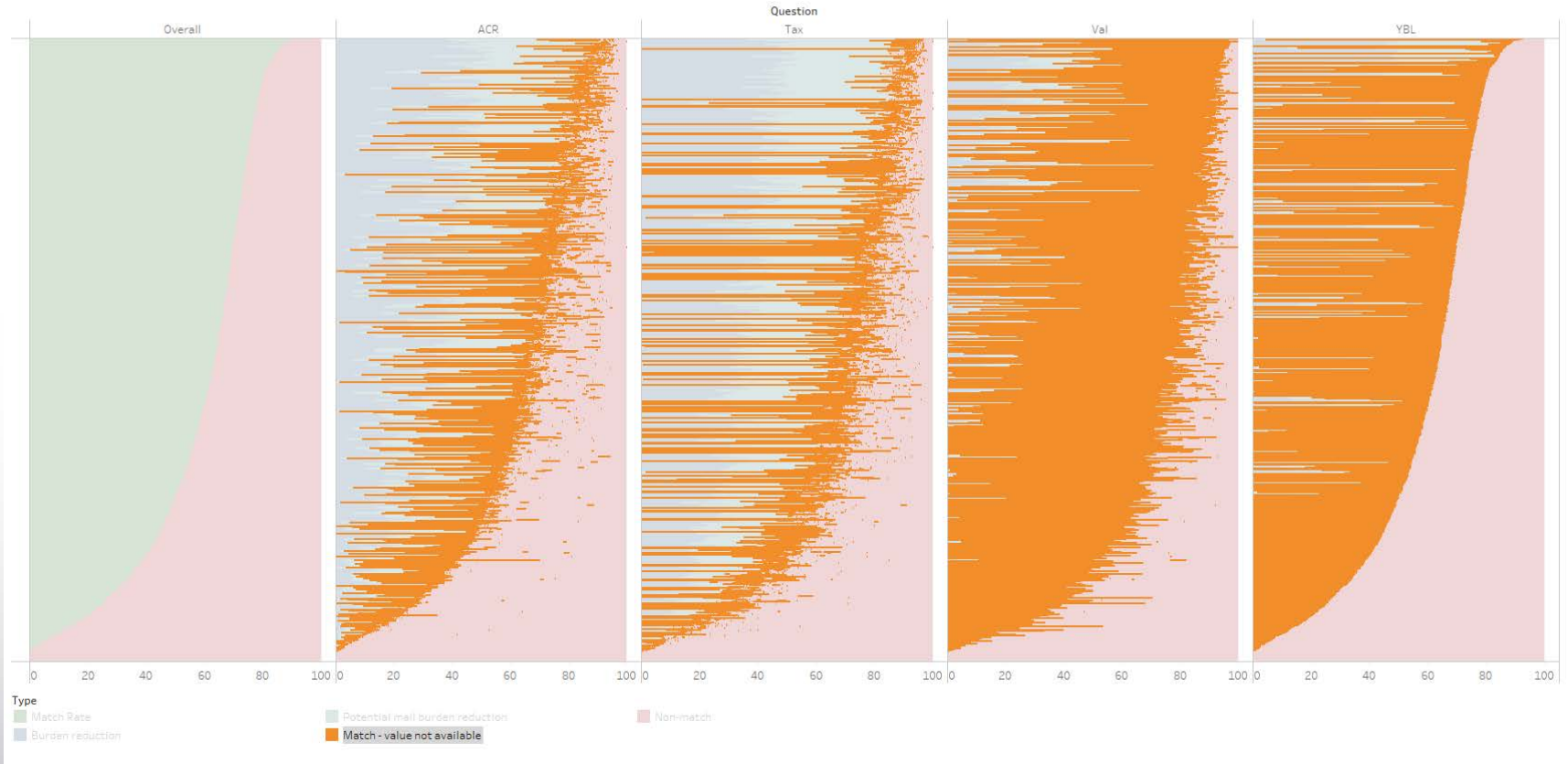## By County and Variable – Match, no value available

What percentage of ACS households have an admin record match in each county?
Sorted by percentage of household with any admin record matches

Type
Match Rate
Burden reduction
Potential mail burden reduction
Match - value not available
Non-match

Burden Reduction Visualized
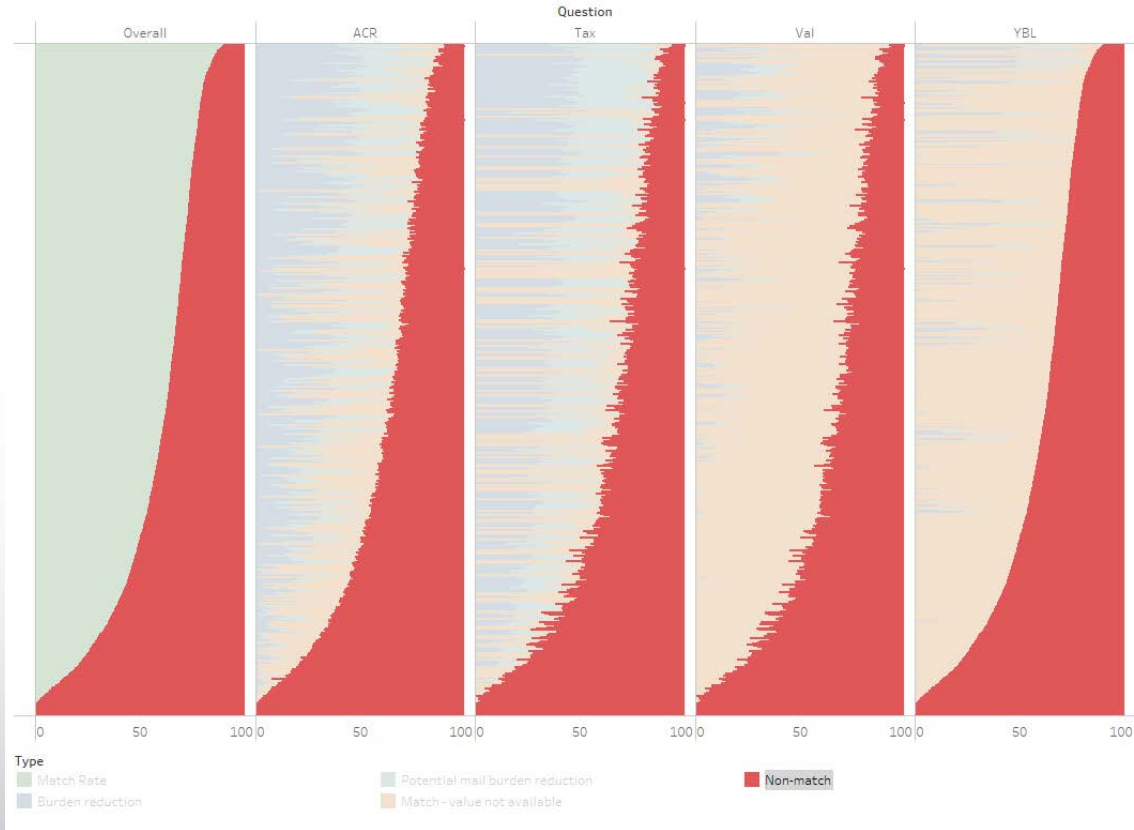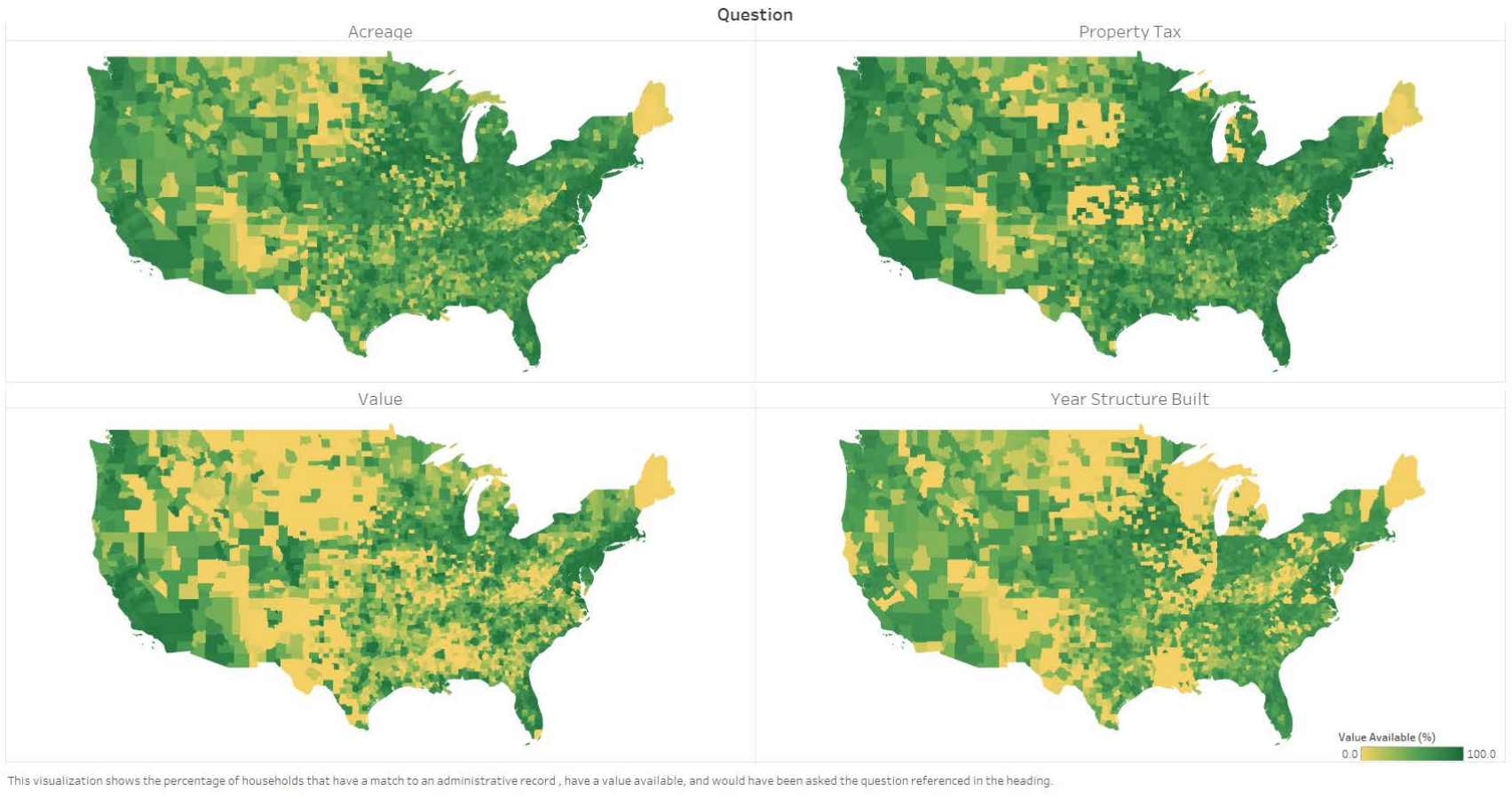## By County and Variable – Match, no value available

What percentage of ACS households have an admin record match in each county?
Sorted by percentage of household with any admin record matches

Burden Reduction Visualized
By County and Variable – Geographic View

Question

Acreage | Property Tax

Value | Year Structure Built

Value Available (%)
0.0 | 100.0

This visualization shows the percentage of households that have a match to an administrative record, have a value available, and would have been asked the question referenced in the heading.

Challenges: Source of ARs

- Rely on outside vendor, CoreLogic (CL)

- Best source of property value data comes from CL's Automated Valuation Model (AVM) and the methodology is unknown

- Collection of AR data differs among jurisdictions

- Data not available for some areas and types of housing units

- Often a time lag between survey year and AR

Challenges: Incorporating AR in the ACS

- Sometime unable to match AR to ACS
- Adapting all modes – difficult for mail mode
- Involves changes to entire survey life cycle
  - Significant resources
  - Risky without thorough testing
- Using AR complicates data disclosure avoidance rules

Preliminary Conclusions

- Significantly reduces respondent burden in terms of asking survey items
- Significant improvement in item allocation rates
- Must overcome several challenges
  - Source of AR
  - Incorporating in ACS
- Admin data is different from ACS response data
  - Break-in series
  - Using two estimate sources - that are different and potentially unfairly represent some geos
  - Inadvertently impacts other survey items