shadac
STATE HEALTH ACCESS DATA ASSISTANCE CENTER

USING SMALL AREA ESTIMATES FOR ACA OUTREACH

2014 AcademyHealth
San Diego, California
June 10

Robert Wood Johnson Foundation

Acknowledgements

- Supported by a grant from the Robert Wood Johnson Foundation to the State Health Access Data Assistance Center (SHADAC) at the University of Minnesota
- Co-Authors
  - Michel Boudreaux, Lead Author, SHADAC
  - Peter Graven, Oregon Health & Science University
- Special Thanks to:
  - Elizabeth Lukanen – Deputy Director, SHADAC
  - Karen Turner—Senior Programmer Analyst, SHADAC
  - Joanna Turner – Senior Research Fellow, SHADAC
  - Lynn Blewett – SHADAC Director

Outline

- Research Objective
- Background
  - ACS
  - ZCTAs
  - Reliability
- Methods
  - CAR model
  - Composite model
- Results
- Interactive Maps
- Findings

Motivation

- The public's knowledge of the ACA is poor
  - As of January 2014, 46% of the uninsured did not know about the availability of financial help for coverage
- Overall, the first open enrollment season was successful
  - But lots of variation across the states and a long way to go
- Success during the 2nd season will depend on outreach

Outreach

- Blanket media campaigns might not be enough
- Need to target the uninsured
- To do that efficiently
  - Need to know where the uninsured are
  - What kind of communities they live in
  - What institutions are present in the local community that can serve as access points

Research Objective

- PROBLEMS:
  - Small Area Health Insurance Estimates (SAHIE) are not granular enough
  - Direct zip code level estimates (ACS) can be unreliable
  - Accessing the data can be difficult
- GOALS:
  - Improve access to ZIP Code level estimates
  - Improve reliability of ZIP Code level estimates

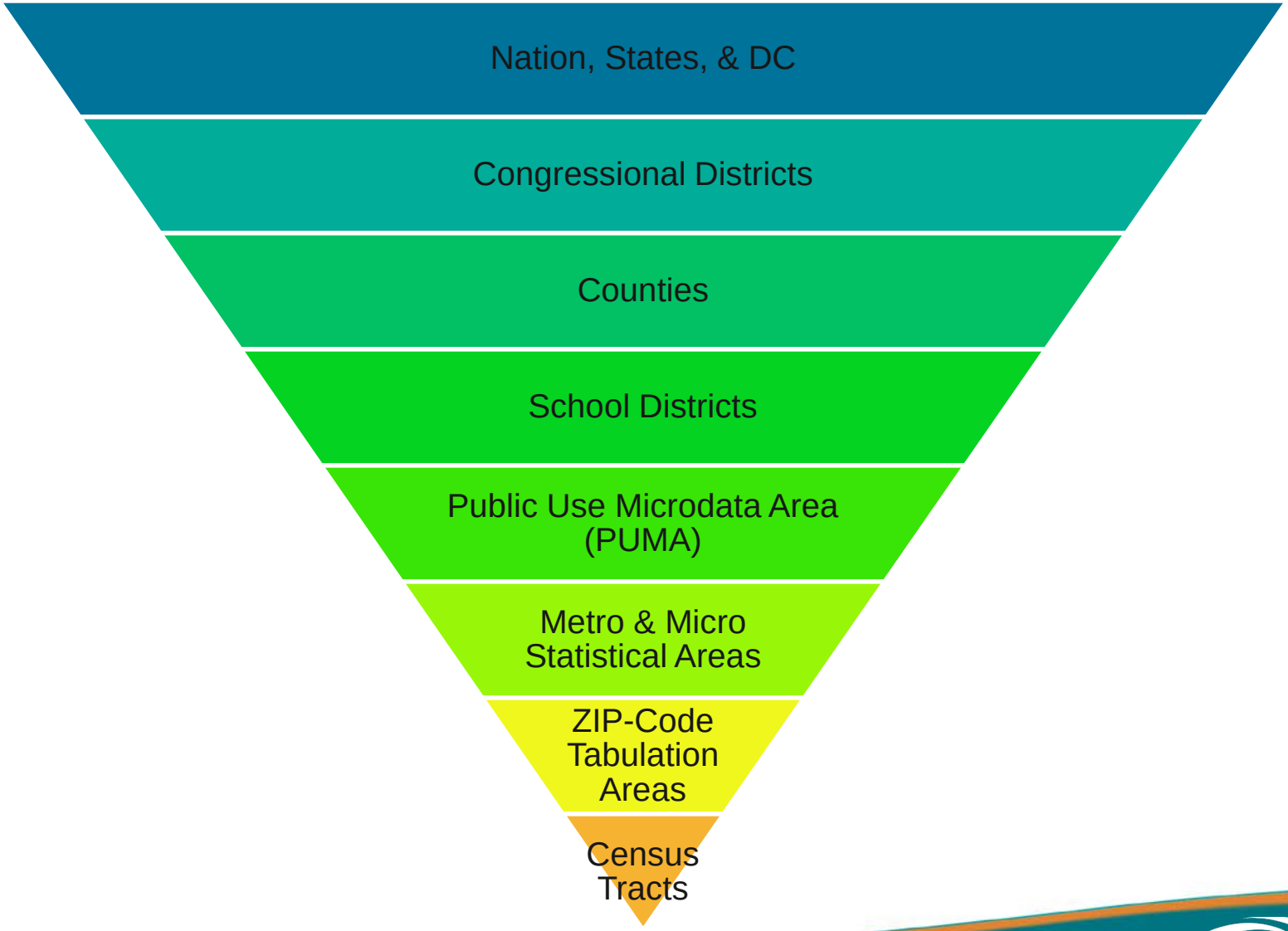
BACKGROUND

American Community Survey (ACS)

- General household survey conducted by the U.S. Census Bureau
  - Mandatory survey in 4 modes (mail, internet, phone, in person)
  - Collects sample in all counties or county equivalents in the U.S. every year

- Replacement for the “long form” of the decennial census
  - Collects detailed economic, social, demographic, and housing information annually instead of once every ten years
  - Collects information on health insurance status information at time of survey (produces point in time insurance estimate)

Identify Location of Potentially Eligible



Counties - Reliability

Minnesota Percent Uninsured Estimates by County

- ACS 2008-2012
  - Highest RSE is 18.4%
  - Average RSE is 8.0%
  - Average RSE top ten (sample size) counties 3.7%
  - Average RSE bottom ten (sample size) counties 12.2%
- SAHIE 2011
  - Highest RSE is 7.9%
  - Average RSE is 6.0%
  - Average RSE top ten (sample size) counties 5.1%
  - Average RSE bottom ten (sample size) counties 6.3%

Note: RSE is relative standard error (standard error/estimate)

ZCTAs – Reliability

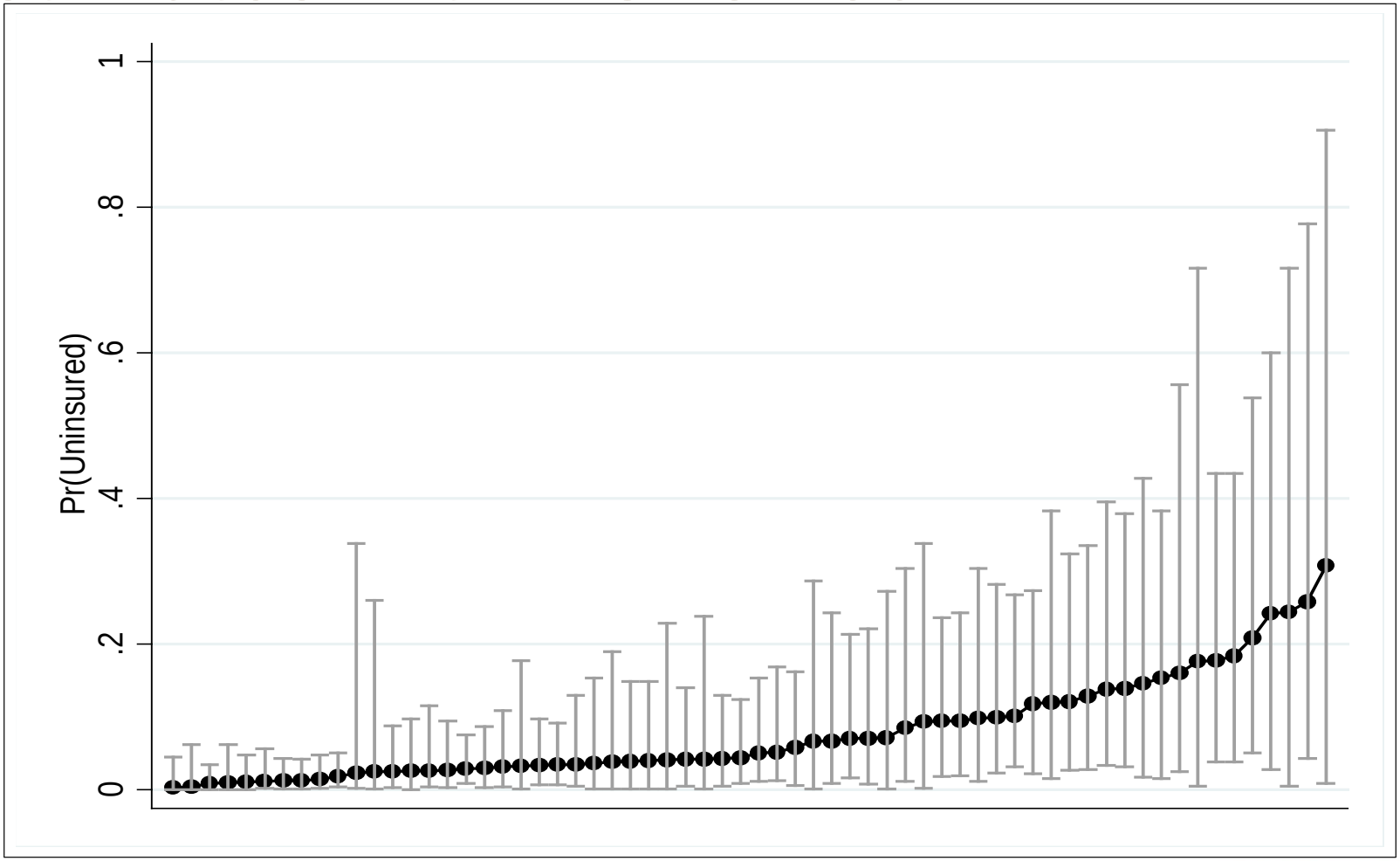
Non-Zero Percent Uninsured Estimates in Minnesota

- Highest RSE is 174%
- Average RSE is 27.6%
- Thirty percent of RSEs>28%
- Seven percent of RSEs>50%
- N ≈ 890

Non-Zero Percent Uninsured Estimates in U.S.

- Highest RSE is 509%
- Average RSE is 27.6%
- Thirty percent of RSEs>31%
- Eleven percent of RSEs>50%
- N ≈ 33,000

95% confidence intervals for ZCTA estimates with RSEs >50%: MN



METHODS

Two Methods for improving precision

- Conditional Auto-Regressive Model (CAR)
  - Advantages: Established method in the statistics literature
  - Disadvantages: High level of complexity, difficult to scale and apply to other types of estimates
- Modified Composite Method
  - Advantages: Easy to scale and apply to different types of estimates
  - Disadvantages: New approach so not peer reviewed

CAR Model

- Auxiliary data (covariates) improves prediction
  - $r_z = \alpha + \beta X_z + v_z$

- Borrows strength from neighbors
  - Creates term for average value of adjacent neighbors
  - $v_z | v_{-z}, \sigma_v^2 \sim N\left(\sum_{j \in \delta_z} \frac{v_j}{|\delta_z|}, \frac{\sigma_v^2}{|\delta_z|}\right)$

Composite Model

- Rough approximation of a composite estimator

$$Comp_{zc} = wt_{zc} * ZIPRATE_{zc} + (1 - wt_{zc}) * COUNTY_c$$

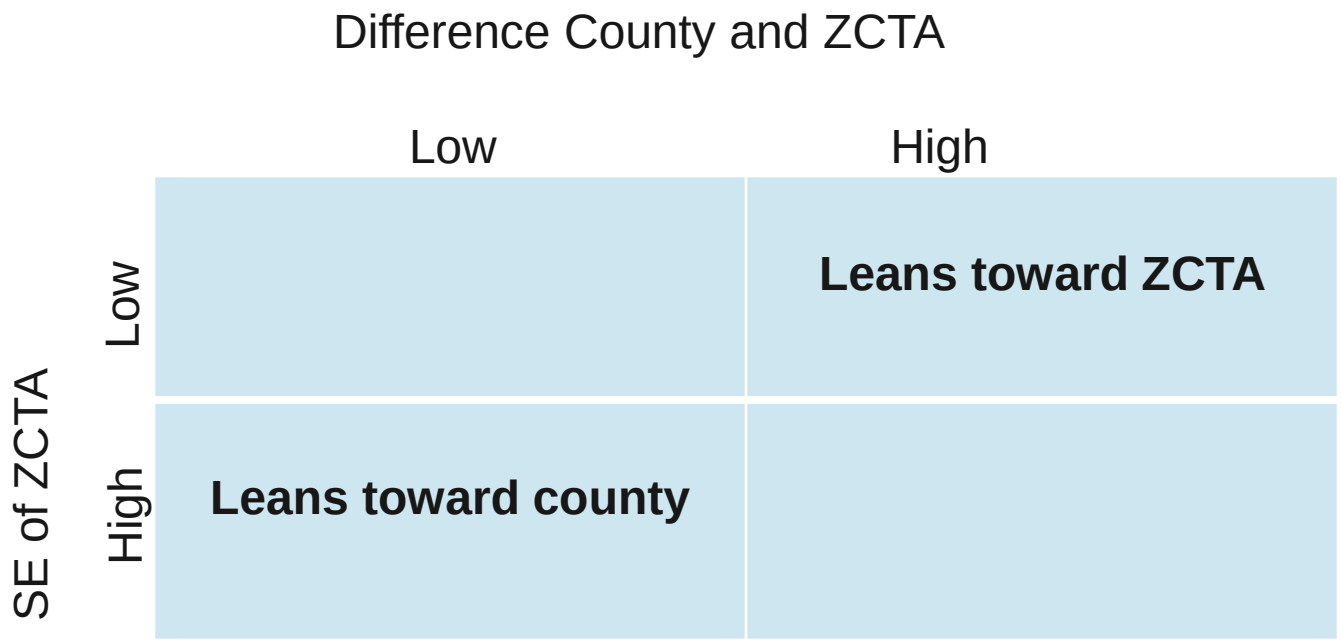
where wt=weight

$$wt_{zc} = (County_c - ZIPRATE_{zc})^2 / Total\ Error_{zc}$$

$$Total\ Error_{zc} = (County_c - ZIPRATE_{zc})^2 + SE^2{}_{zc}$$

See Rao (2003)

Composite Model Intuition



Model Results

CAR Model Results	Coef.	95% Credible Interval
% White	-0.02	(-0.02,-0.017)
% Living w/Kids	-0.01	(-0.016,-0.009)
SD of Spatial Effects	0.32	

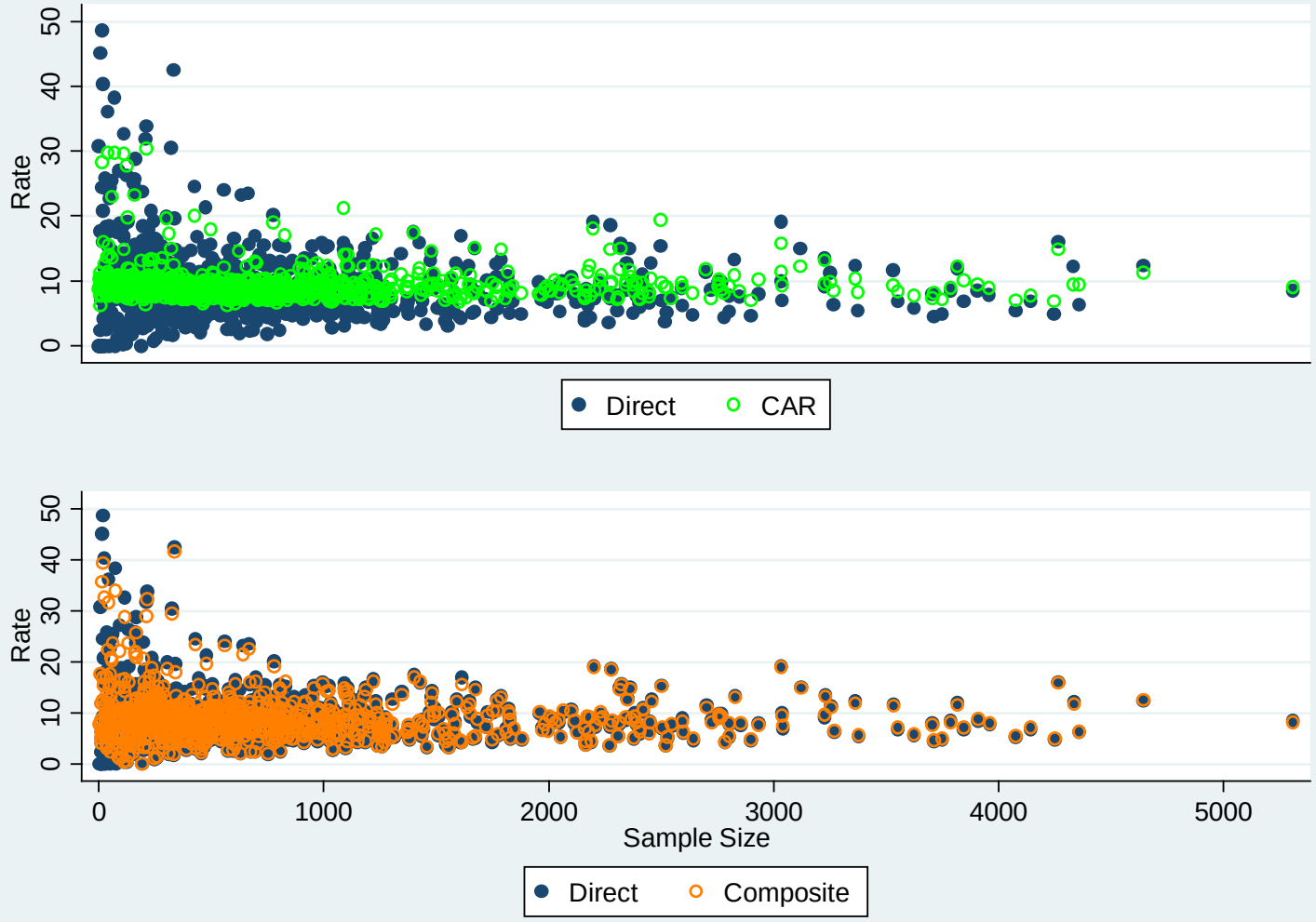
Composite Model	Mean	SD
Weight	.53	.32

Comparison of Methods

Average across estimates: Minnesota

	Direct	CAR	Composite
Rate, %	9.3	9.5	9.1
SE	2.6	1.1	1.6
RSE, %	27.6	11.1	20.9
RSE>30, %	28.7	0.1	11.2
RSE>50, %	8.5	0.1	3.4

Distributions by Sample Size

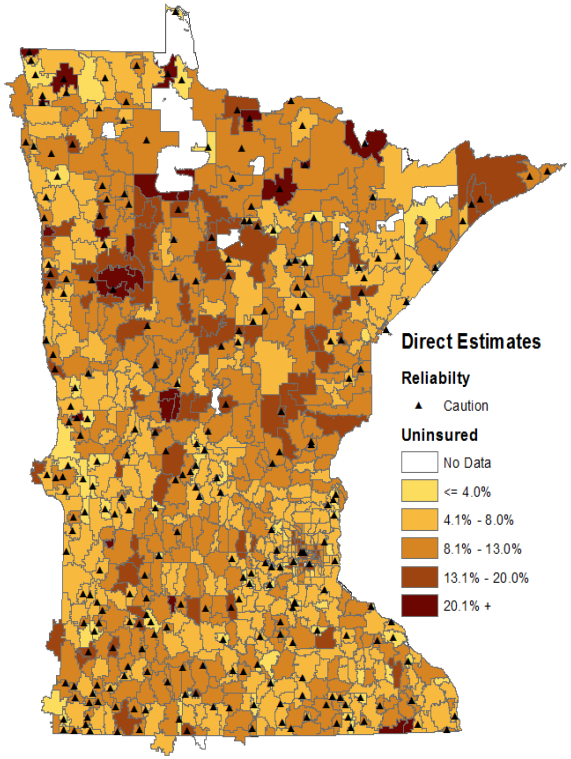


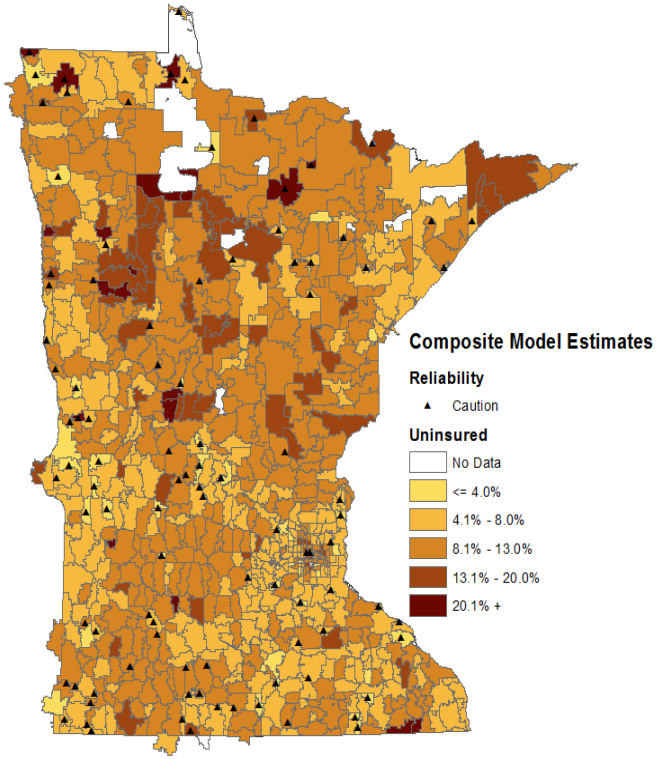
Which method is better?

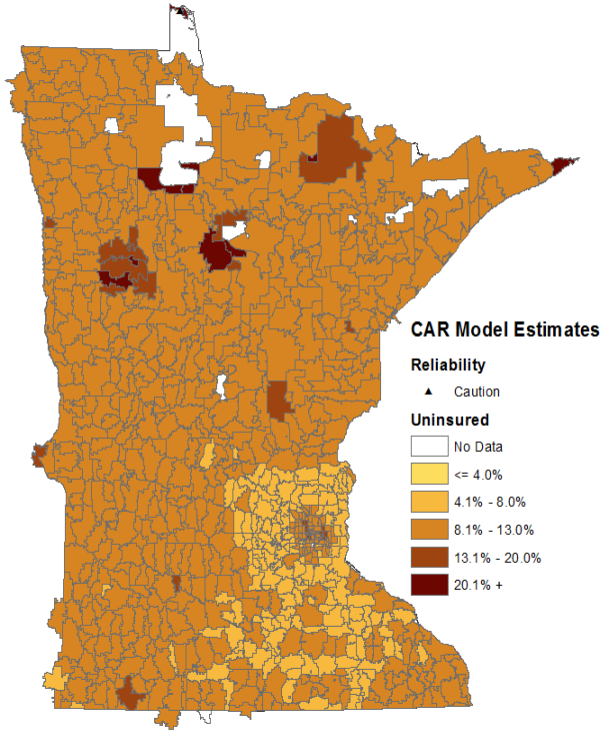
	Direct	CAR	Composite
Complexity	Low	High	Low
Scalability	Easy	Hard	Easy
Reliability	Not very reliable	Very reliable	More reliable but still not great
Bias	?	?	?

Maps!

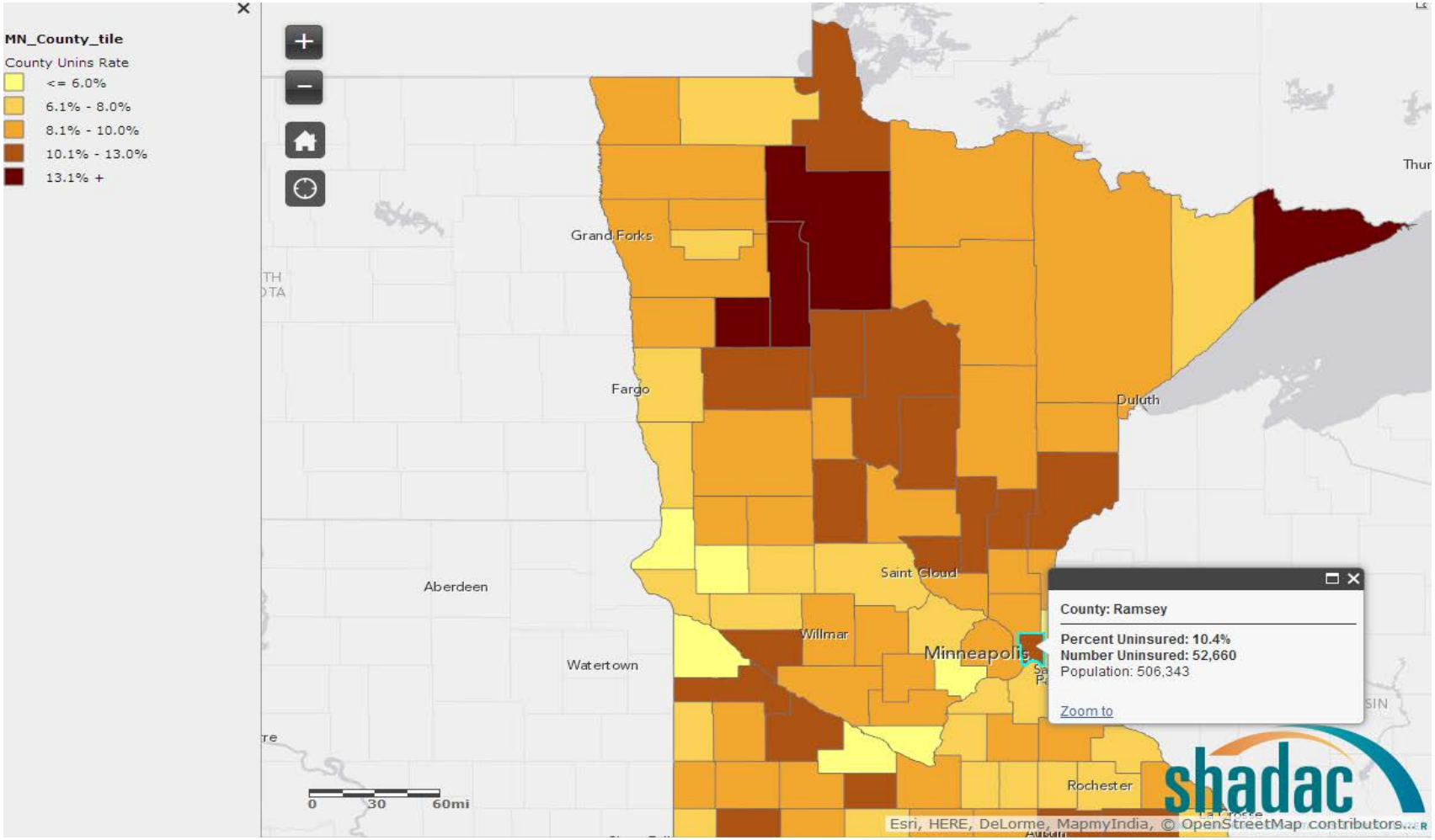
Minnesota: Comparing Estimates



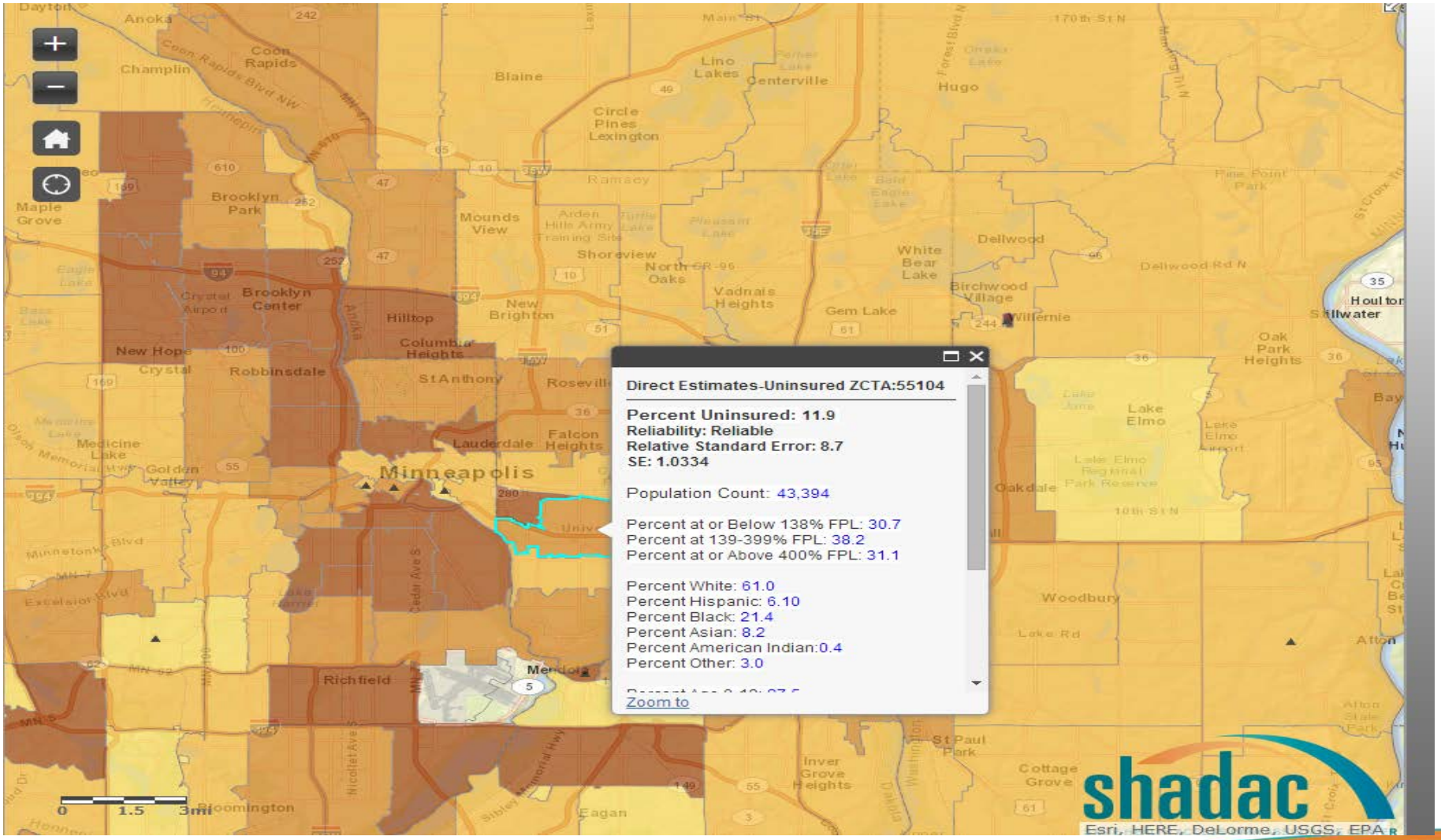




Start with the County



Then look at ZCTA estimates



Findings

- Providing uninsurance estimates at the ZCTA level is problematic both from the standpoint of reliability and accessibility
- Possible solutions to the problem of reliability is to use small area methods such as CAR or a moderated composite estimator
- CAR is the more established method and provides more reliable estimates but is complex and difficult to scale
- A potential compromise is to use the modified composite estimator but more testing is needed
- Interactive mapping can make these estimates available at the ZCTA level