



Design Effect Anomalies in the American Community Survey

Michel Boudreaux

Joint Statistical Meetings

July 29, 2012

San Diego, CA

Overview

- Background
- Irregularities in the DEFF of the ACS
- Potential drivers
- Practical significance for analysts

Design Effects

- A relative measure of sample efficiency

$$DEFF = \frac{\sigma_{complex}^2}{\sigma_{SRS}^2}$$

- Sensitive to design elements
 - Stratification (-)
 - Intra-cluster correlation (+)
 - Variance in the weights (+)
- Unique to each variable
- Typically 2-4

ACS Sample Design

- Separate design for HU and GQ
- Housing Unit Design
 - Frame: MAF
 - Each county chosen with certainty
 - Sub-county strata defined on population and RR
 - 1/3 Non-responders sampled for personal interviews
 - PUMS created as a systematic sample such that a 1% sample of each state is formed

Construction of Full Sample Weight

- Inverse of the probability of selection (BW)
- CAPI Sub-sample adjustment (SSF)
- Seasonal response adjustment
- Non-interview adjustment
- Mode bias adjustment
- Raking to control totals at the sub-county level

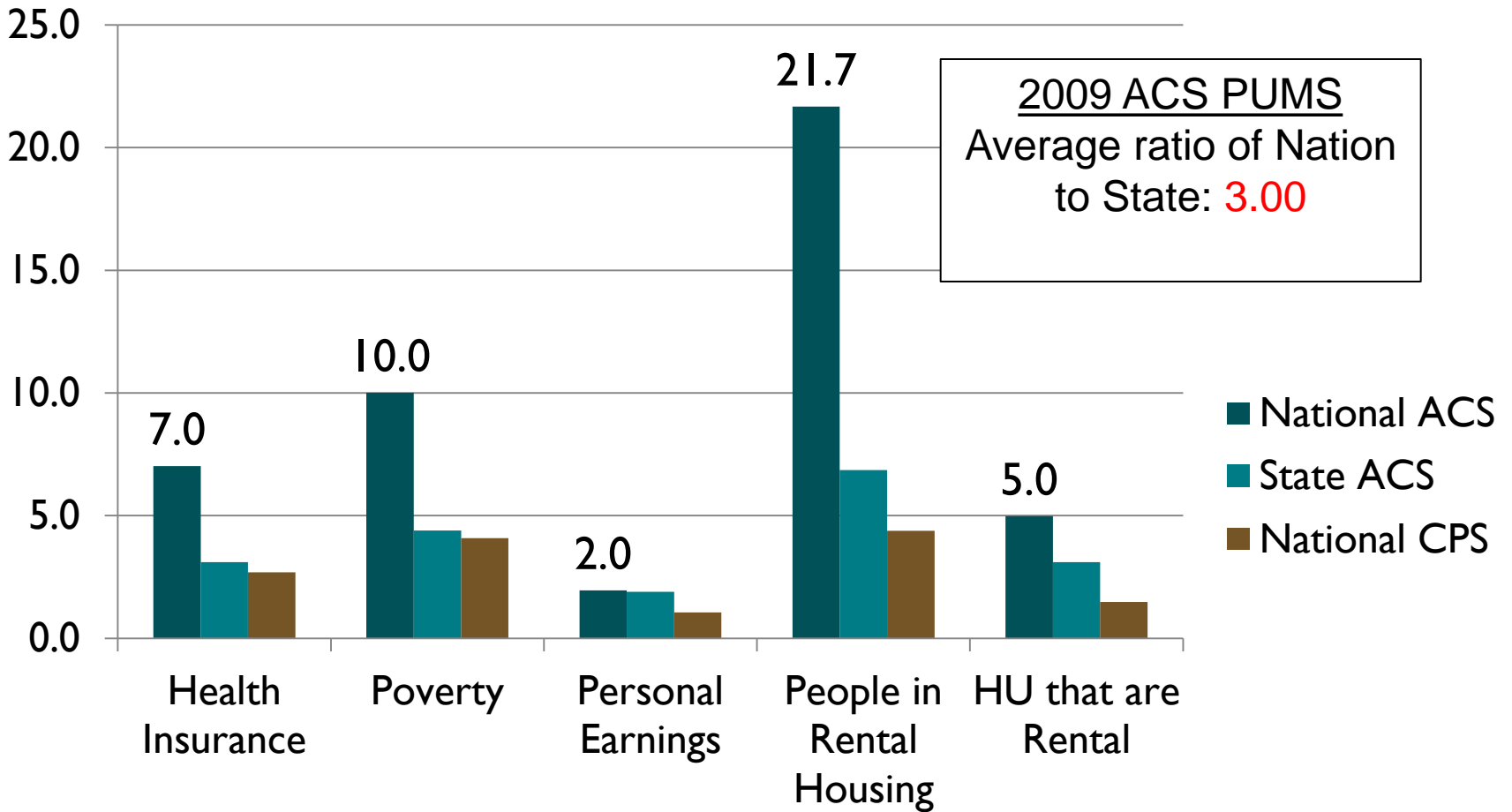
Complex Variance Estimation

- Successive Difference Replication
 - Similar to BRR w/Fay's adjustment
 - Geographic sort order is informative
- Replicate Weights (80 replicates)
 - 1 of 3 replicate factors applied to each case
 - 1.0 (50% of cases), 0.3, 1.7
 - Factor assigned using a Hadamard Matrix (RF)
 - BW adjusted with replicate factor
 - Weighting process repeated

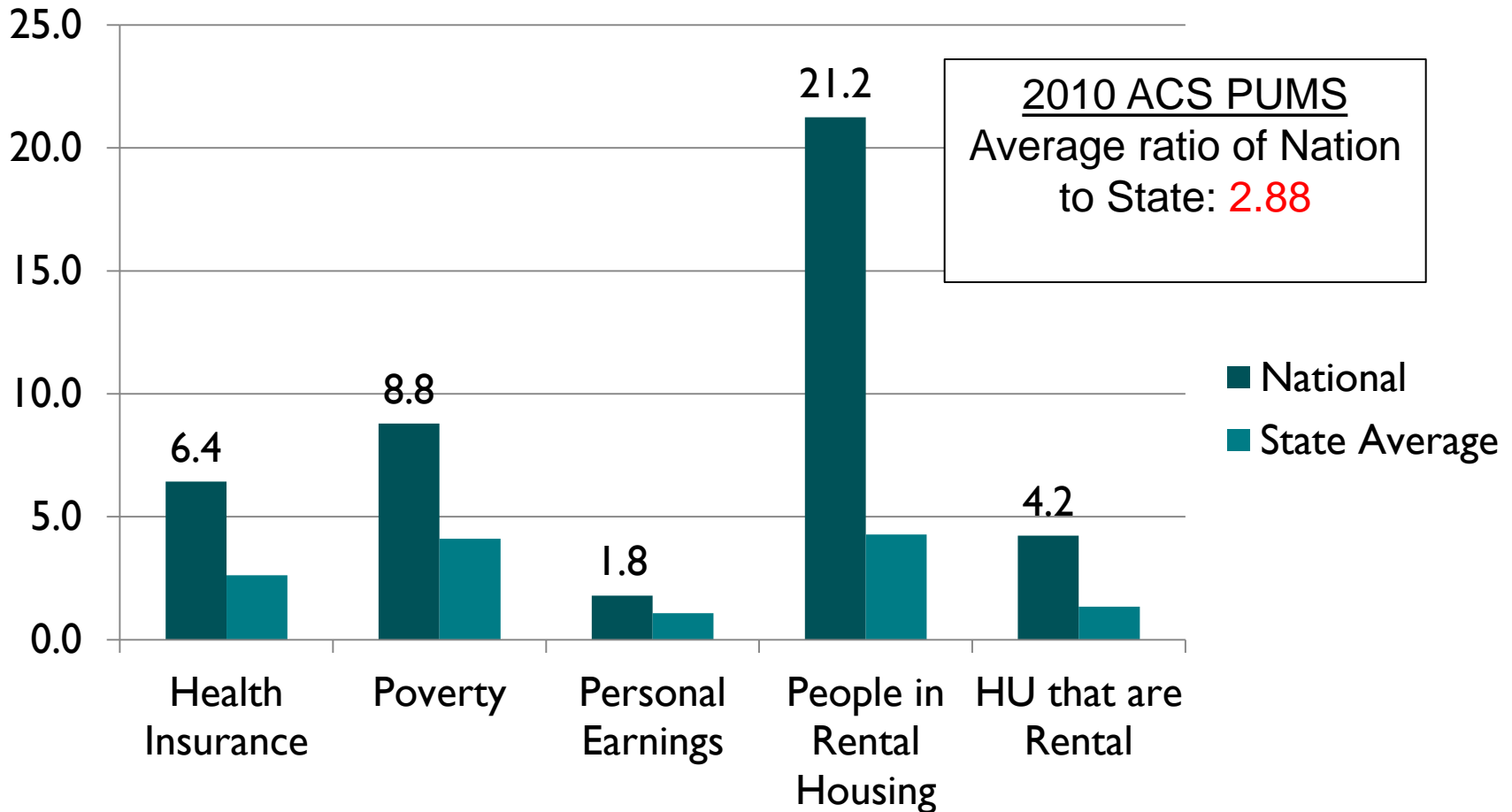
SDR Formula

$$\sigma_x = \sqrt{\frac{4}{R} \sum_{r=1}^R (x_r - x_0)^2}$$

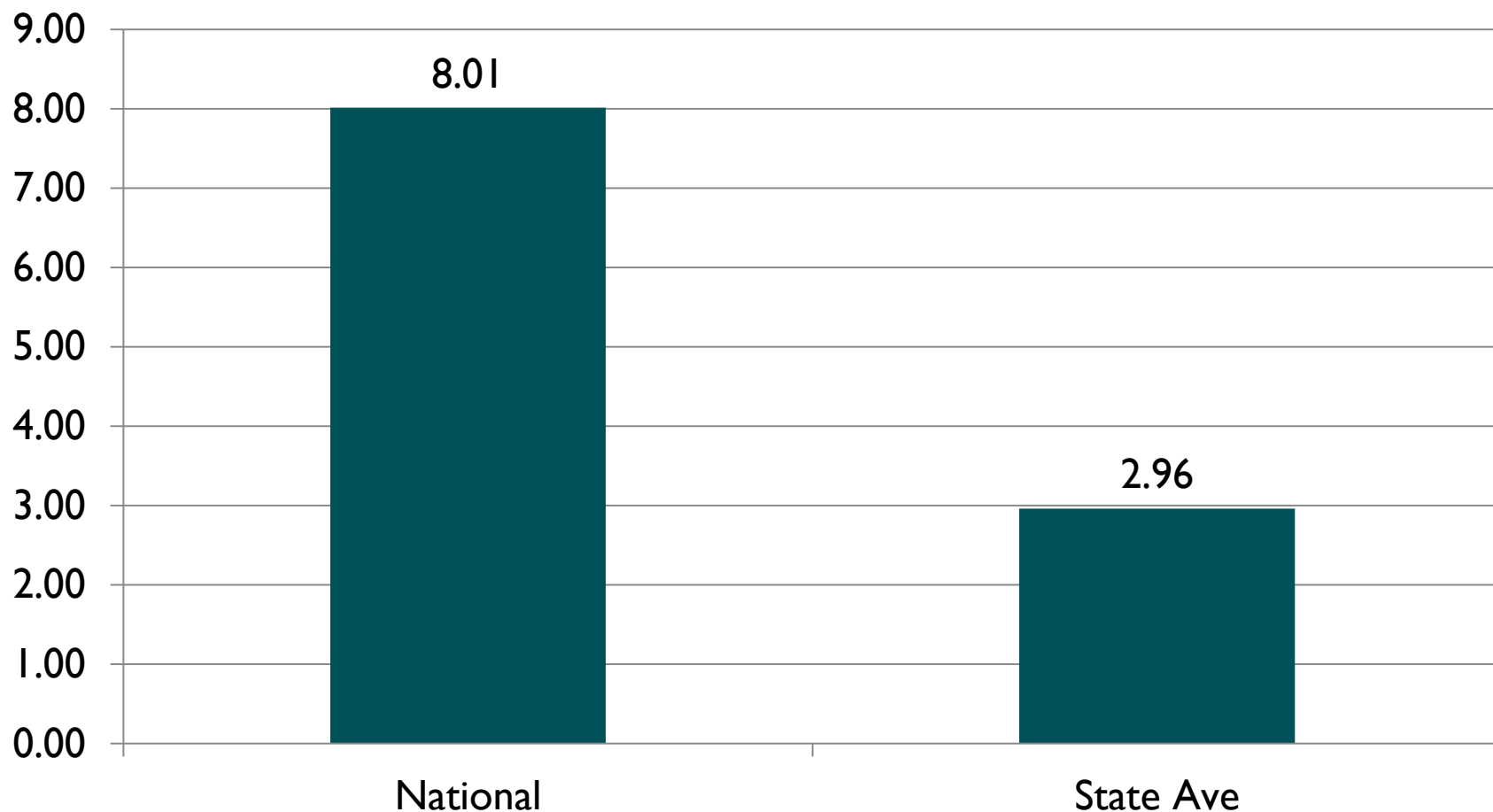
Design Effects in the 2009 PUMS



Design Effects in the 2010 PUMS

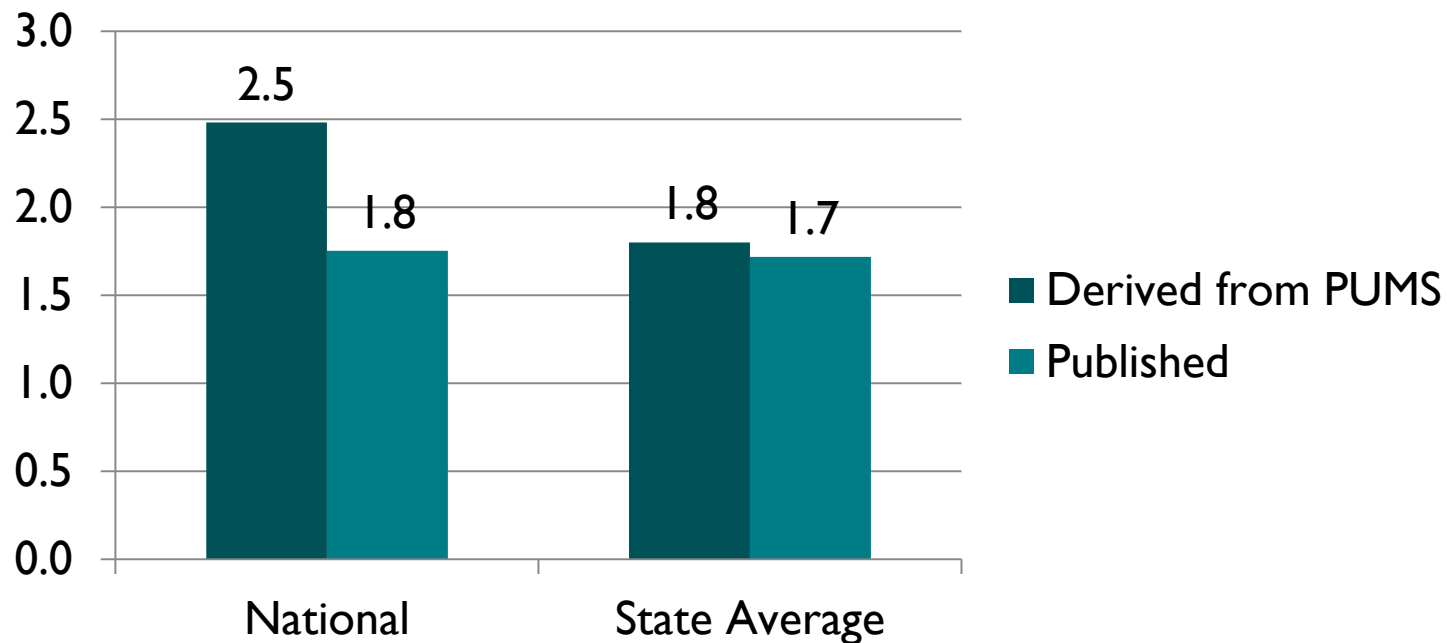


Design Effect for Health Insurance in 2008 Internal File (Derived from AFF)



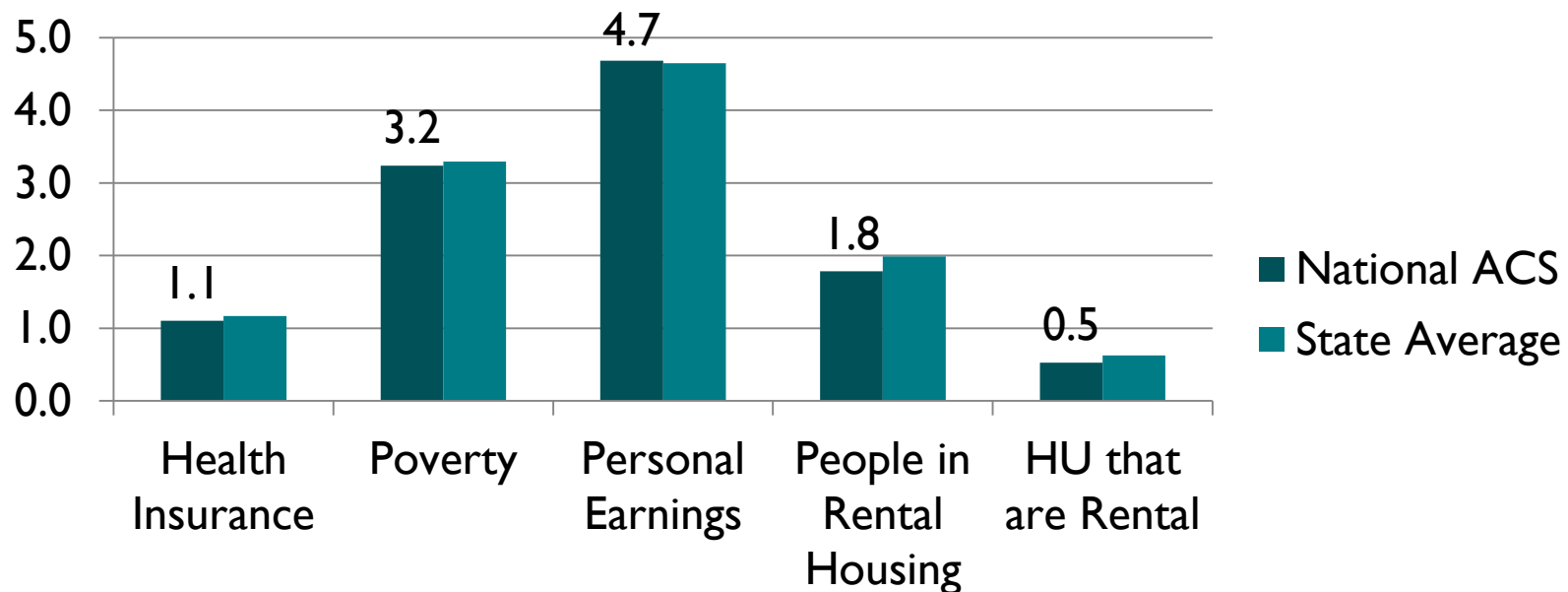
2008 Design Factors (Published vs. Derived)

- DF's are ratios of standard errors
 - Used as a ratio-adjuster in place of SDR



Using an Alternative Complex Variance Estimator

- Taylor Series
 - Strata: PUMA; Cluster: Household



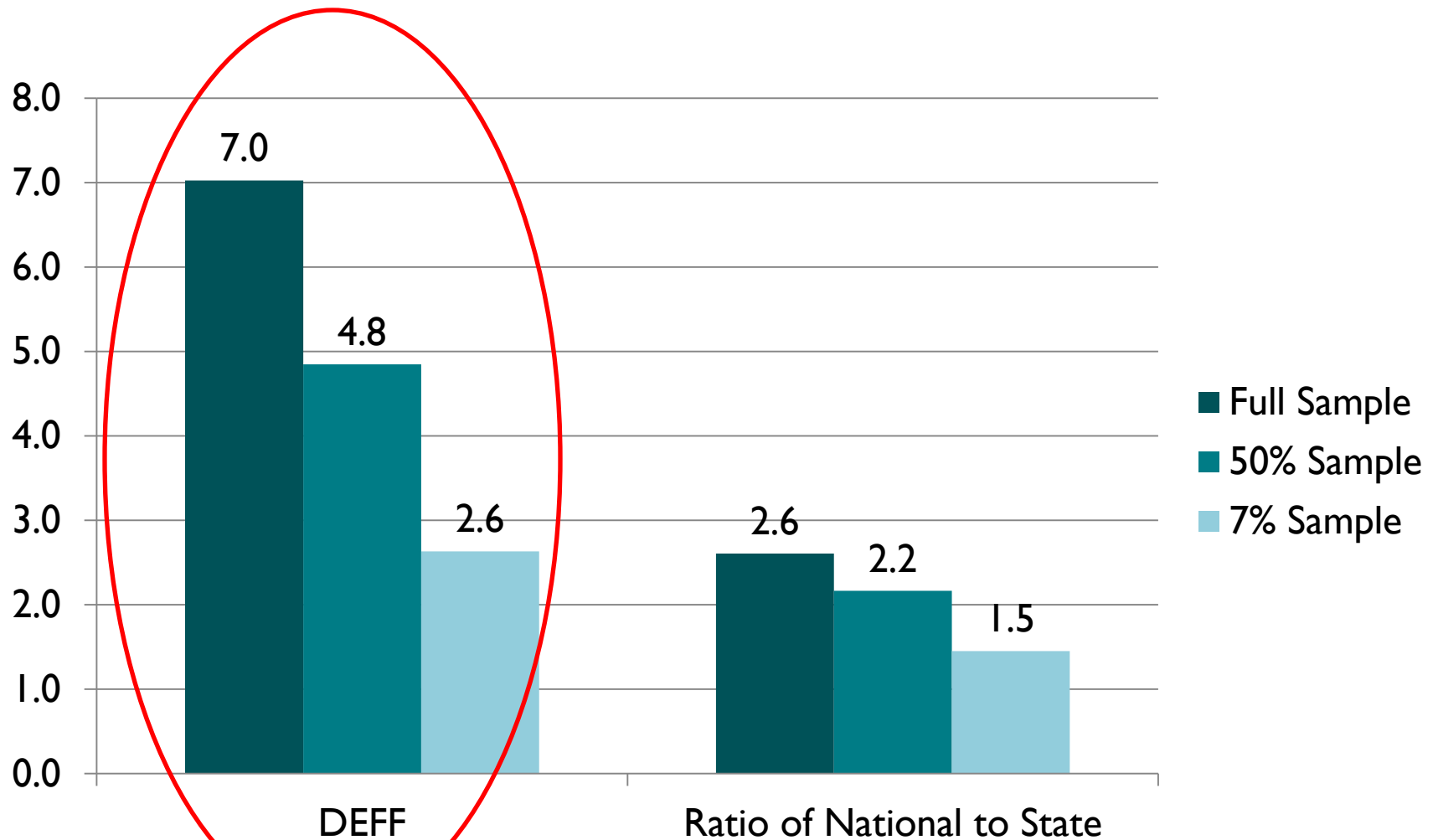
Summary of Anomaly

- National DEFF is far larger than expectation
 - Versus literature (Kish, 2003)
 - Versus state average
 - Versus CPS
- Consistently present
 - Across Variables
 - Across years and PUMS/internal file
- Important Exceptions
 - Personal Earnings (low correlation)
 - Published Design Factors
 - Taylor Series estimator

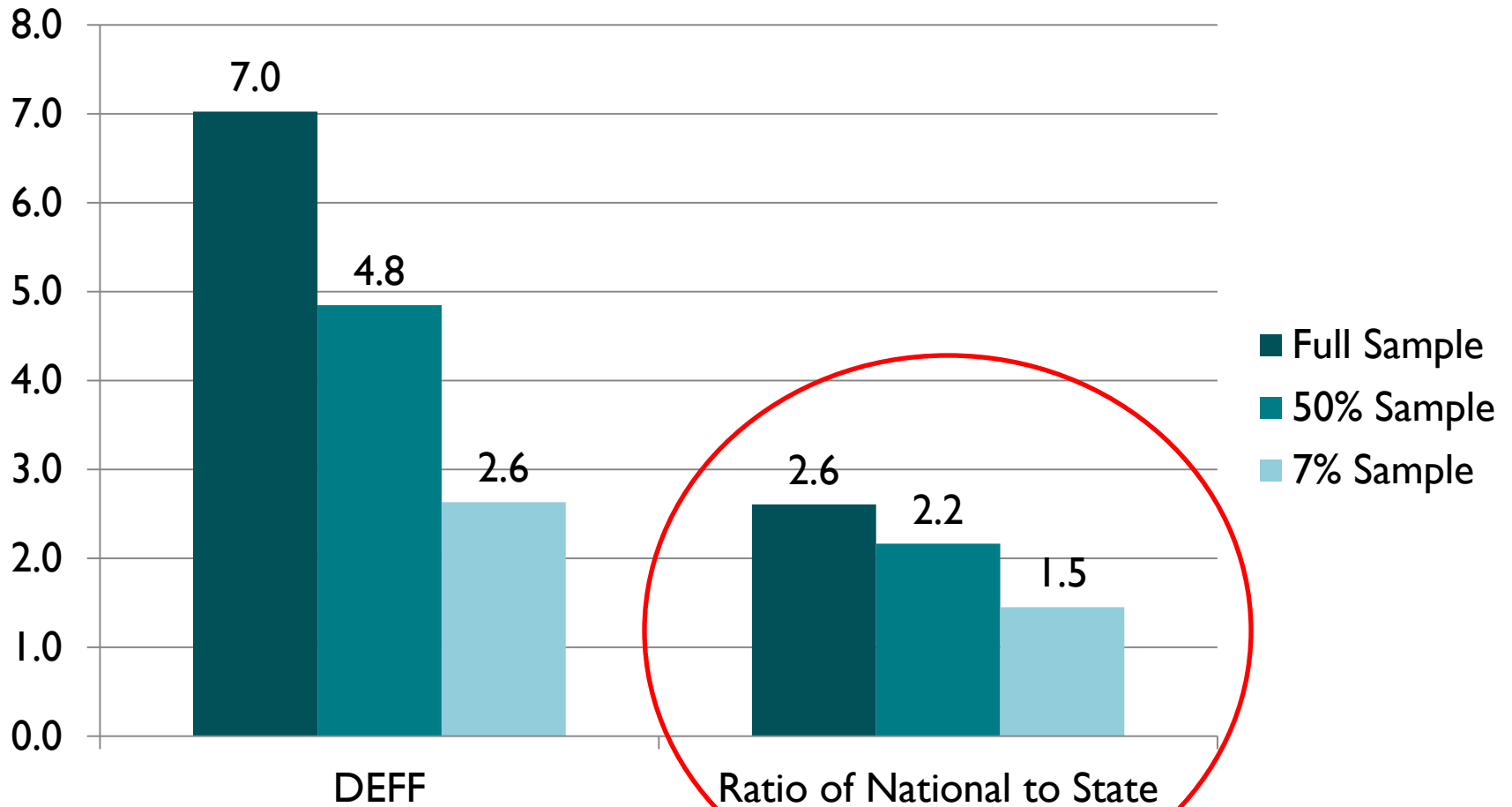
Potential Causes

- Over-sampling
 - PUMS sampling rate set in each state to 1%
 - Moderate geographic oversampling, but not likely
- Heterogeneity in the weights (I+L)
 - I+L in line with expectations; not correlated with geography (for full weight and replicates)
- Rounding of the weights
 - ACS rounds, CPS does not
 - Ruled this out
- Sample Size
 - Definitely differentiates states from nation

Results by sample size (2009)



Results by sample size (2009)



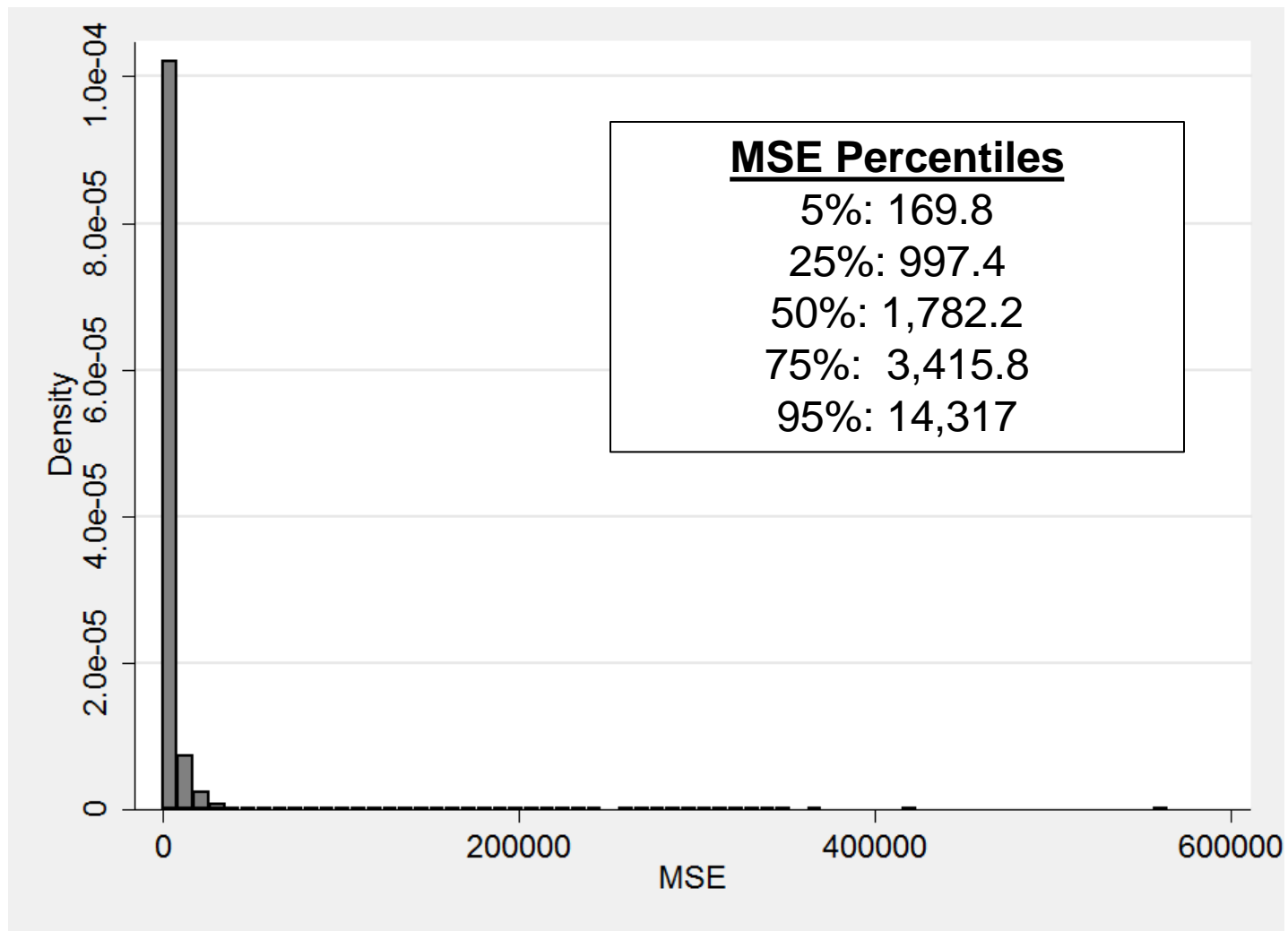
Why Sample Size?

- As sample size increases, the probability of capturing a meaningful outlier increases
- Only potential outlier in SDR versus Taylor Series is the variation *across* weights.

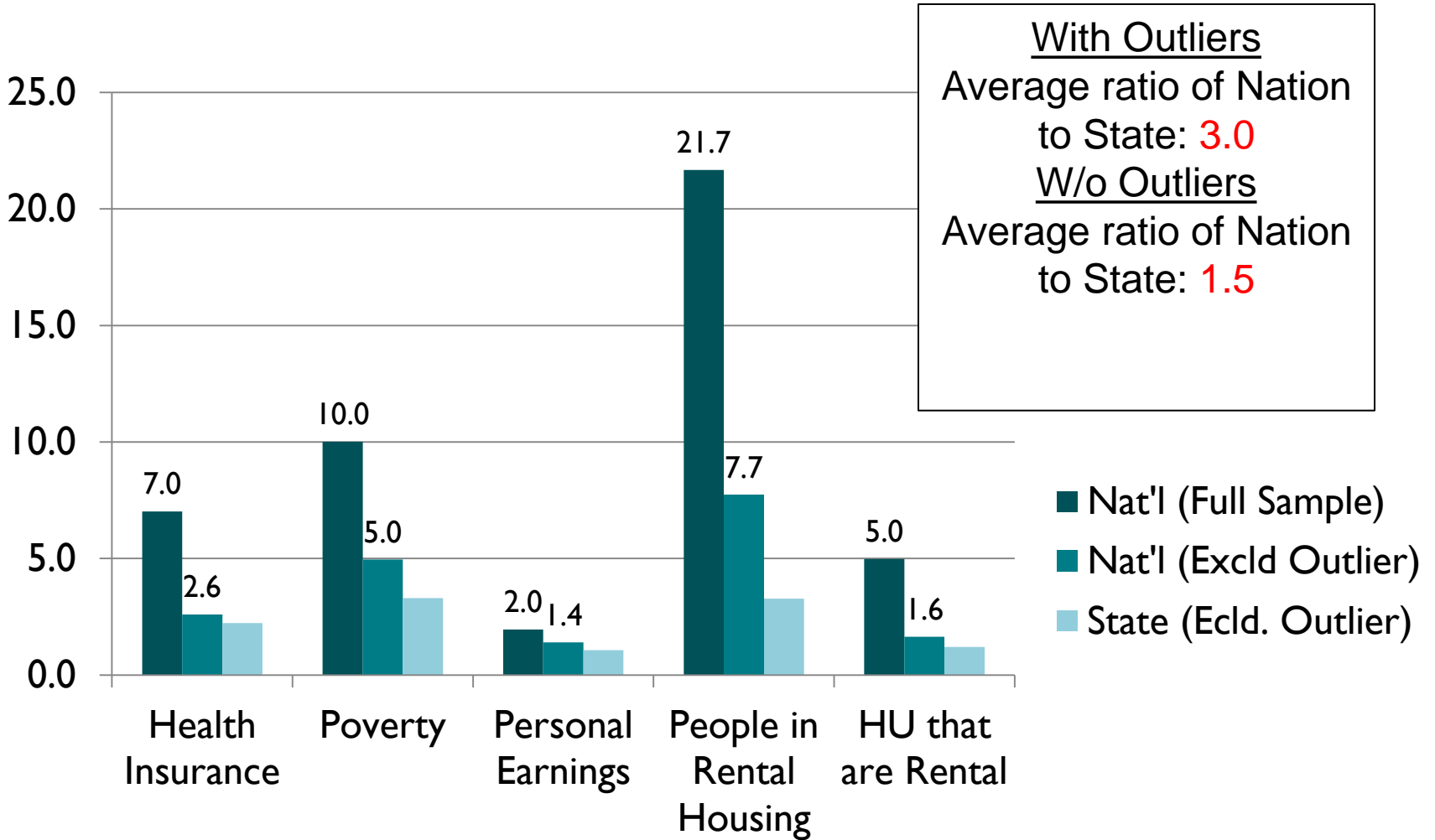
$$MSE_i = \frac{1}{80} \sum_{r=1}^{80} (w_{ir} - w_{i0})^2$$

For $i = (1 \dots n)$ and $r = (1 \dots 80)$

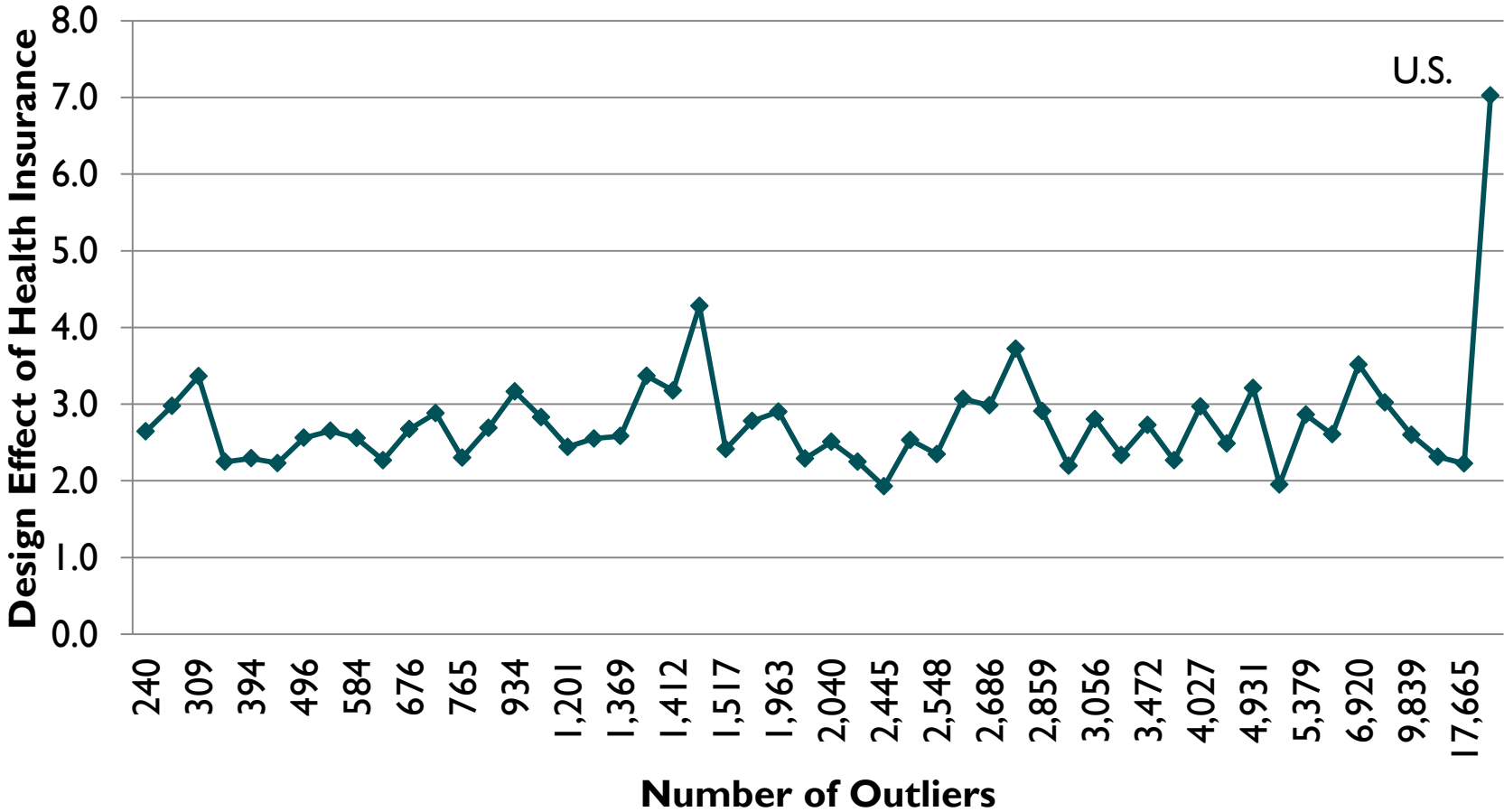
Histogram of MSE



Results w/o Upper 5% of MSE Distribution



Effect Appears Non-Linear



Outlier Correlates

	MSE	
	0-14,316	14,317+ (Outlier)
Full Sample Wt, Mean (SD)	89 (47.6)	322 (85.3)
Median Replicate, Mean (SD)	89 (47.8)	323 (85.9)
Mode/GQ*		
HU Mail, %	99.6	.34
HU CATI/CAPI, %	85.3	14.9
GQ, %	97.3	2.7
* Significant at $p < 0.05$; Remains significant after adjusting for age, race, and sex		
NOTE: CATI/CAPI is grouped together in PUMS		

One Potential Reason for Outlying MSE

- Recall the weighting strategy for the replicates
 - $BW * RF * SSF \dots$
- Potentially some correlation between RF and SSF that causes larger MSE in CAPI cases, relative to Mail/CATI

Limitations

- Our hypothesis that outlying MSE cause this anomaly fails to explain 2 results
 - Why do we fail to find an effect for characteristics with low intra-class correlations (earnings)?
 - The group (CATI/CAPI) with the highest rate and frequency of outliers does not have the highest DEFF

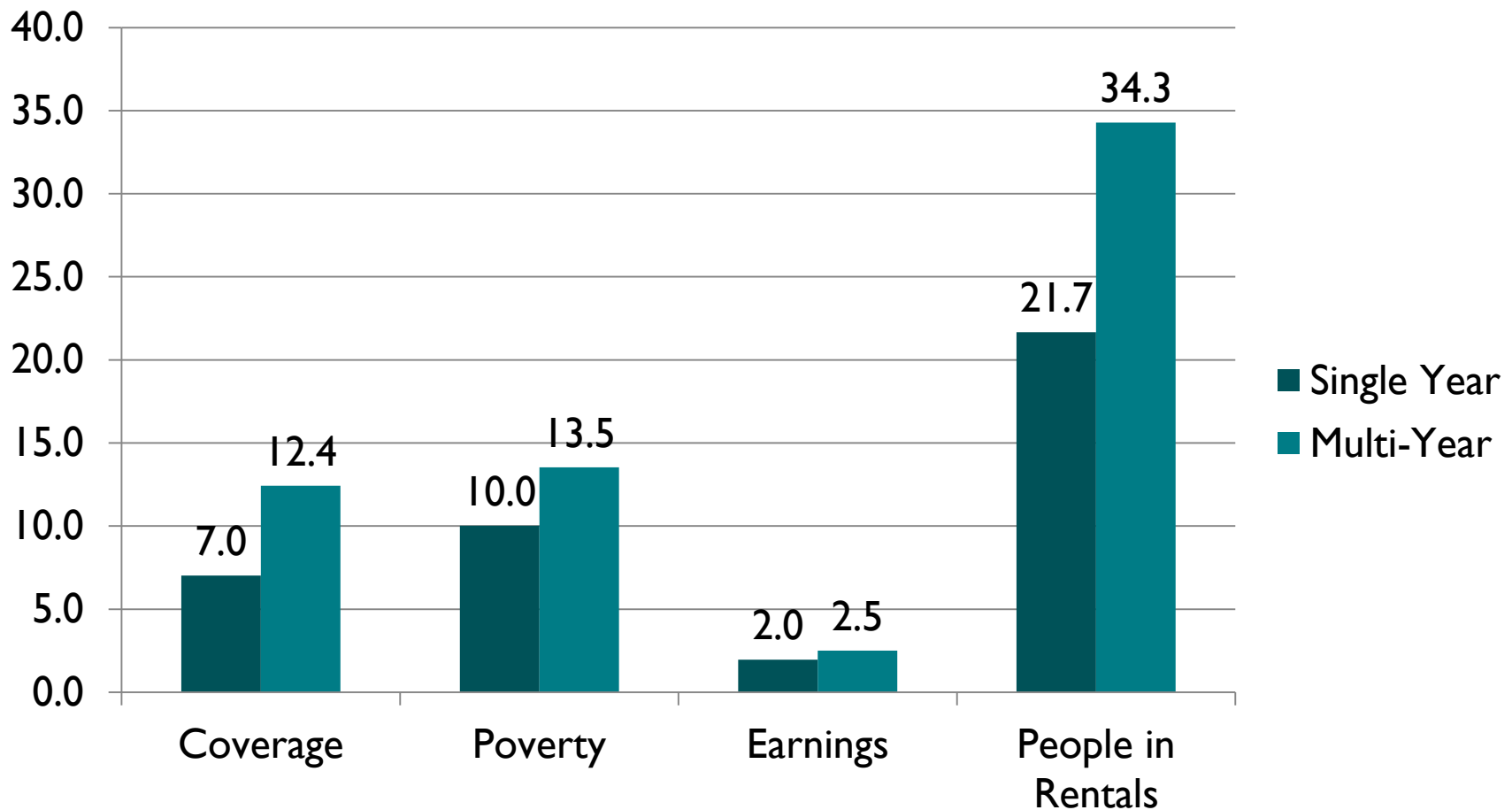
Design Effects by Mode/GQ

	MSE (mean)	% Outliers	# Outliers	DEFF of Insurance
HU Mail	2005	.34	6,814	4.5
HU CATI/CAPI	7110	14.9	142,427	1.9
GQ	3713	2.7	2,256	2.0

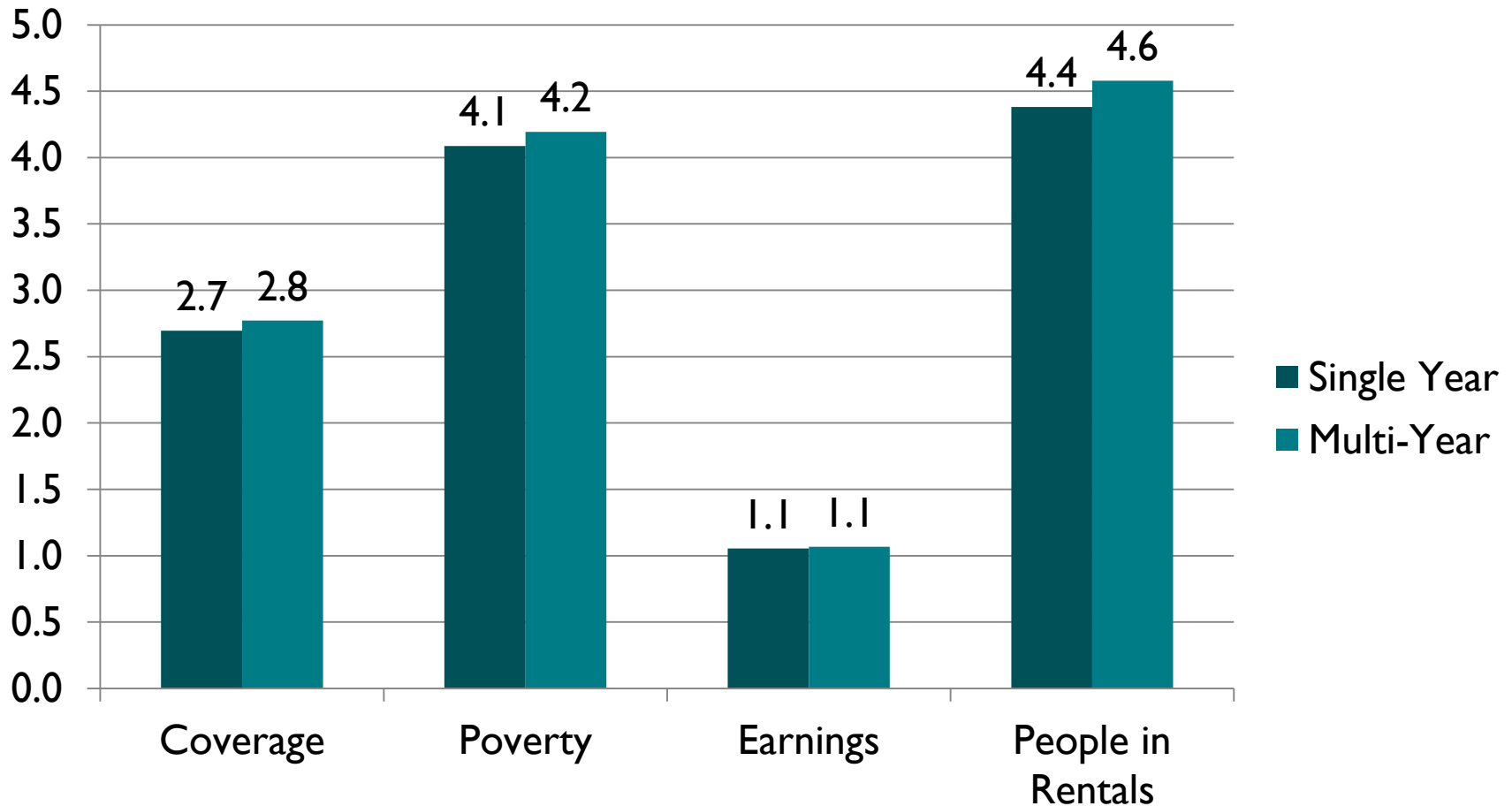
Practical Significance

- ACS is designed for local area estimation
- National standard errors already small
 - National S.E. for health insurance: 0.06
 - National S.E. w/o outliers: 0.03
 - National S.E. assuming state DEFF: 0.01
- Two practical areas of concern
 - Multi-year file: larger sample size
 - Large sub-groups

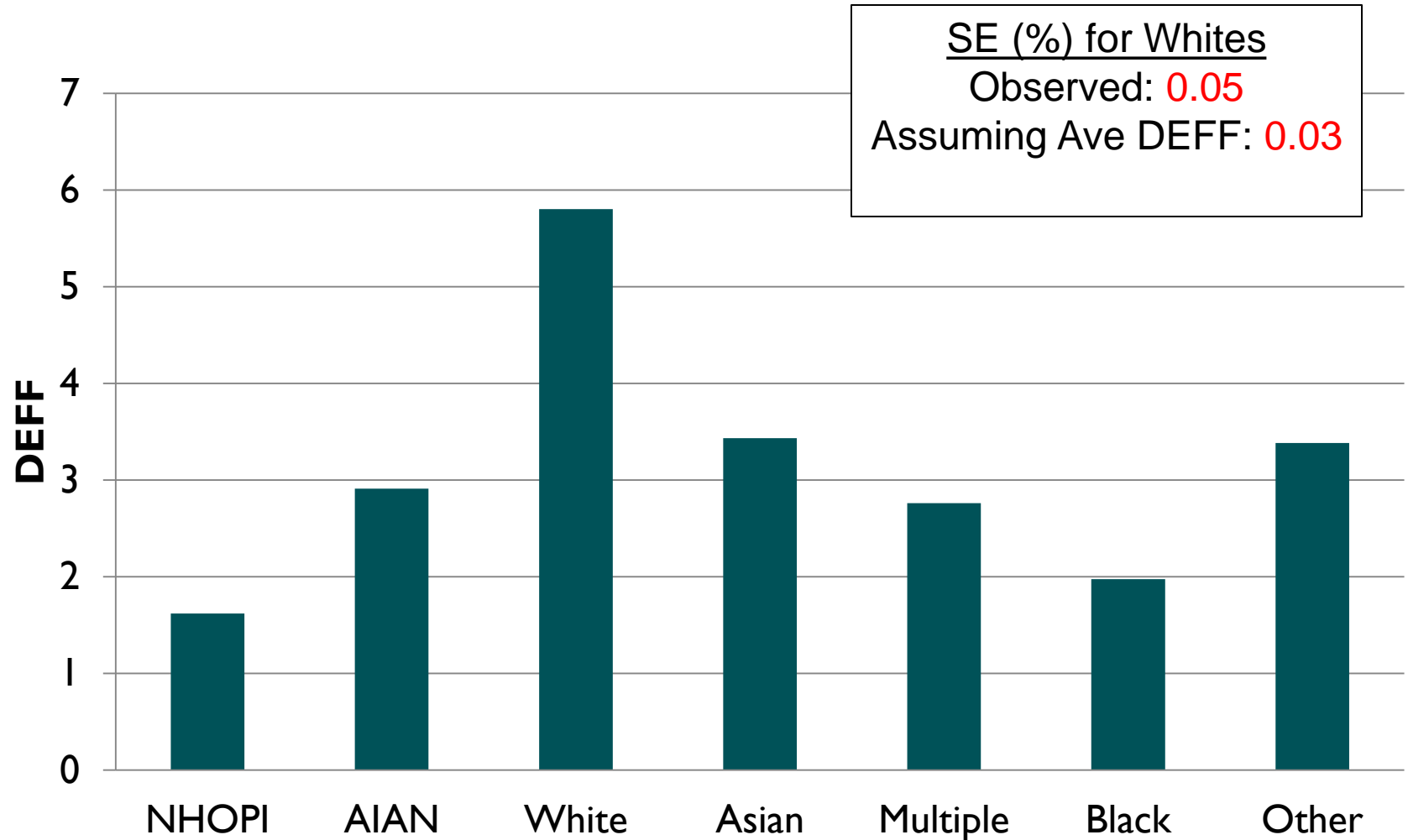
National Design Effects in Single Year vs. Multi-year



Average State Design Effects in Single Year vs. Multi-year



Large Sub-Groups



Summary/Recommendations

- National Design Effects appear to large
 - National SE's upwardly biased
- Potentially driven by CAPI Sub-sampling
 - Only apparent at aggregated domains
- Further investigation into RF by SSF interaction
 - Accurate reflection of sample design or fixable in the weights?
- Analysts that wish to avoid this can adopt alternative variance method (TS)

Acknowledgements

- **Funding**
 - RWJF grant to SHADAC
 - Interdisciplinary Doctoral Fellowship Program (Univ. of MN)
- **Collaborators**
 - Peter Graven
 - Michael Davern
 - Kathleen Call

Michel Boudreaux

boudr019@umn.edu



Sign up to receive our newsletter and updates at

www.shadac.org



UNIVERSITY OF MINNESOTA

School of Public Health