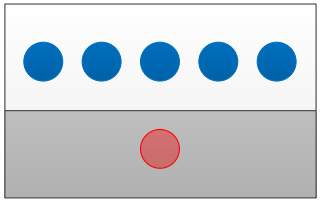
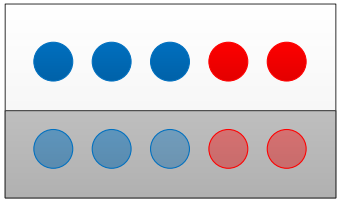
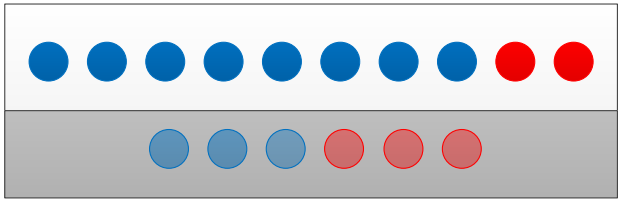
Developing and Evaluating a New Metric for ABS Frame Quality Assessment

Stephanie Zimmer and Ashley Amaya

Background

- CDS used as foundation for sampling frame for an address-based sample (ABS)
- ABS frame has high coverage nationally, but some areas have undercoverage
- Many surveys use frame enhancement in segments with low estimated net coverage
- We argue net coverage may not be best metric to identify segments to enhance and have developed a metric to indicate potential bias in estimates

Within Segment Coverage Error

Segment 1		Segment 2		Total	
					
N_{frame}:	5	N_{frame}:	5	N_{frame}:	10
N_{pop}:	6	N_{pop}:	10	N_{pop}:	16
Coverage Rate:	83.3%	Coverage Rate:	50.0%	Coverage Rate:	62.5%
Risk of Bias:	100%	Risk of Bias:	0.0%		
$Red_{NoEnhance}$:	0.0%	$Red_{NoEnhance}$:	40.0%	$Red_{NoEnhance}$:	20.0%
				$Red_{NoEnhanceWT}$:	25.0%
$Red_{Current}$:	0.0%	$Red_{Current}$:	40.0%	$Red_{Current}$:	26.7%
				$Red_{CurrentWT}$:	25.0%
Red_{New}:	16.7%	Red_{New}:	40.0%	Red_{New}:	27.3%
				Red_{NewWT}:	31.25%
Red_{True}:	16.7%	Red_{True}:	40.0%	Red_{True}:	31.25%

Outline of method

1. Rank segments by risk of bias due to coverage

2. Use enhance listing in segments with highest risk of bias – this addresses within segment coverage error

3. Add weighting step to address undercoverage at segment level

4. After data collection, post-stratify the coverage adjusted weights

Coverage Bias Risk Index (CBRI): a metric to identify segments at high risk of coverage error

1. For each segment, estimate coverage rate for a subgroup (e.g. those with less than HS education or renters) using a model

2. Find estimate of proportion of the population belonging to each subgroup from reliable source such as ACS

3. Given the coverage rate and population distribution, estimate the percent of each subgroup that will be covered

4. Calculate difference between estimate in step 2 and step 3

5. These differences are combined using an average of absolute values to create the CBRI

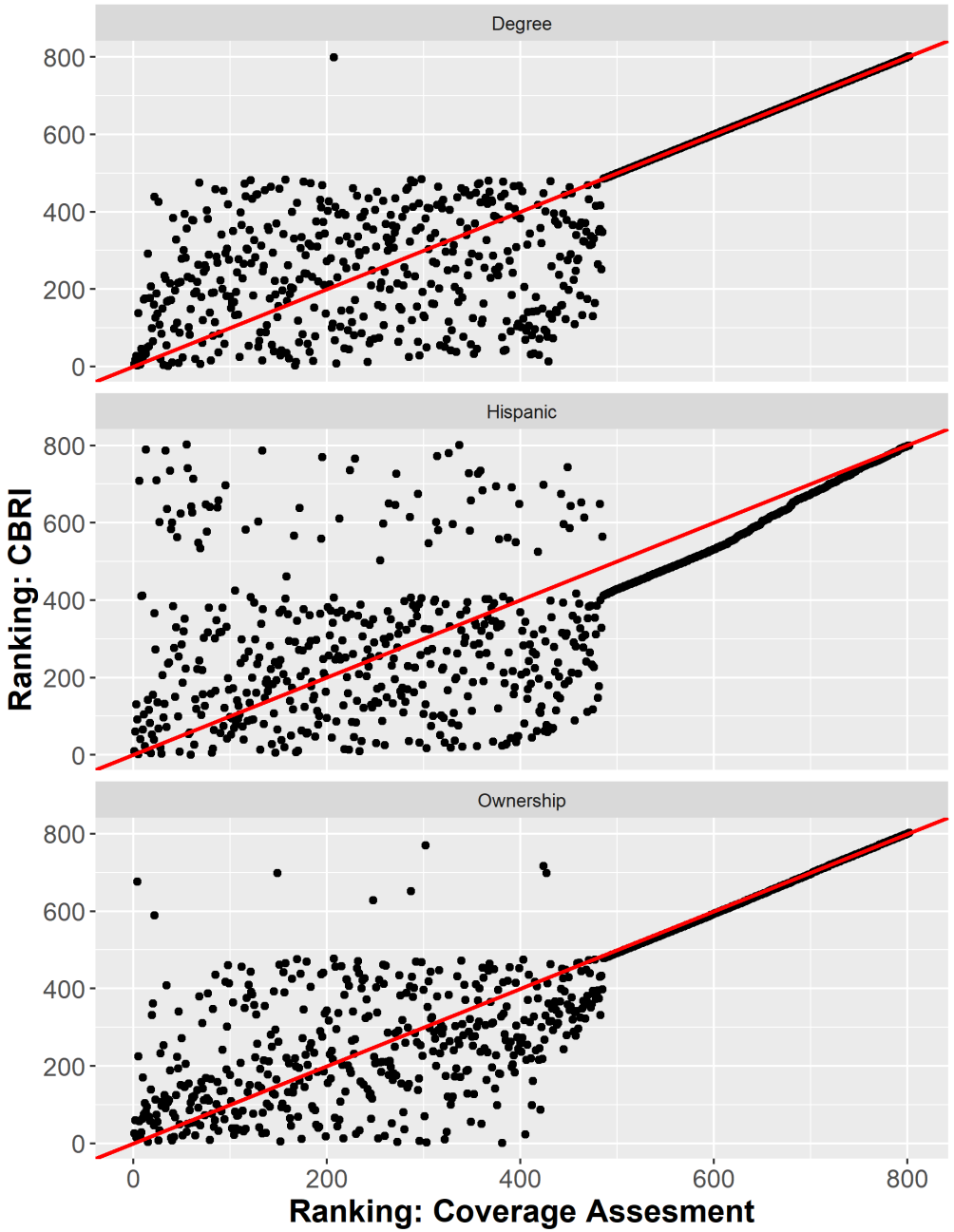
Adjust weights for coverage

- Design weight is defined as the inverse of probability of selection from the sampling frame used

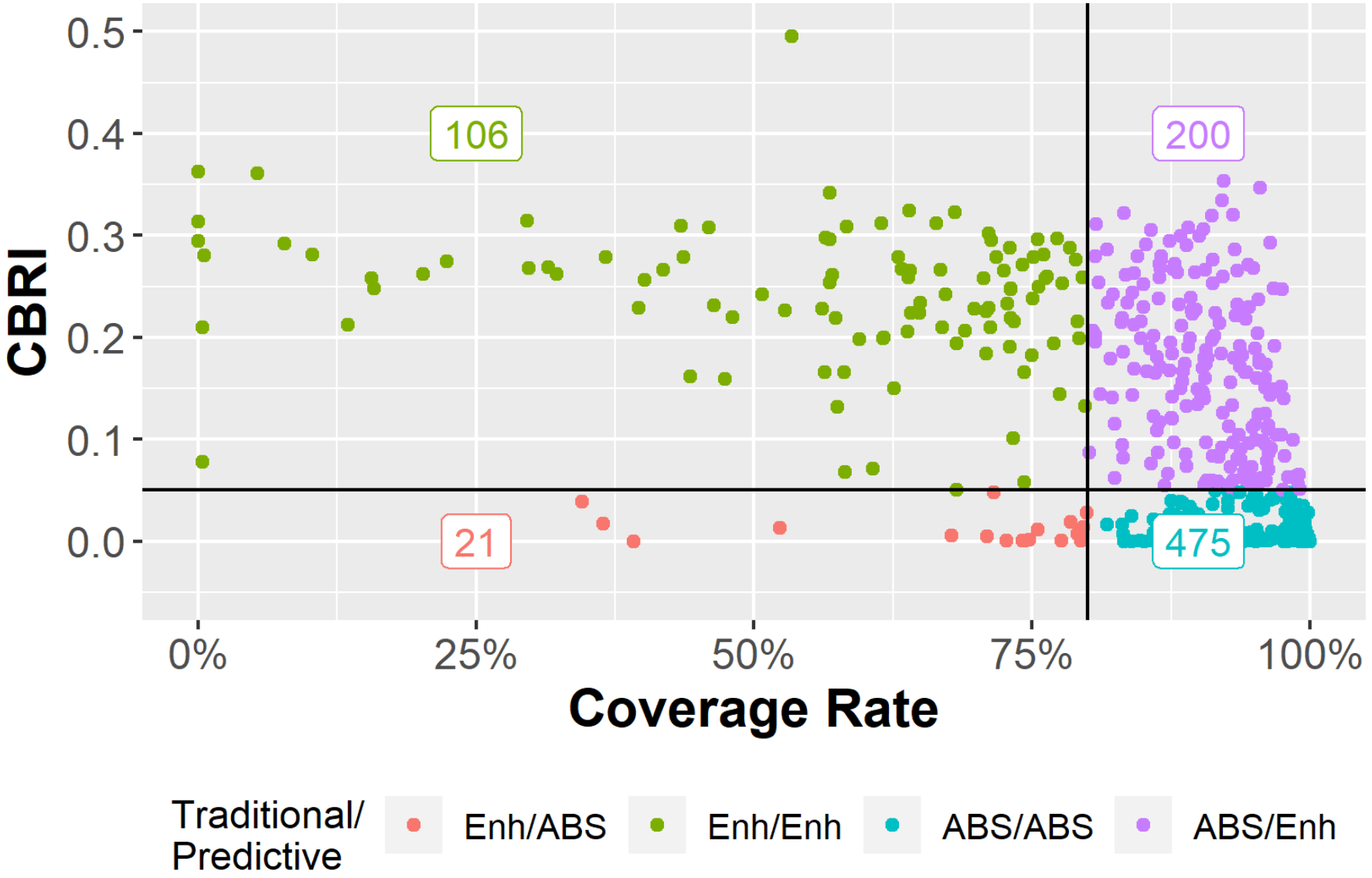
- Then adjust this weight for coverage as follows:

$$w_{covadj,i} = w_{design,i} * \frac{N_{pop}}{N_{frame}}$$

Comparing ranking of coverage rates to CBRI



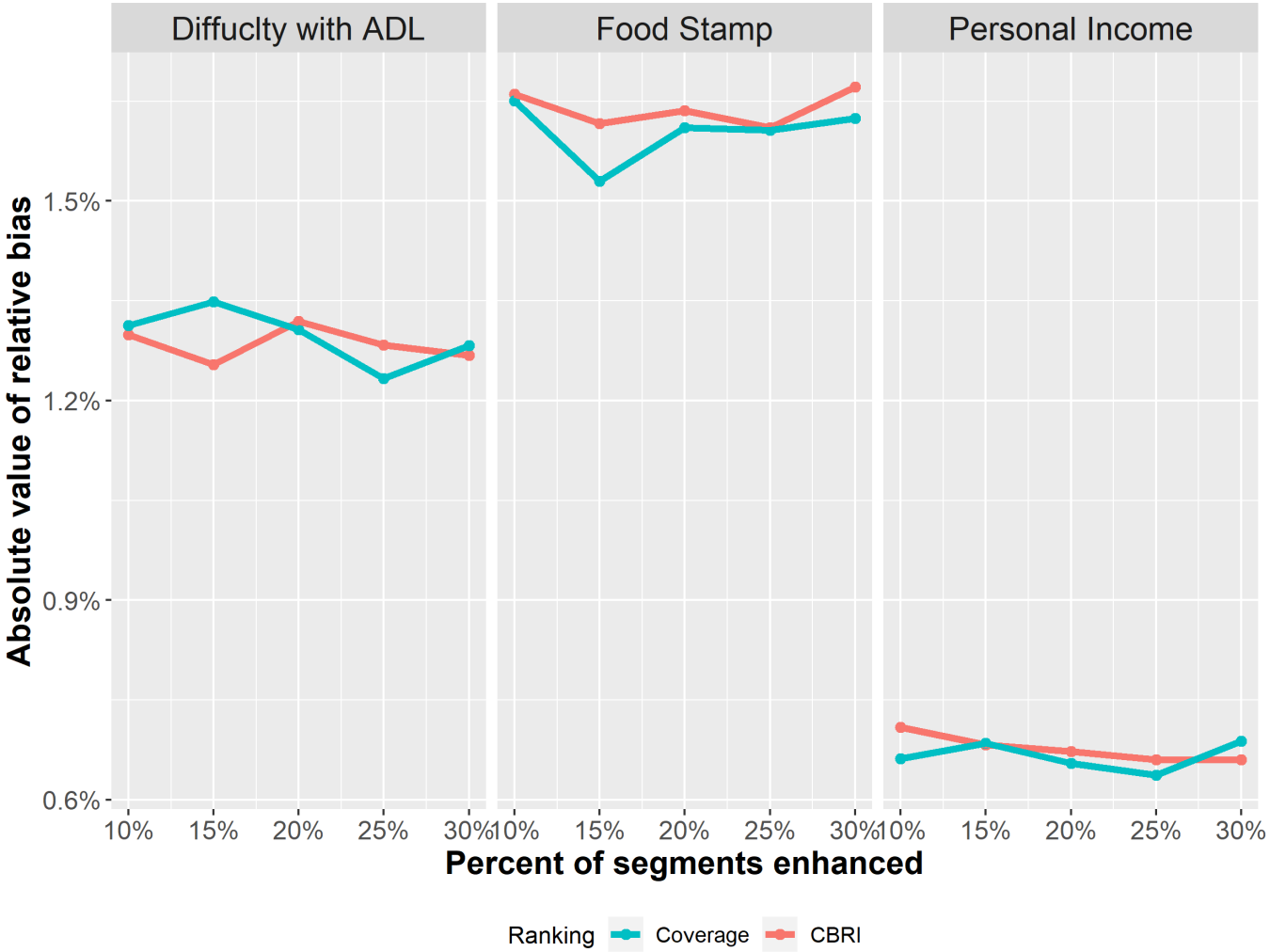
Comparing coverage rates to CBRI for decision making



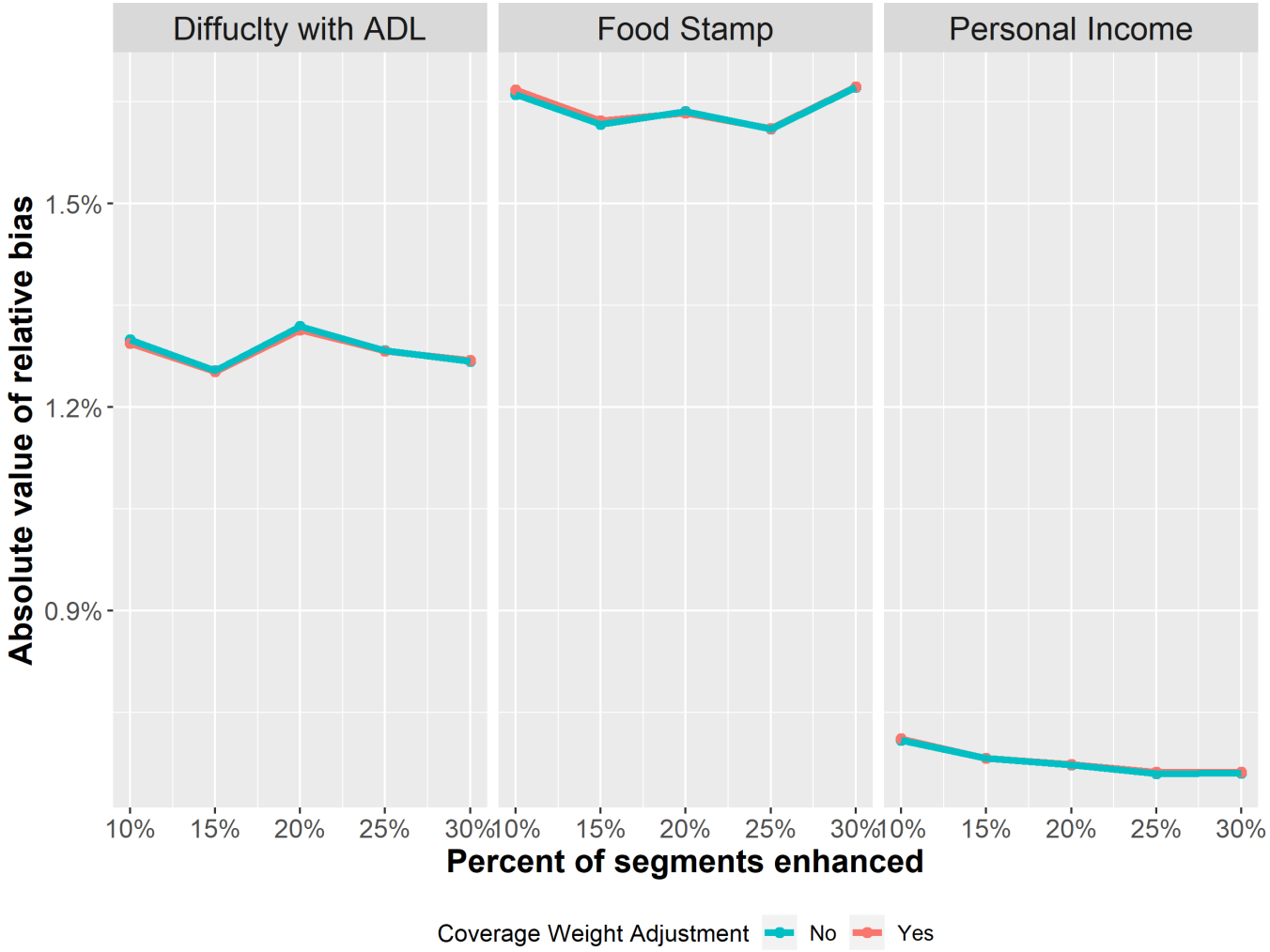
Simulation

- Generated a population based on ACS and assigned coverage propensity for each housing unit based on model
- Ranked segments according to CBRI and coverage
- Segments with low coverage were “enhanced” and simulation considered them having 100% coverage, segments with high CBRI were “enhanced” and simulation considered them having 100% coverage
- Sampled from frames with varying coverage levels – varied whether CBRI or coverage used for decision making, whether weight was adjusted for coverage, and whether post-stratification was used.
- 12 outcomes were simulated and bias estimated

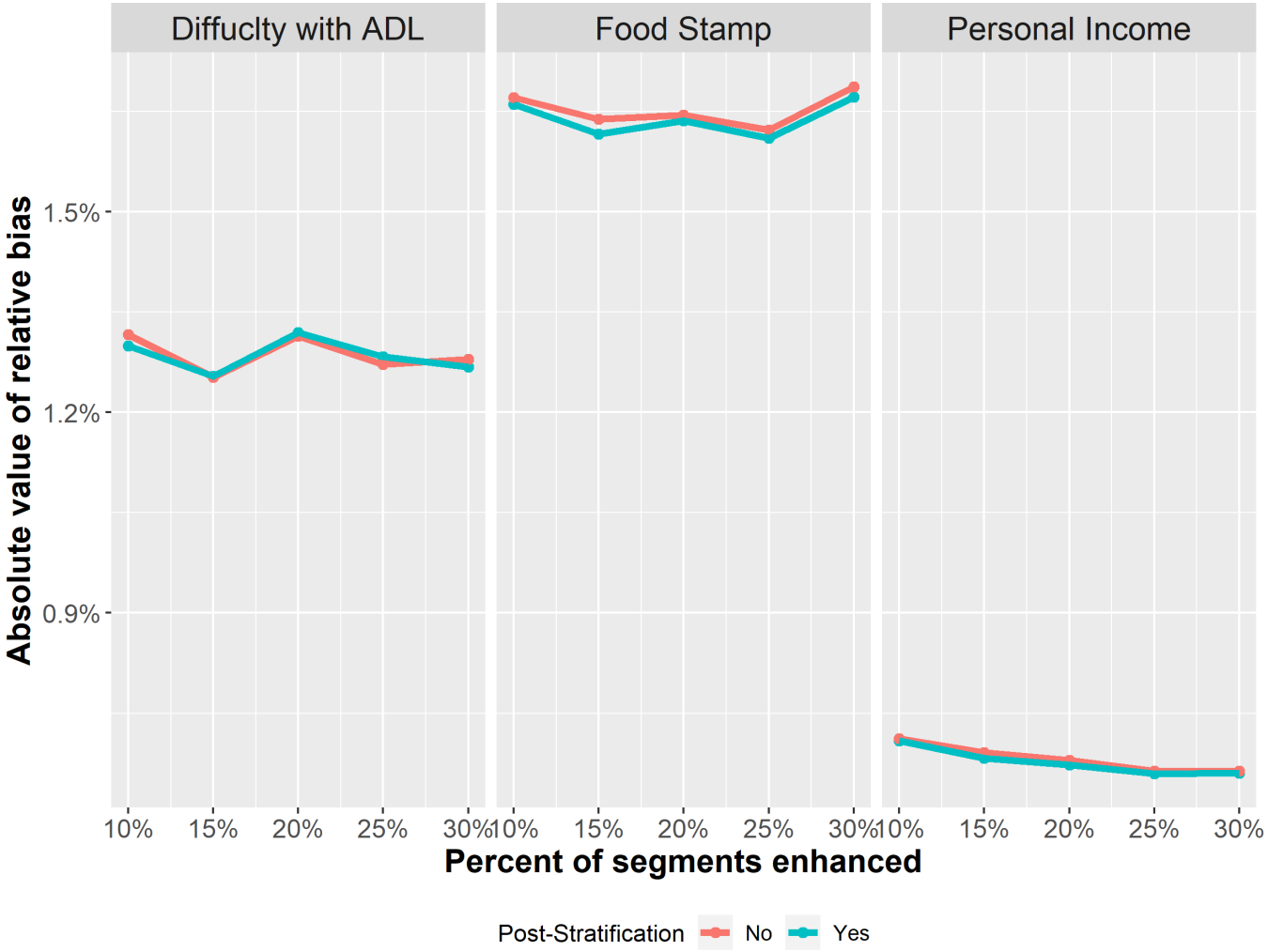
Simulation Results: Ranking Segments



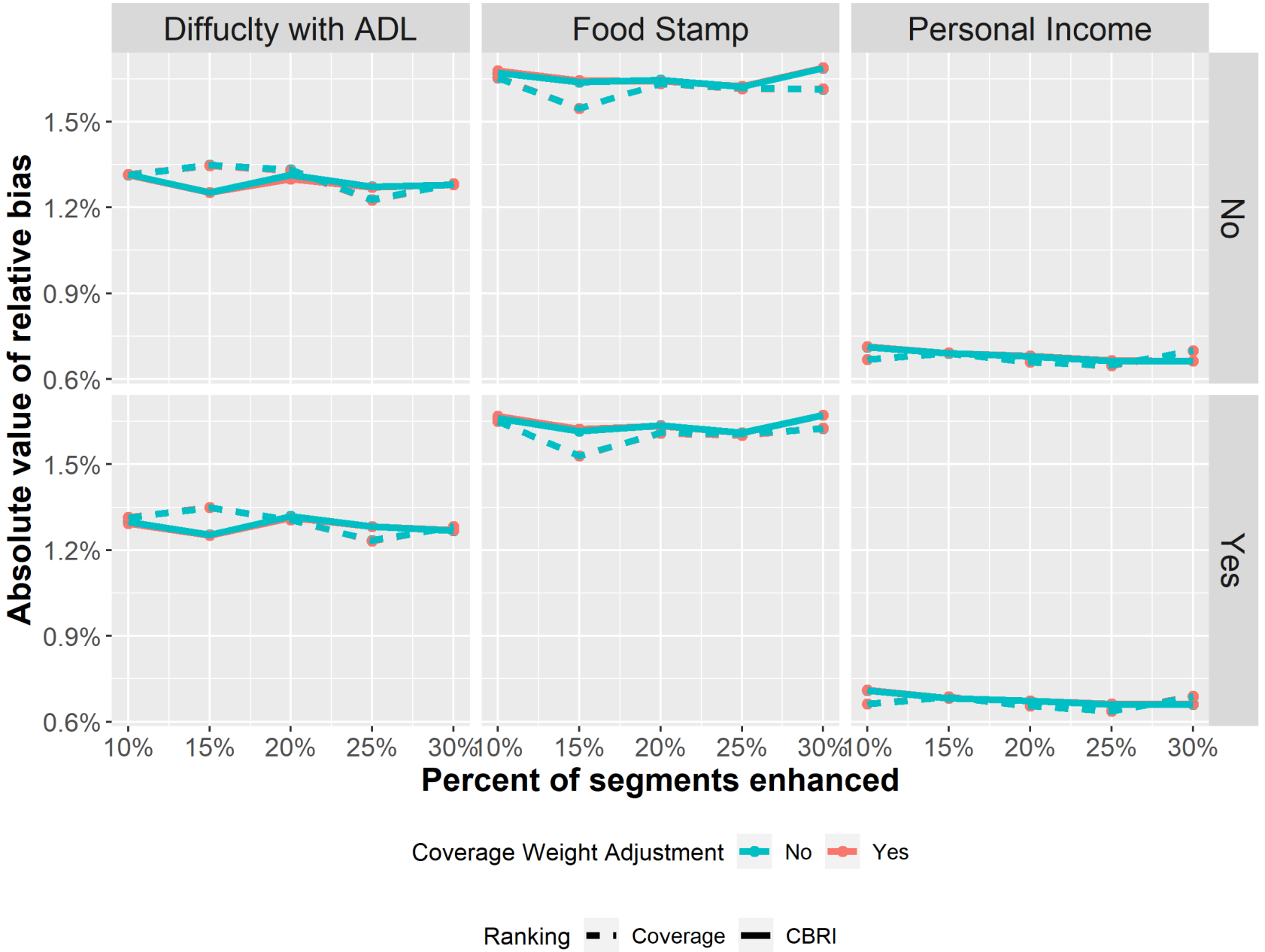
Simulation Results: Modified Coverage Weight



Simulation Results: Post-Stratification



Simulation Results



Simulation results

- No significant difference found by using CBRI compared to using coverage rate
- No significant difference from using coverage weighting adjustment
- Post-stratification reduced bias for some variables, namely citizenship, birth place, and insurance status. Post-stratification variables included Census division, race/ethnicity, and sex only

Conclusion

- This new ranking method does not out-perform standard method of using net coverage rate for decision making
- Better models may improve the method – needs ground truth data which is expensive
- Weighting for undercoverage does not impact bias of estimates

More Information

Stephanie Zimmer
Research Statistician
sazimmer@rti.org