

Switching from Field Enumeration to an ABS Frame: The Effect on Coverage Bias

Ashley Amaya, Matthew Williams, Devon Cribb, Rachel Harter, &
Katherine Morton

August 2, 2018



Acknowledgements

- The presentation is sponsored by RTI International's Survey Research Division, with the research included in the presentation stemming from ongoing methodological work conducted under the contract for the **National Survey on Drug Use and Health (NSDUH)**. The NSDUH is funded by the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality under contract no. 283-2017-00002C and project no. 0215638.
- The views expressed in this presentation do not necessarily reflect the official position or policies of SAMHSA or the U.S. Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

Introduction to ABS

- What is address-based sampling?
 - Address list based on the US Postal Service's database
 - Geocode addresses into sample segments
 - Draw a sample
- Why do we use it?
 - Great coverage (over 90%) (AAPOR 2016)
 - Less expensive
 - Eliminates human listing error
 - Timely

Undercoverage on the ABS Frame

- The ABS frame is not perfect
 - Undercoverage is much higher in rural areas
 - 23-35% in rural areas vs. 1-10% in urban areas (Dohrmann et al 2006; Dohrmann et al 2007; O’Muircheartaigh et al 2007)
 - Purposely excludes:
 - Unique ZIP codes (e.g., AIAN tribal areas and universities)
 - Vacant units in rural areas
 - Includes “unusable” addresses:
 - PO Boxes
 - Simplified addresses (e.g., rural routes)

Research Question

- Does the use of an ABS frame introduce coverage bias?
 - If so, how much?

Methods - Datasets

- The NSDUH provides national, state and substate data on substance use and mental health in the civilian, noninstitutionalized population age 12 and older.
- Data are collected on a quarterly basis each year.
- Approximately 700 field interviewers (FIs) staffed.
- Approximately 140,000 household screenings and 67,500 interviews completed annually.
- Conducted by RTI under contract with SAMHSA.
- Currently uses a field enumerated frame.

Methods – Datasets (cont'd)

3 datasets constructed from 2015-2016 NSDUH

– Field enumerated (FE) dataset

- All respondents
- N = 136,000

– ABS Subsample 1

- FE dataset minus residents of description-based addresses
- N = 129,000

– ABS Subsample 2

- Subsample 1 minus residents of tribal areas and group quarters
- N = 125,000

Methods – Variables of Interest

▪ Alcohol

- Past month binge alcohol use
- Past month alcohol use
- Past year alcohol use disorder

▪ Other drugs

- Past month stimulant use
- Past year substance use disorder
- Past month pain reliever use
- Past year illicit drug use disorder
- Past year specialty substance use treatment

- Past month **cigarette** use
- Past month **marijuana** use

▪ Mental health

- Past year serious mental illness
- Past year any mental illness
- Past year mental health service use
- Past year major depressive episode
- Past year major depressive episode

Methods – Domains of Interest

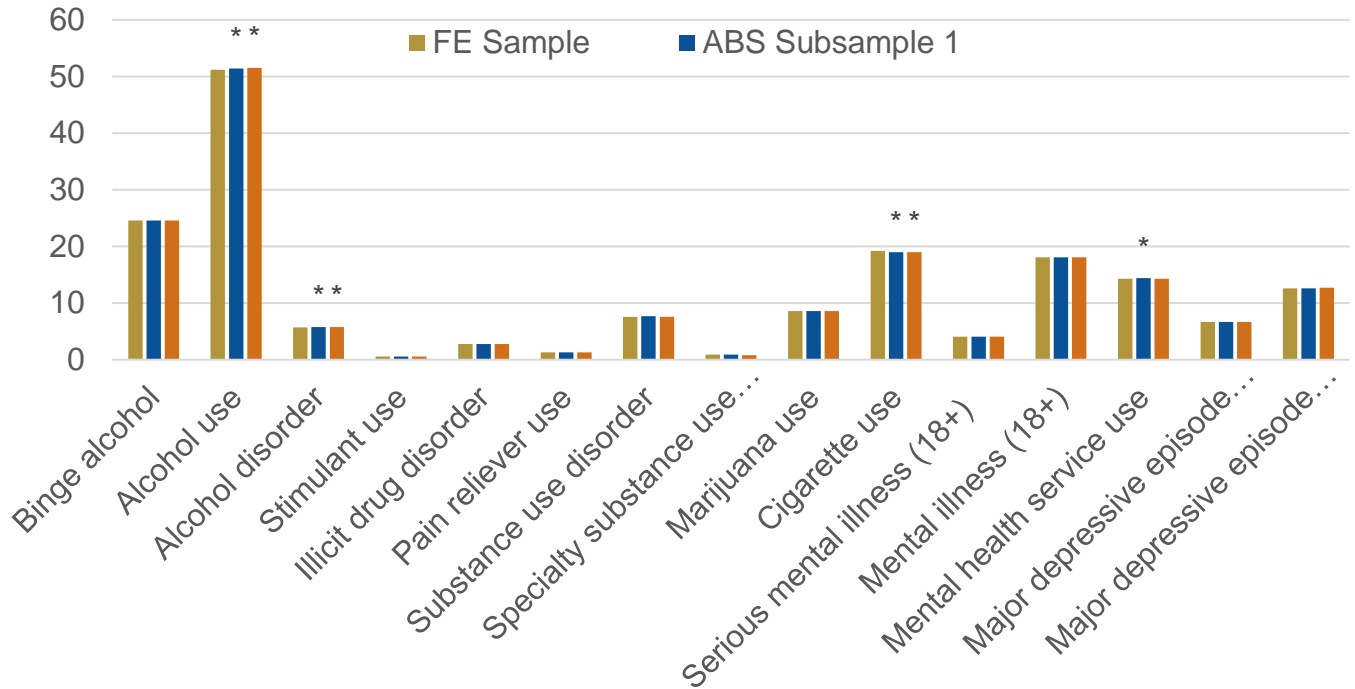
- Age
- Sex
- Race
- Hispanicity
- Census division
- County type
- College education status
- Pregnancy status
- 13 two-way cross domains
- Up to 325 comparisons for each outcome variable

Methods – Overview of Analyses

1. **Overall** differences by sample
2. Summary of differences by **measure**
3. Summary of differences by **sample size**
4. Summary of differences in **conclusions** drawn

1. Overall differences by sample

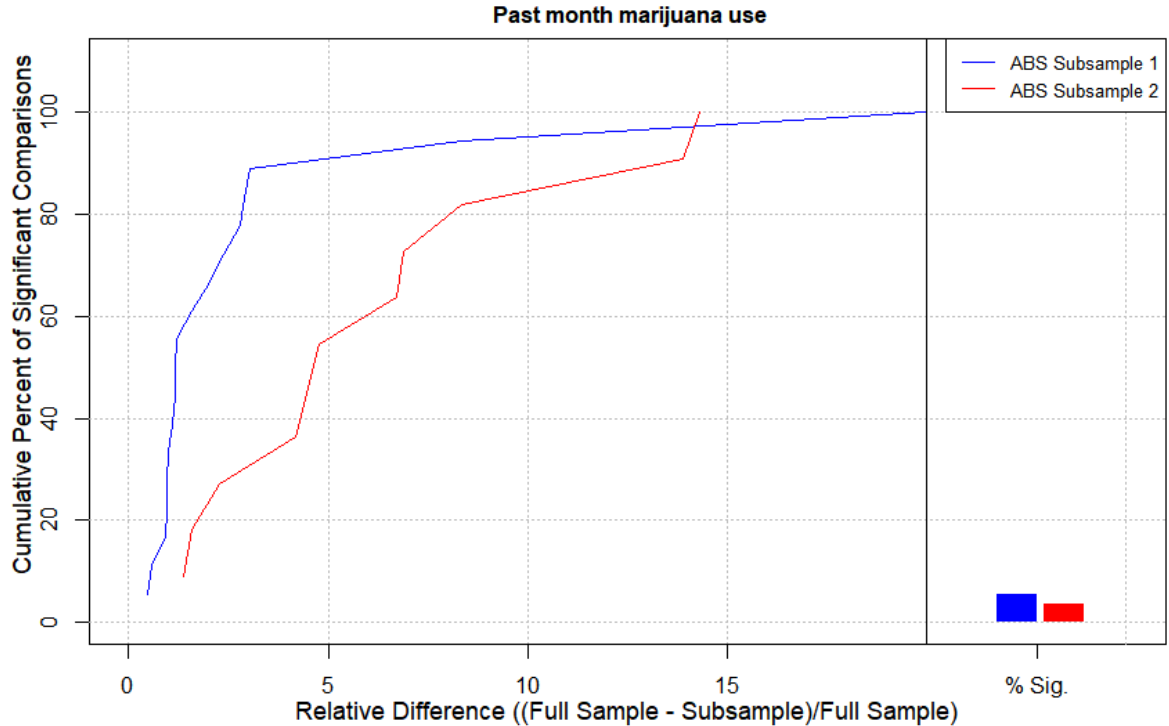
1. Overall Differences by Sample



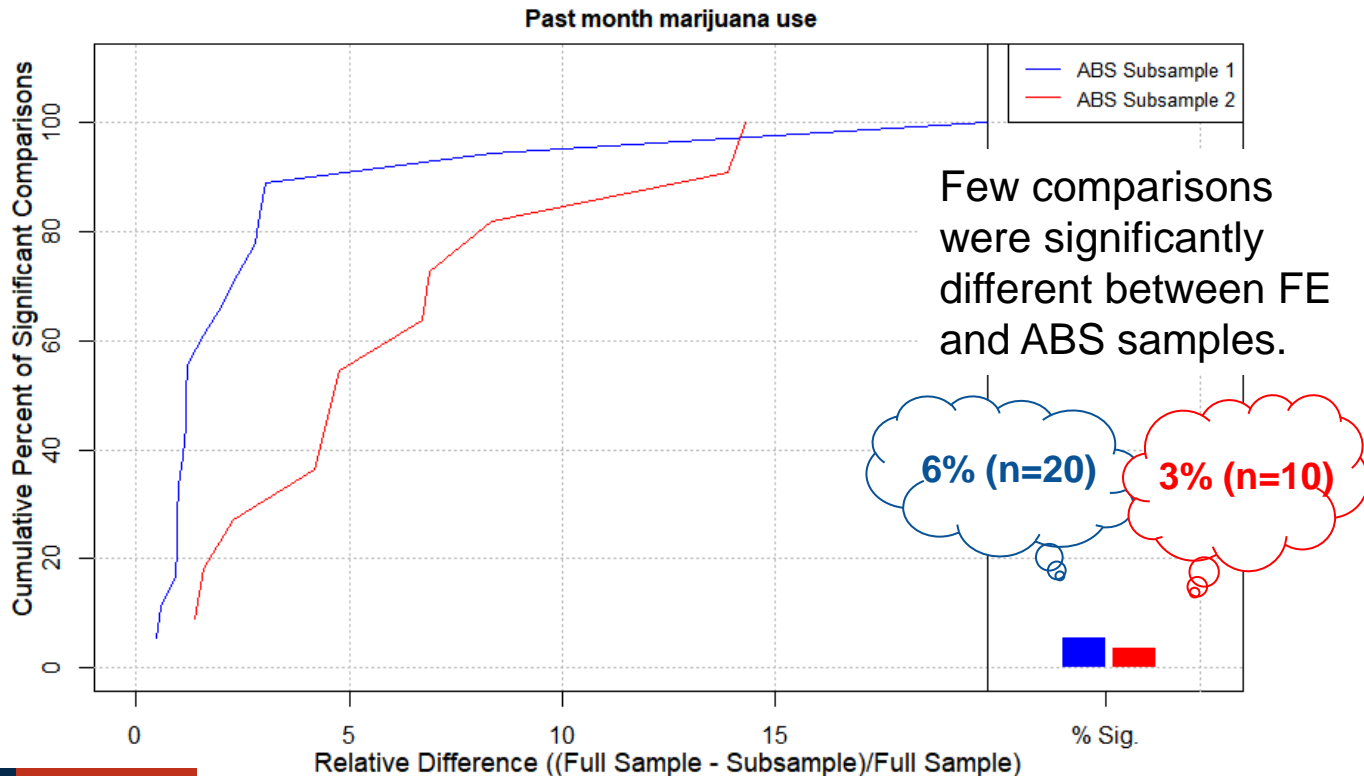
* $p < 0.05$

2. Summary of significant subdomain comparisons by measure

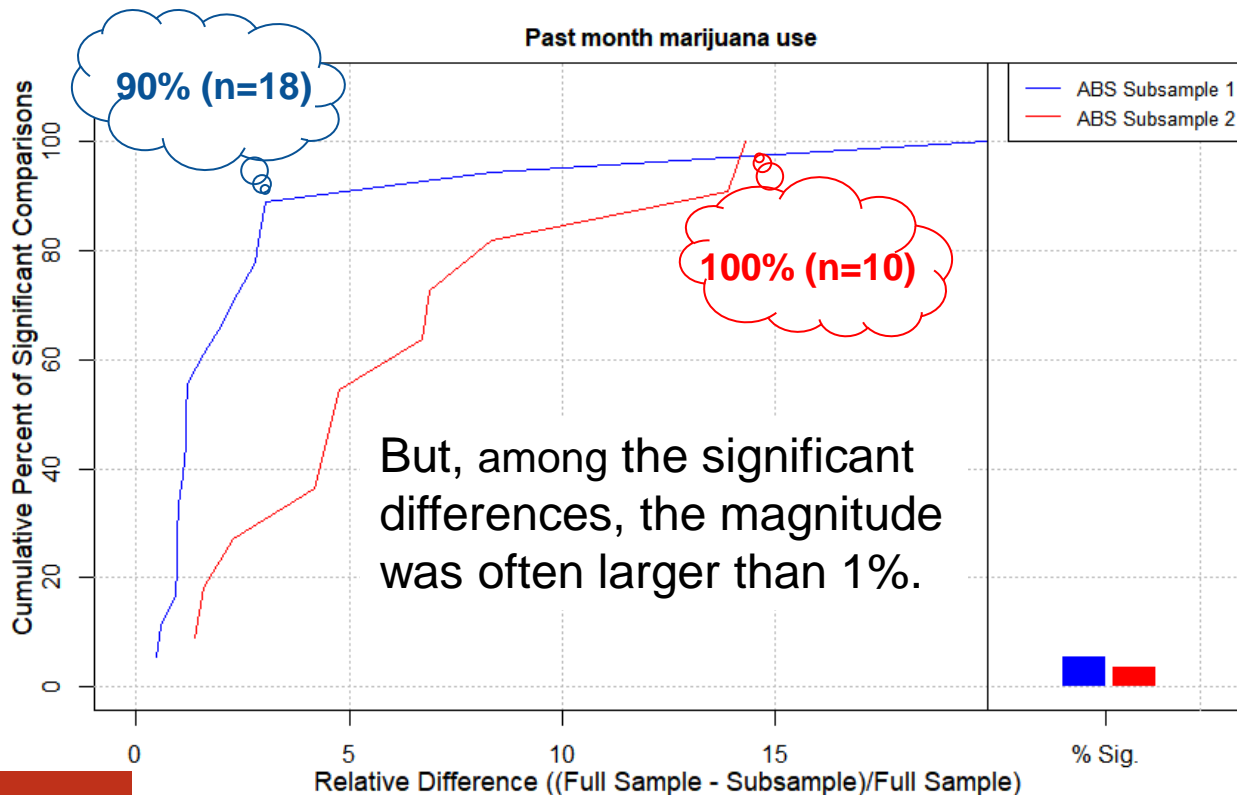
2. Summary Relative Difference in Estimates: Marijuana Use



2. Summary Relative Difference in Estimates: Marijuana Use



2. Summary Relative Difference in Estimates: Marijuana Use



2. Summary of Findings by Measure

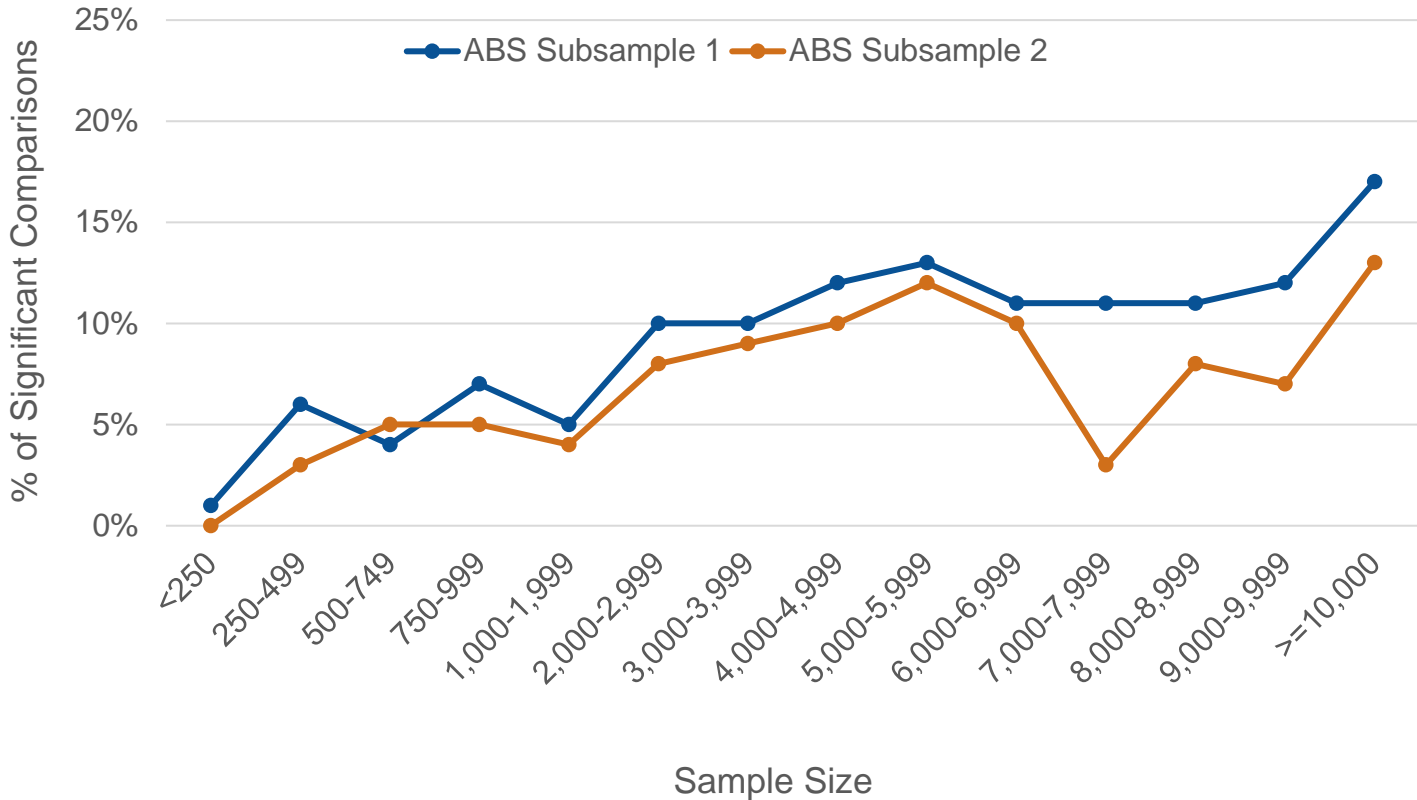
		ABS Subsample 2			
		Unaffected	Bias on few domains, but large	Bias on many domains but small	Bias on many domains and large
ABS Subsample 1	Unaffected	<ul style="list-style-type: none"> • Binge alcohol • Stimulant use • Serious mental illness (18+) 			
	Bias on few domains but large	<ul style="list-style-type: none"> • Substance use disorder • Specialty substance use treatment • Marijuana use 	<ul style="list-style-type: none"> • Past month pain reliever use 		<ul style="list-style-type: none"> • Major depressive episode (18+)
	Bias on many domains but small		<ul style="list-style-type: none"> • Mental health service use (18+) 	<ul style="list-style-type: none"> • Alcohol use • Cigarette use • Any mental illness (18+) 	
	Bias on many domains and large		<ul style="list-style-type: none"> • Alcohol use disorder 		<ul style="list-style-type: none"> • Illicit drug use disorder

2. Summary of Findings by Measure

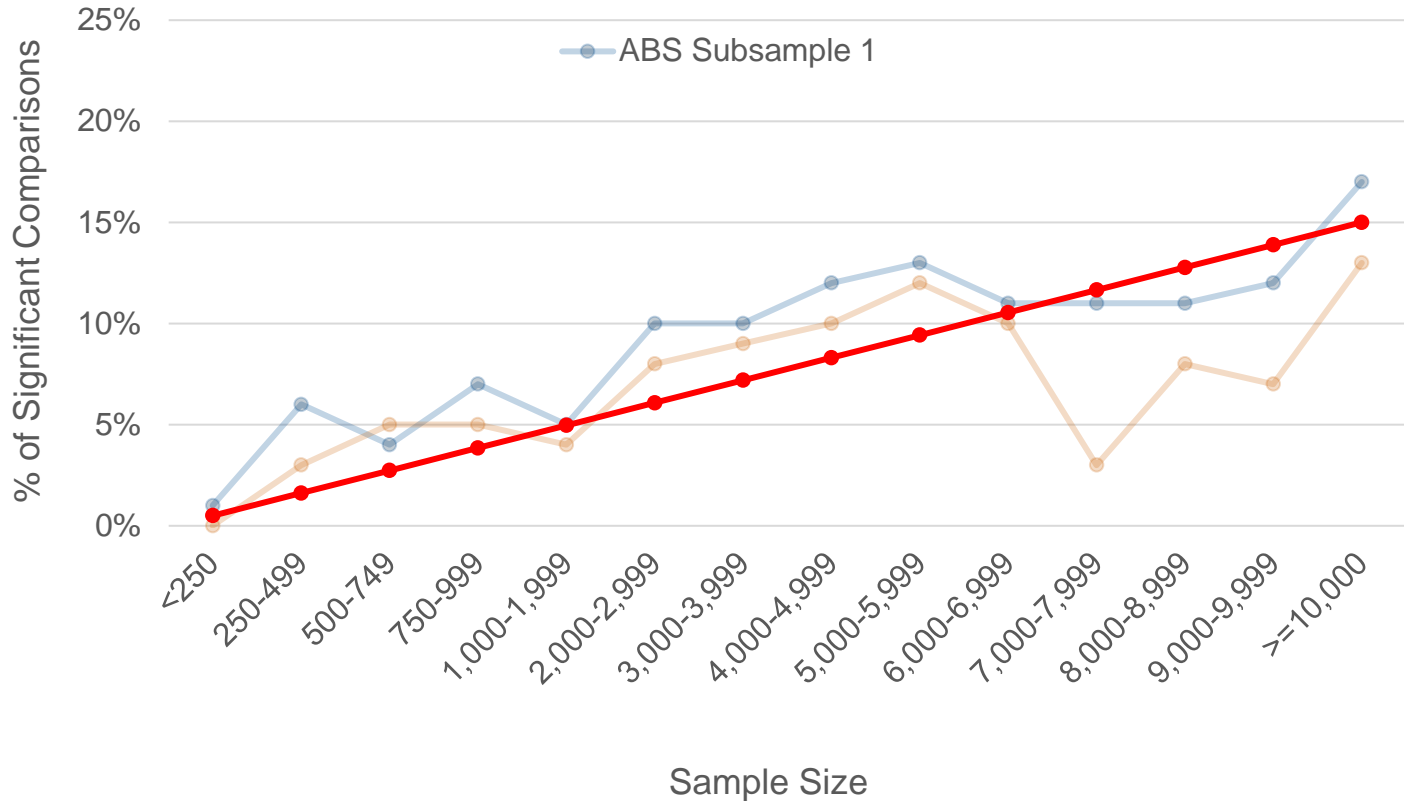
		ABS Subsample 2			
		Unaffected	Bias on few domains, but large	Bias on many domains but small	Bias on many domains and large
ABS Subsample 1	Unaffected	<ul style="list-style-type: none"> • Binge alcohol • Stimulant use • Serious mental illness (18+) 			
	Bias on few domains but large	<ul style="list-style-type: none"> • Substance use disorder • Specialty substance use treatment • Marijuana use 	<ul style="list-style-type: none"> • Past month pain reliever use 		<ul style="list-style-type: none"> • Major depressive episode (18+)
	Bias on many domains but small		<ul style="list-style-type: none"> • Mental health service use (18+) 	<ul style="list-style-type: none"> • Alcohol use • Cigarette use • Any mental illness (18+) 	
	Bias on many domains and large		<ul style="list-style-type: none"> • Alcohol use disorder 		<ul style="list-style-type: none"> • Illicit drug use disorder

3. Summary of significant subdomain comparisons by sample size

3. Summary of Differences by Sample Size

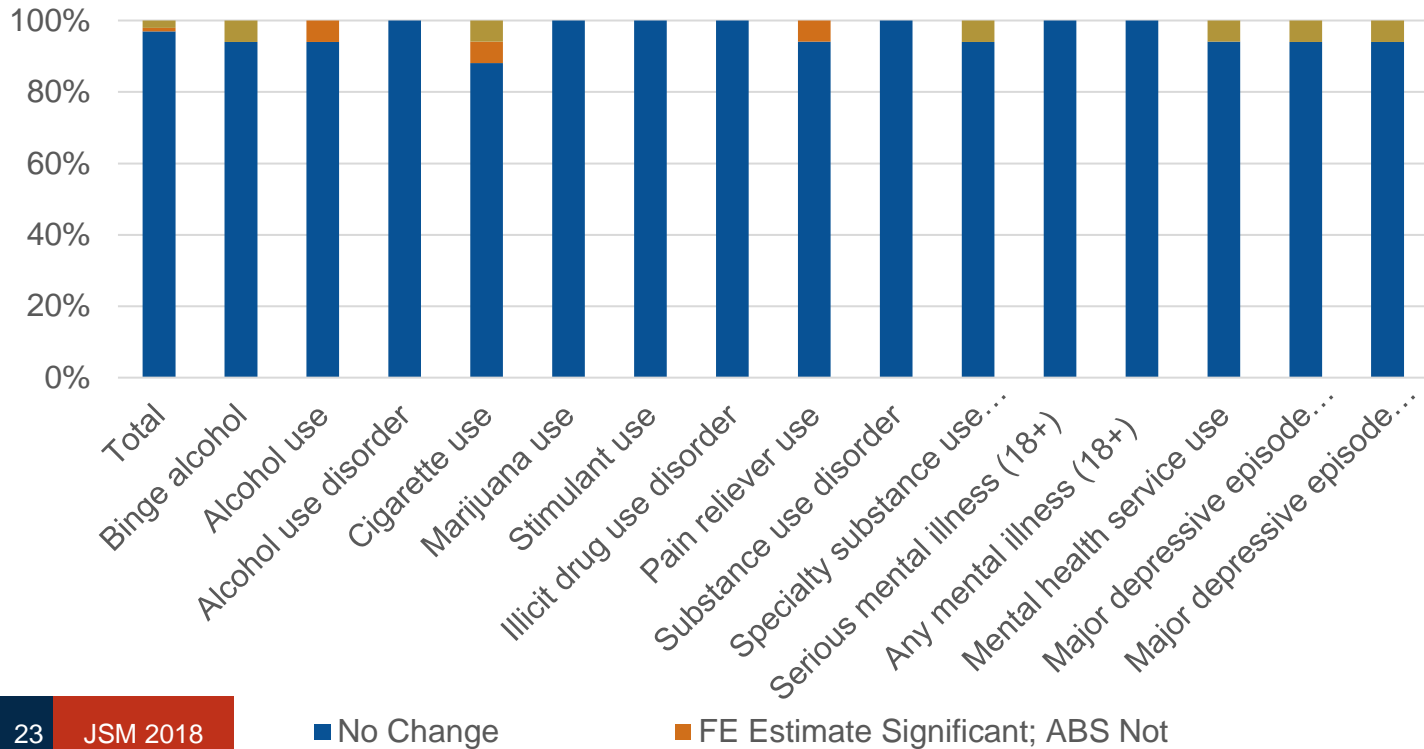


3. Summary of Differences by Sample Size



4. Comparisons between subdomain and overall measures

4. Proportion of Comparisons that would Change Significance between Subsample 1 and FE Sample



Summary

Does an ABS Frame Introduce Coverage Bias?

- **Overall** differences by sample
 - 7 of 30 comparisons were significant, but the differences were small.
- Summary of differences by **measure**
 - Variables differed on the frequency and size of change across frames.
 - No pattern or consistency across frames
- Summary of differences by **sample size**
 - Large samples drove many of the significant findings.
- Summary of differences in **conclusions** drawn
 - Only 3% of conclusions changed, possibly due to chance

Final Take-Away

- An ABS frame has the potential to introduce coverage bias, but...
 - It will depend on the variable of interest
 - It will depend on desired precision and sample size
 - The magnitude of the bias will vary
 - Substantive conclusions in bivariate analyses are unlikely to be affected

Limitations

- Identified differences are worst case scenarios.
 - Large sample sizes increase number of significant differences.
 - All differences were attributed to error in the ABS frame.
 - Areas with known coverage problems would be enumerated, in practice.
 - The simulations are imperfect.
- These findings are limited to health indicators.

More Information

Ashley Amaya

Research Survey Methodologist

202.728.2486

aamaya@rti.org

Works Cited

- American Association for Public Opinion Research. (2016). *Address-based sampling* (prepared for AAPOR Council by the Task Force on Address-Based Sampling; R. Harter, Chair). Oakbrook Terrace, IL: Author. Retrieved from <http://www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx>
- Dohrmann, S., Han, D., & Mohadjer, L. (2006). Residential address lists vs. traditional listing: Enumerating households and group quarters. In *Proceedings of the 2006 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Seattle, WA* (pp. 2959-2964). Alexandria, VA: American Statistical Association.
- Dohrmann, S., Han, D., & Mohadjer, L. (2007). Improving coverage of residential address lists in multistage area samples. In *Proceedings of the 2007 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Salt Lake City, UT* (pp. 3219-3126). Alexandria, VA: American Statistical Association.
- O'Muircheartaigh, C., English, N., & Eckman, S. (2007). Predicting the relative quality of alternative sampling frames. In *Proceedings of the 2007 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Salt Lake City, UT* (pp. 3239-3248). Alexandria, VA: American Statistical Association.