Supplementing Address-based Sample with Prepaid Cell Sample to Help Improve Sample Representativeness

ssrs
research. refined.

JONATHAN BEST
JAMES McKINSTRY
ARIFAH HASANBASRI

CHIA.
center
for health
information
and analysis

CHRISTINE LOVERIDGE
HUONG TRIEU

AAPOR
Annual Conference
2022
Chicago, IL

May 13, 2022

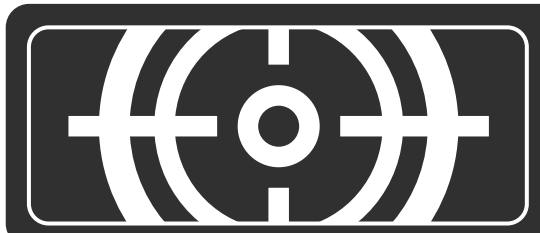
Introduction

- Declining response rates remain a concern for social scientists who still depend on sample surveys.
- The increasing cost of telephone research has made addressed-based sampling more appealing.
- Address-based samples offer their own challenges.
- Often subgroups of interest are less likely to respond to surveys using address-based samples. (Fesksens et al. 2007; Hu, Link and Mokdad 2010; Smith 2014).
- Prepaid cell sample can be used to supplement ABS push to web design.

Agenda

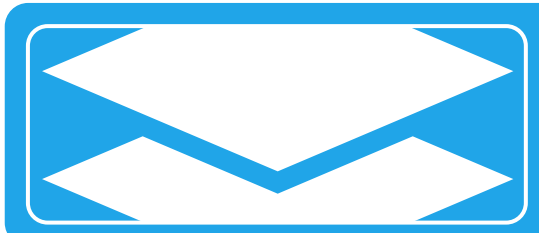
- Show how ABS and prepaid cell samples can be combined in base weighting.
- Investigate optimal sample allocation between the two frames given their demographic characteristics.
- Explore relative costs of ABS and prepaid cell data collection, both overall and by subgroup.

Massachusetts Health Insurance Survey



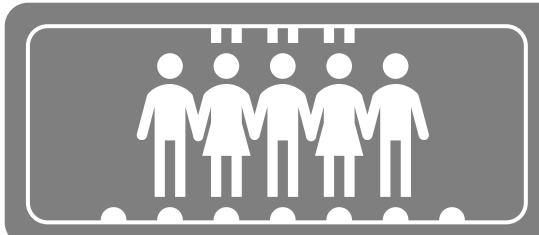
TARGET POPULATION

- Residential population of MA



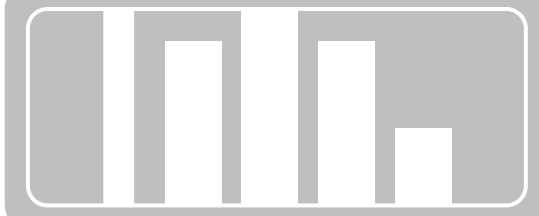
SAMPLE DESIGN

- Stratified ABS design pulled from Computerized Delivery Sequence File
- Simple random sample pulled from list of all prepaid cell numbers in MA



CONTACT PROCEDURES

- ABS: Letter-postcard-letter w/ HC-postcard; Allow people to call in; Calls made to records with matched phone numbers
- Prepaid cell: 6(+) calls

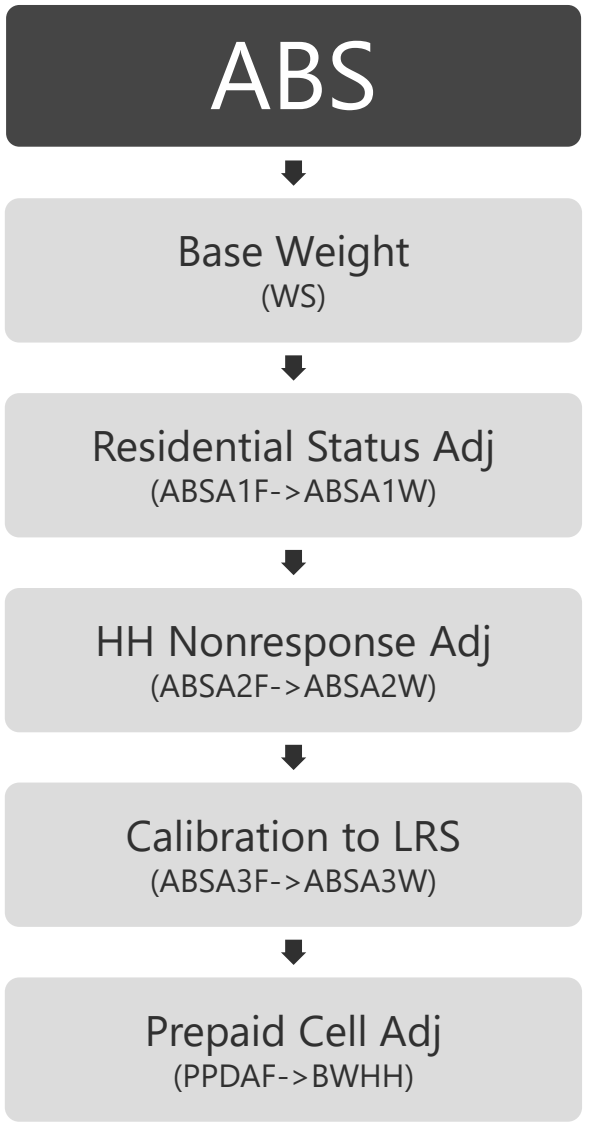


FINAL SAMPLE SIZES AND RESPONSE RATES

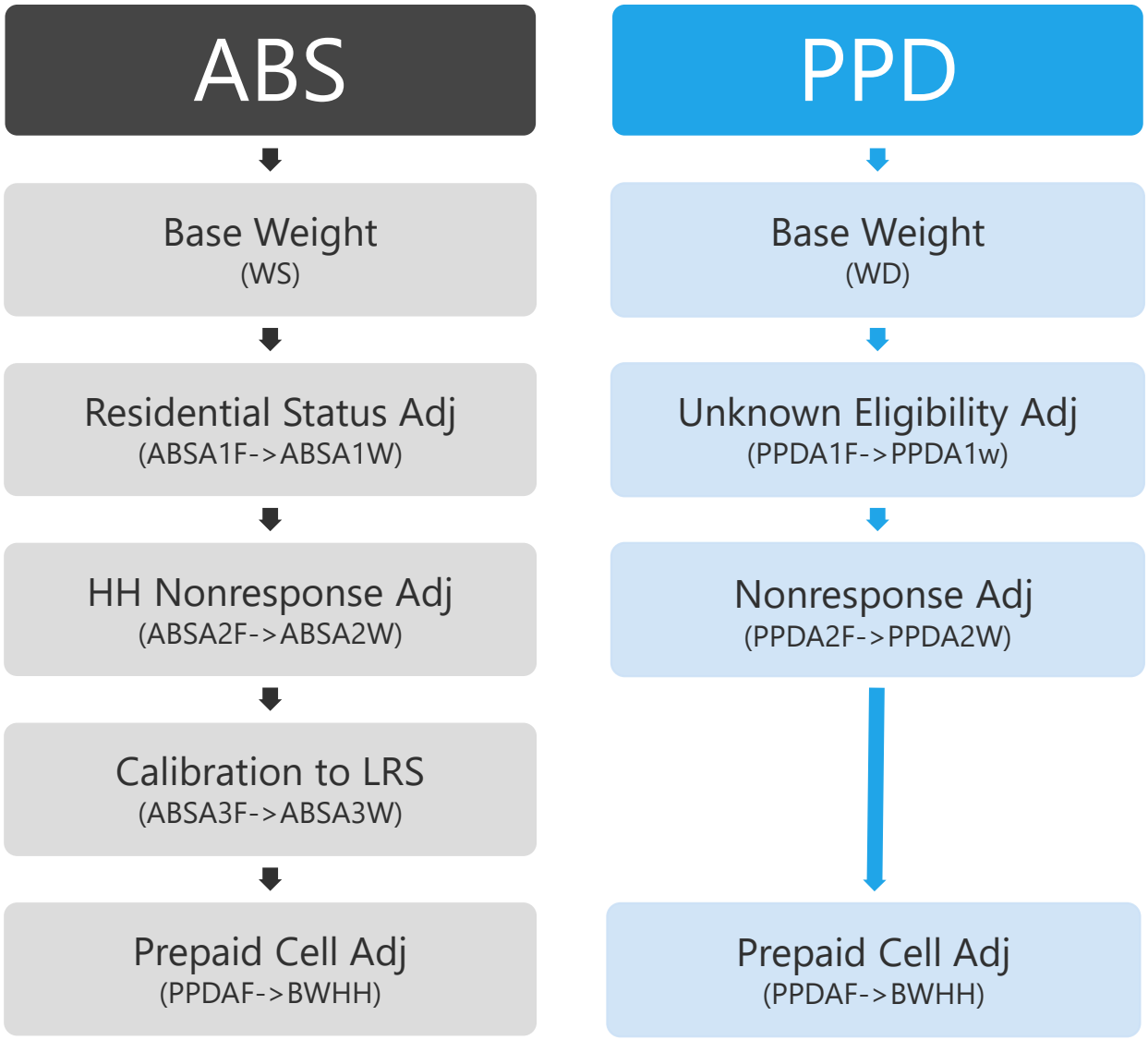
- ABS: N = 41,238; n = 4,206; AAPOR RR3 = 13.2%
- PPD cell: N = 79,737; n = 794; AAPOR RR3 = 2.2%

HH Base Weight Adjustments by Frame

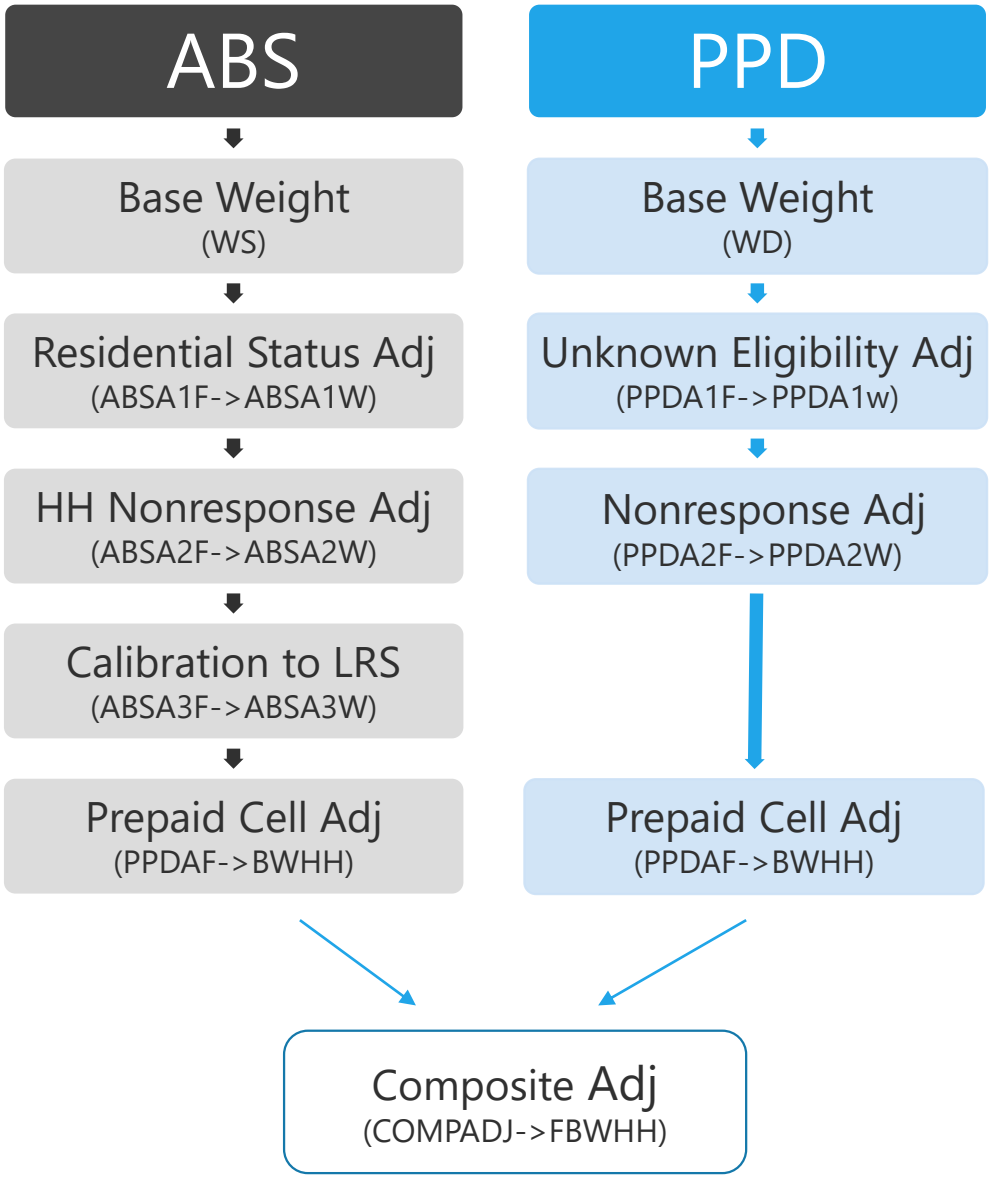
HH Base Weight Adjustments by Frame



HH Base Weight Adjustments by Frame



HH Base Weight Adjustments by Frame



Prepaid Cell Adjustment

- Household selection probabilities need to account for the number of adults in the household with a prepaid cell phone.
- $$PPDAF_i = \begin{cases} 1/PPD_i\,, & i \in PPD \cup ABS(PPD) \\ 1, & i \in ABS(\sim PPD) \end{cases}$$
- PPD_i is the number of adults in HH who have a prepaid cell phone

Composite Adjustment

- Households with an adult who has a prepaid cell phone could be sampled from both frames.
- When the two samples are combined, we make a composite adjustment to account for the frame overlap.
- Reweights all households with a prepaid cell phone, regardless of the sample frame, so that they represent the proportion of households in the ABS frame that have at least one adult with prepaid cell phone.
- $$COMPADJ_i = \begin{cases} 1, \ i \in ABS(\sim PPD) \\ \sum_{i \in ABS(PPD)} ABSA3W_i / \sum_{i \in PPD \cup ABS(PPD)} BWHH_i, \ i \in PPD \cup ABS(PPD) \end{cases}$$

Composite Adjustment

Adjusted Base Weight Sums

		ABS(PPD)	ABS(~PPD)	Total ABS	PPD Frame	
Calibration to LRS	$\sum ABSA3W$	373,422	2,244,075	2,617,497	NA	
Prepaid Cell Adjustment	$\sum BWHH$	295,525	2,244,075	2,539,600	736,416	
Composite Adjustment						
$\frac{\sum_{ABS(PPD)} ABSA3W}{\sum_{PPD \cup ABS(PPD)} BWHH} = \frac{373{,}422}{295{,}525 + 736{,}416}$		0.3619	1.0000		0.3619	
		ABS(PPD)	ABS(~PPD)		PPD Frame	Total ABS+PPD
Final HH-level BW	$\sum FBWHH$	106,940	2,244,075		266,482	2,617,497

Composite Adjustment

Adjusted Base Weight Sums

		ABS(PPD)	ABS(~PPD)	Total ABS	PPD Frame	
Calibration to LRS	$\sum ABSA3W$	373,422	2,244,075	2,617,497	NA	
Prepaid Cell Adjustment	$\sum BWHH$	295,525	2,244,075	2,539,600	736,416	
Composite Adjustment						
$\frac{\sum_{ABS(PPD)} ABSA3W}{\sum_{PPD \cup ABS(PPD)} BWHH} = \frac{373,422}{295,525 + 736,416}$		0.3619	1.0000		0.3619	
		ABS(PPD)	ABS(~PPD)		PPD Frame	Total ABS+PPD
Final HH-level BW	$\sum FBWHH$	106,940	2,244,075		266,482	2,617,497

Composite Adjustment

Adjusted Base Weight Sums

		ABS(PPD)	ABS(~PPD)	Total ABS	PPD Frame	
Calibration to LRS	$\sum ABSA3W$	373,422	2,244,075	2,617,497	NA	
Prepaid Cell Adjustment	$\sum BWHH$	295,525	2,244,075	2,539,600	736,416	

Composite Adjustment

$\frac{\sum_{ABS(PPD)} ABSA3W}{\sum_{PPD \cup ABS(PPD)} BWHH} = \frac{373,422}{295,525 + 736,416}$		0.3619	1.0000		0.3619	
		ABS(PPD)	ABS(~PPD)		PPD Frame	Total ABS+PPD
Final HH-level BW	$\sum FBWHH$	106,940	2,244,075		266,482	2,617,497

Optimal ABS/PPD Sample Allocation

Optimal ABS/PPD Sample Allocation

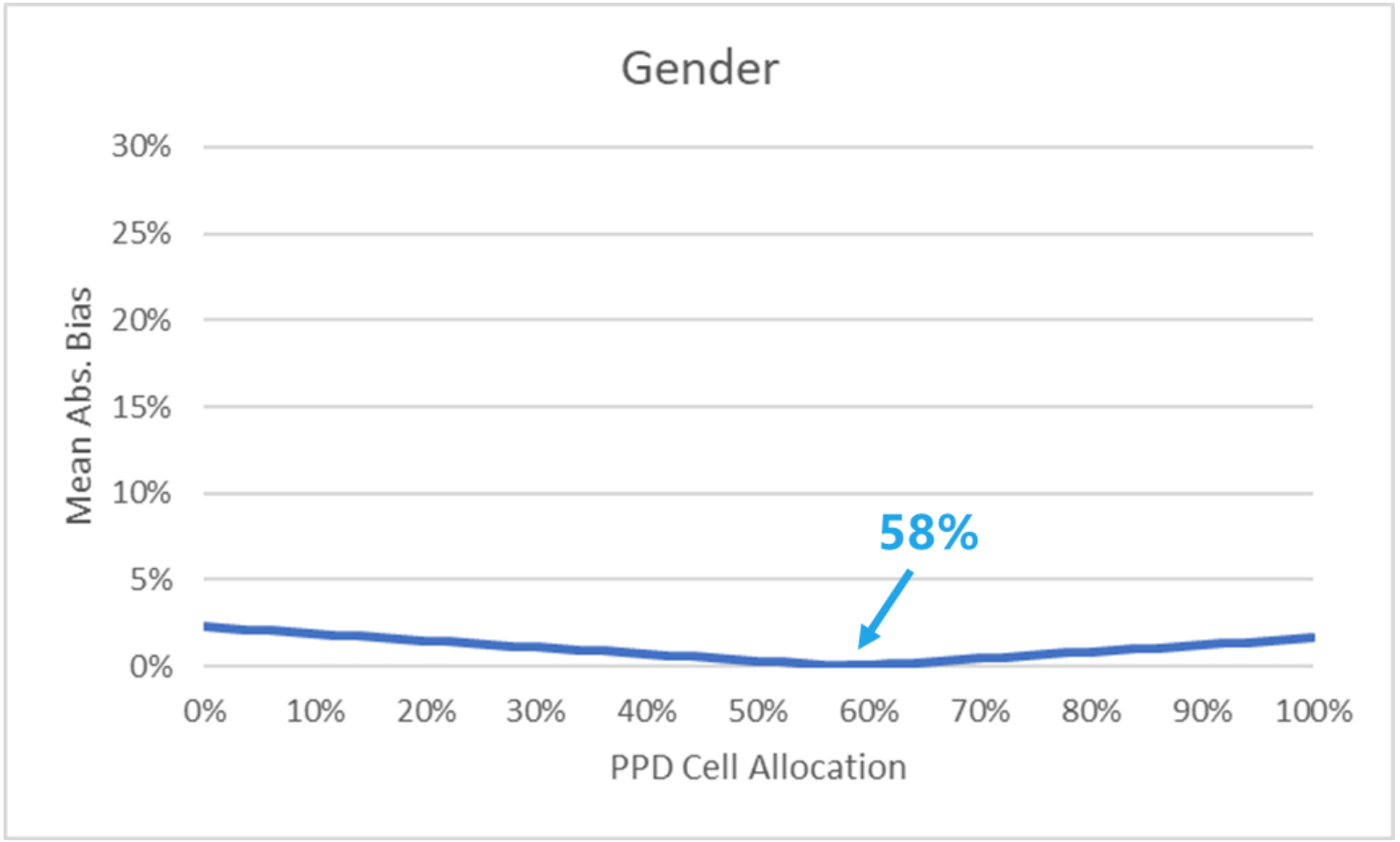
- ABS and prepaid cell samples reach groups that are quite different demographically.
- What is the optimal allocation of ABS and PPD samples?
- We define optimal allocation as the one that minimizes mean absolute bias in demographic characteristics.

Optimal ABS/PPD Sample Allocation

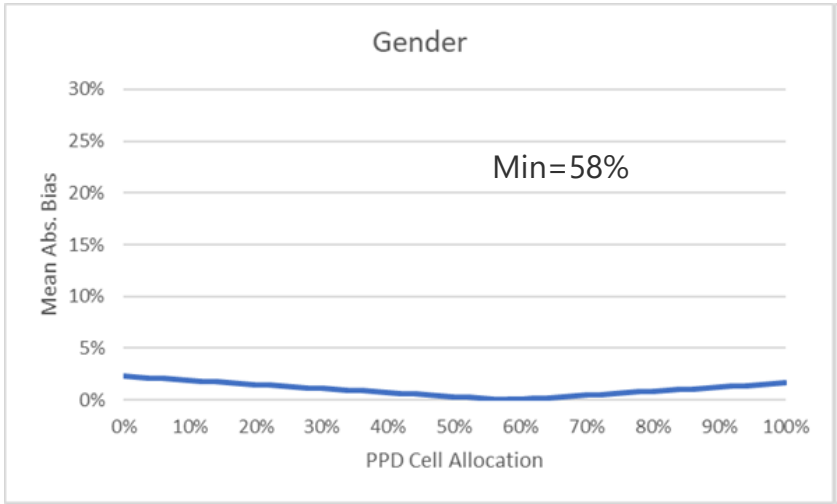
- Mean Absolute Bias
- $MAB = n^{-1} \sum_{i=1}^{n} |p_i - P_i|$

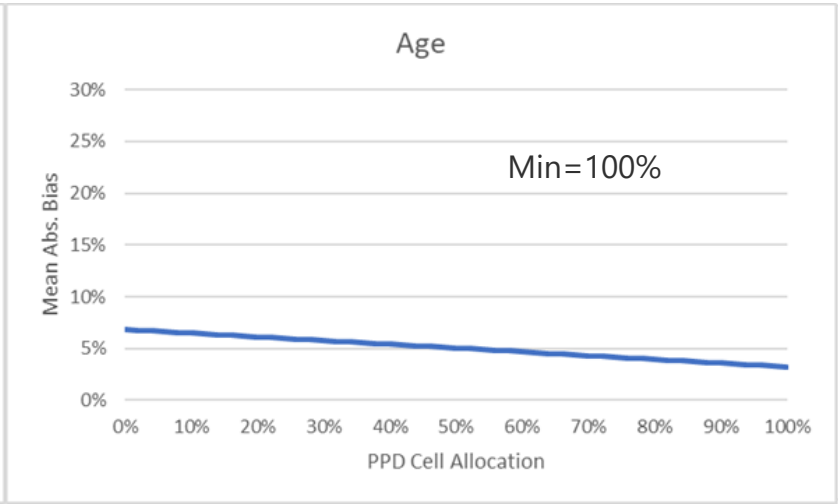
	Benchmark P_i	Sample p_i	$\lvert P_i - p_i \rvert$
Less than HS grad.	8.6%	4.8%	3.8%
HS graduate	24.4%	15.5%	8.9%
Some college	23.7%	21.4%	2.3%
College grad.	43.2%	58.4%	15.2%
	100.0%	100.0%	
		$\sum \lvert P_i - p_i \rvert =$	30.2%
	$MAB = \frac{1}{4} \sum \lvert P_i - p_i \rvert =$		7.6%

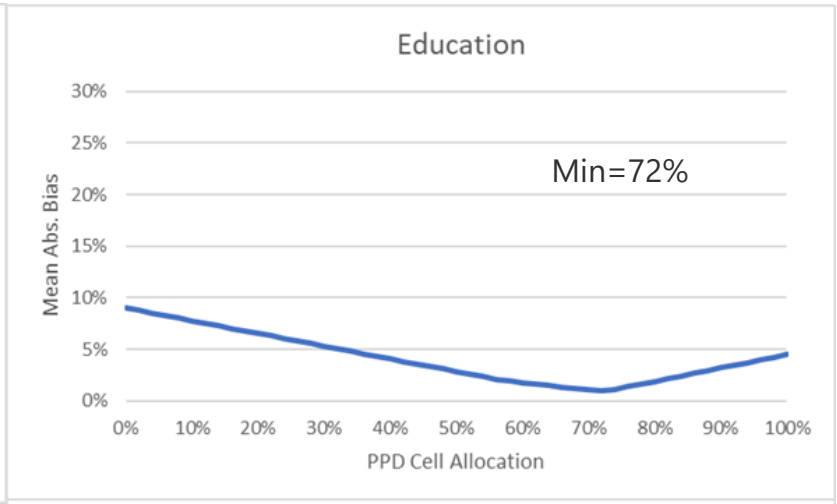
Gender Optimal ABS/PPD Sample Allocation

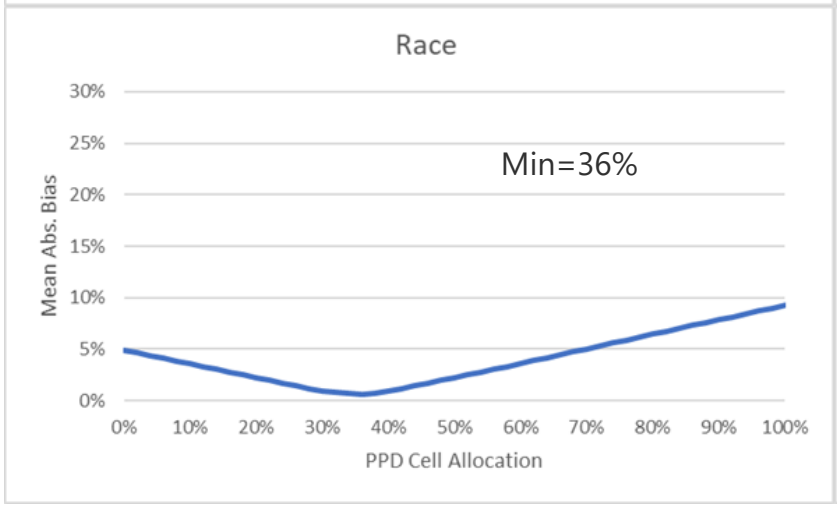


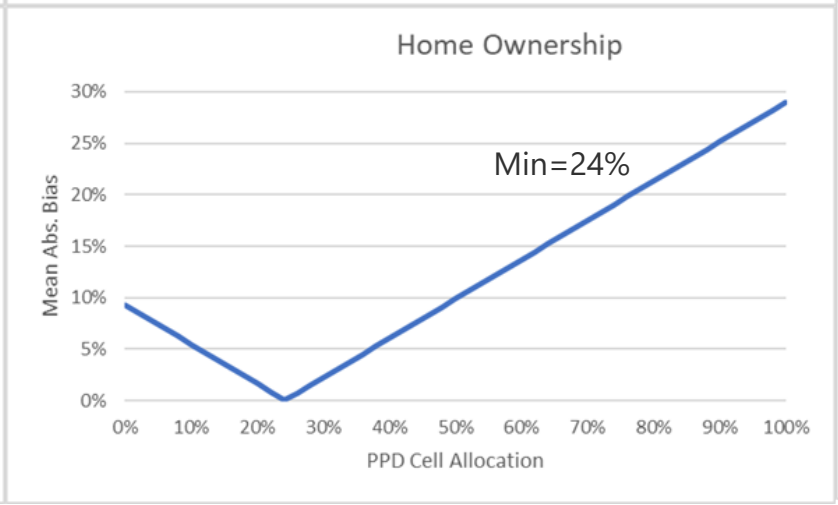
Optimal ABS/PPD Sample Allocation

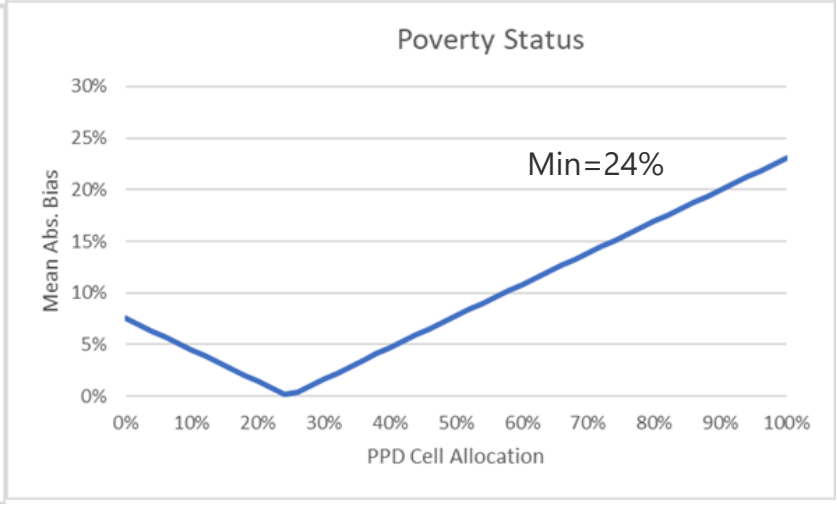




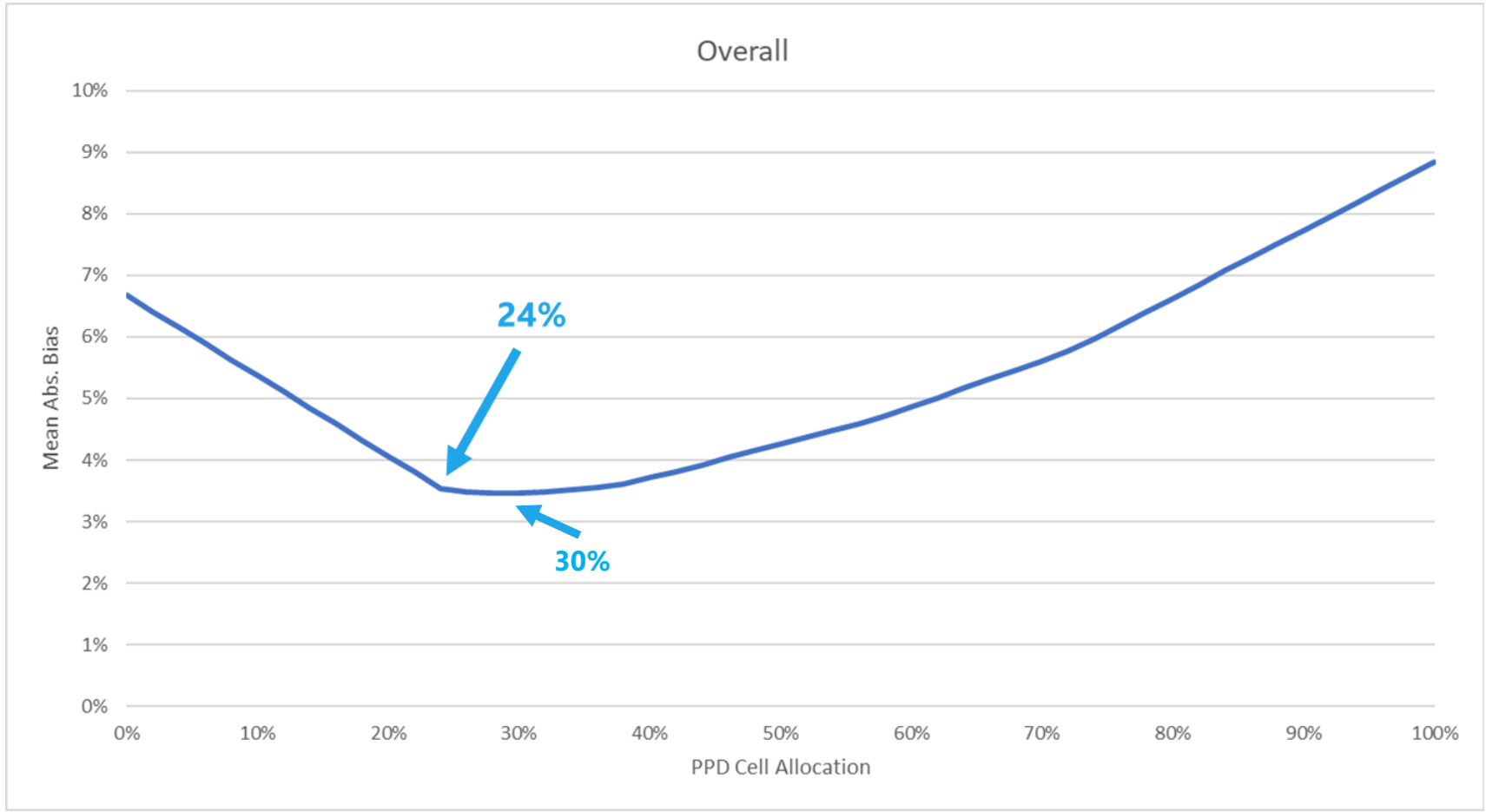








Optimal ABS/PPD Sample Allocation - Overall



Data Collection Costs – ABS vs. PPD

Data Collection Costs – ABS vs. PPD

- Data collection costs are a function of many variables, including
  - Length of survey
  - Telephone contact procedures
  - Mailing protocol
  - Incentive structure
- Interviews from PPD cell frame are more expensive that interviews from ABS.
- For MHIS which targeted the MA general population, PPD interviews were approx. 4.7 times more expensive than ABS interviews

Data Collection Costs– ABS vs. PPD

- Relative CPI is different for subgroups that have different incidences in each frame.*

Target Group		ABS	PPD	CPI Ratio (PPD/ABS)
18+	Incidence	100.0%	100.0%	4.7
	CPI	1	4.7	

*Incidences in table are base weighted

Data Collection Costs– ABS vs. PPD

- Relative CPI is different for subgroups that have different incidences in each frame.*

Target Group		ABS	PPD	CPI Ratio (PPD/ABS)
18+	Incidence	100.0%	100.0%	4.7
	CPI	1	4.7	
Renter	Incidence	24.3%	62.6%	1.8
	CPI	4.1	7.5	

*Incidences in table are base weighted

Data Collection Costs– ABS vs. PPD

- Relative CPI is different for subgroups that have different incidences in each frame.*

Target Group		ABS	PPD	CPI Ratio (PPD/ABS)
18+	Incidence	100.0%	100.0%	4.7
	CPI	1	4.7	
Renter	Incidence	24.3%	62.6%	1.8
	CPI	4.1	7.5	
Below 200% FPL	Incidence	13.2%	43.8%	1.4
	CPI	7.6	10.7	

*Incidences in table are base weighted

Data Collection Costs– ABS vs. PPD

- Relative CPI is different for subgroups that have different incidences in each frame.*

Target Group		ABS	PPD	CPI Ratio (PPD/ABS)
18+	Incidence	100.0%	100.0%	4.7
	CPI	1	4.7	
Renter	Incidence	24.3%	62.6%	1.8
	CPI	4.1	7.5	
Below 200% FPL	Incidence	13.2%	43.8%	1.4
	CPI	7.6	10.7	
African- American	Incidence	3.1%	12.5%	1.2
	CPI	32.3	37.6	

*Incidences in table are base weighted

Data Collection Costs– ABS vs. PPD

- Relative CPI is different for subgroups that have different incidences in each frame.*

Target Group		ABS	PPD	CPI Ratio (PPD/ABS)
18+	Incidence	100.0%	100.0%	4.7
	CPI	1	4.7	
Renter	Incidence	24.3%	62.6%	1.8
	CPI	4.1	7.5	
Below 200% FPL	Incidence	13.2%	43.8%	1.4
	CPI	7.6	10.7	
African- American	Incidence	3.1%	12.5%	1.2
	CPI	32.3	37.6	
LT HS Grad	Incidence	2.6%	13.2%	0.9
	CPI	38.5	35.6	

*Incidences in table are base weighted

Data Collection Costs– ABS vs. PPD

- Relative CPI is different for subgroups that have different incidences in each frame.*

Target Group		ABS	PPD	CPI Ratio (PPD/ABS)
18+	Incidence*	100.0%	100.0%	4.7
	CPI	1	4.7	
Renter	Incidence	24.3%	62.6%	1.8
	CPI	4.1	7.5	
Below 200% FPL	Incidence	13.2%	43.8%	1.4
	CPI	7.6	10.7	
African- American	Incidence	3.1%	12.5%	1.2
	CPI	32.3	37.6	
LT HS Grad	Incidence	2.6%	13.2%	0.9
	CPI	38.5	35.6	
Hispanic	Incidence	5.0%	29.4%	0.8
	CPI	20.0	16.0	

*Incidences in table are base weighted

Conclusion

1. Easy to combine ABD and PPD samples with a base weight composite adjustment.
2. Optimal sample allocation depends on which demographics are chosen to optimize.
3. Cost differential between the two frames varies widely and is dependent on target group incidences. PPD cell could be more cost effective for some groups.
4. I have never been called "Special J" before today.

References

- Bruce, Antonio, J. Gregory Robinson, and Monique V. Sanders. 2001. "Hard-to-Count Scores and Broad Demographic Groups Associated with Patterns of Response Rates in Census 2000." Paper presented at the Joint Statistical Meetings of the American Statistical Association, Atlanta, Georgia, USA.
- Feskens, Remco, Joop Hox, Gerty Lensvelt-Mulders, and Hans Schmeets. 2007. "Nonresponse Among Ethnic Minorities: A Multivariate Analysis." Journal of Official Statistics 23:387–408.
- Hu, S. Sean, Michael W. Link, and Ali H. Mokdad. 2010. "Reaching Linguistically Isolated People: Findings from a Telephone Survey Using Real-Time Interpreters." Field Methods 22:39–56.
- Smith, Tom W. 2014. "Hard-to-Survey Populations in Comparative Perspective." In Hard-to-Survey Populations, edited by Roger Tourangeau, Brad Edwards, Timothy P. Johnson, Kirk M. Wolter, and Nancy Bates, 21–36. Cambridge: Cambridge University Press.

Thank You, AAPOR

ssrs
research. refined.

JONATHAN BEST
JBEST@SSRS.COM

JAMES McKINSTRY
JMCKINSTRY@SSRS.COM

ARIFAH HASANBASRI
AHASANBASRI@SSRS.COM

CHIA.
center
for health
information
and analysis

CHRISTINE LOVERIDGE
CHRISTINE.LOVERIDGE@CHIAMASS.GOV

HUONG TRIEU
HUONG.TRIEU@CHIAMASS.GOV

AAPOR
Annual Conference
2022
Chicago, IL

May 13, 2022