Applications of Predictive Modeling to Survey Design and Operation in Address-based Samples

Presented by: Cameron McPhee | SSRS Chief Methodologist

THURSDAY, MARCH 17, 2022

ssrs
research. refined.

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Schedule

INTRODUCTION

OVERVIEW OF PREDICTIVE MODELING

APPLICATION EXAMPLES RELATED TO:

Sample Design	Operational Improvement	Cost Manipulation (Responsive/Adaptive Design)	Sensitivity Models	Weighting

SUMMARY AND CONSIDERATIONS

Q&A

Introduction: The Changing Landscape

- ~~Techno-apocalypse~~ Rise of cellular and alternative telephone technology leads to decline in landline coverage
  - Traditional RDD methods need to shift dual frame
- Caller-ID and other privacy concerns makes it more difficult to reach potential respondents
  - Response rates decline rapidly
  - Costs increase
  - Risks of bias increase
- Notable and public evidence of polling error due to nonignorable nonresponse (AAPOR, 2020)

INCREASED INNOVATION!

Smart People Start Trying Smart Things…

- Growth of carefully tailored designs (Dillman, et. al., 2014) driven by leverage/saliency theory (Groves, Singer, and Corning 2000)
  - Link, M., and A.T. Burks. 2013 test differential incentives based on auxiliary characteristics
  - Olson, Smyth, and Wood (2012) examine tailoring follow-up survey request modes based on demonstrated mode preference
- Increased use of mixed modes for recruitment and response to minimize burden
  - Beginning in 2000s, large-scale shift to alternative sampling frames (e.g., ABS) and self-administered or mixed-modes (AAPOR, 2019)
- Adaptive and responsive design techniques that incorporate results from previous surveys or earlier waves to target resources and/or interventions
  - Rise in collection and use of paradata to monitor and manipulate resource allocation (Groves et. al., 2005, Groves & Heeringa, 2006; Peychev et. al., 2009; Wagner, 2013; West et. al., 2015).

What is Predictive Modeling?

ssrs
research. refined.

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Predictive Modeling is Simply…

- A collection of **statistical tools and techniques** to fine tune study methods using multivariate models
  - Can what we already know tell us anything about what we don’t know?
- “Response Propensity” models are a subset of predictive models
  - Most models described today are versions or adaptations of response-propensity models
- There are many statistical options for building models
- For this application, typically the goal is to identify or classify sample units based on information known at T1 to differentially deploy treatment or intervention in order to impact outcomes at T2.
- Advantage of the multivariate approach is the aggregation of the impact of different characteristics on response or another outcome while limiting the number of treatment groups or conditions.

What kinds of questions can we use predictive modeling to answer in the context of survey methodology?

1. Who will most likely respond to my survey? Who will be more difficult to reach?
2. How do I get more ________________ in my sample?
3. Who should receive bilingual materials and who should not?
4. How can I use more expensive interventions strategically?
5. When should I "give up" on certain cases?
6. How do I reduce cost without impacting representation?

Predictive Modeling Process*: Step 1 - Prediction

- Requires prior survey data to build predictive model
  - A prior wave of a longitudinal collection (historically a common application)
  - A previous iteration of a repeated cross-sectional study
  - A different study using the same methodology with the same/similar population
  - We have seen *some* success with cross-population applications.
- Often models will require full sample data (including nonrespondents), but there are situations where the models can be built on the respondent data alone

This is our Training Data

*See Lavrakas, P.J., Jackson, M. and McPhee, C. (2018) for a general overview of this process.

Dependent Variable(s)

- Outcome (dependent) variable is usually dichotomous; can be multinomial categorical or continuous
  - Often dependent variable is the response indicator from the previous study (response propensity model)
  - Can be more nuanced response indicator
    - Responded after first contact attempt
    - Responded by a specific mode vs. an alternative mode
- In some situations, the dependent variable is specific information about the respondents from the training data
  - Demographic indicator(s)
  - Other eligibility criterion indicator

Independent Variables

- Predictor variables may come from many places and depend on whether the adaptive design is static or dynamic (Coffey, Reist, and Miller, 2019)
  - Static adaptive designs use information available prior to data collection to tailor contact methods and timing
  - Dynamic adaptive designs employ information available before and during data collection to enable interventions at multiple decision points (e.g., paradata)
- For ABS, *most* applications are static because the timing of mailings require pre-determined classifications.
- Therefore, independent variables typically must exist prior to data collection

Independent Variables continued

- May be based on substantive hypotheses or prior experience, but may be more exploratory to begin with
  - Sampling frame data
  - Auxiliary data that can be appended to the sampling frame at the sampled unit level
  - In longitudinal or panel samples, there is all the data from the previous round(s)
  - Geographic-area data
    - Census bureau-collected tract-level or block group level demographic distributional data (including the "Low Response Score (LRS)" and ACS mail-back response rate
    - State, region, city-level data
    - GIS-based auxiliary information

Great Resource: https://www.census.gov/topics/research/guidance/planning-databases.html

  - Operational paradata when applicable (e.g., call history)

Modeling Method

- No single "best" way to build the models. May be driven by extent of *a priori* understanding of the data and variable relationships
- Logistic Regression is a great option for dichotomous outcomes
  - Straightforward, easily interpretable
  - Great if there is prior knowledge of likely predictors
  - Can be implemented in stepwise method to obtain a parsimonious set of predictors
- Nonparametric Methods (e.g., CART, CHAID, Random Forests and similar)
  - More data-driven in their approach
  - Can identify 2nd and 3rd order interactions
  - Should be pared down or "pruned" to avoid overfitting
  - Can be used directly to assign propensity scores or used for variable selection to be applied to a simpler logistic model

Predictive Modeling Process: Step 2 - Classification

- Utility is in the ability to use the propensity scores for classification into treatment groups
  - Less importance placed on independent variable-level coefficients
- Formation of treatment groups depends on how they are being used
  - May need to identify cut scores or thresholds
  - May use propensity scores to form strata
- Some statistical modeling processes (e.g., classification trees) may classify units directly based on parameters you set, others require you to determine appropriate thresholds
- Propensity scores may be converted to “sensitivity scores” to identify cases for whom a specific treatment or protocol is likely to be the most effective (Jackson, Medway, and Megra, 2021) or transformed to R-indicators to implement treatments based on representation of target groups (Coffey, Reist, and Miller, 2019)
- Number of groups will be based on number of alternative treatments being assigned

Model Evaluation

- All predictive models should be evaluated for accuracy prior to implementation
- Accuracy of the classification should always be **tested on an "out-of-sample"** **test data** set in order to prevent overfitting and help protect against unintended consequences when it is applied to you "real" study.
  - Some statistical processes (e.g., Random Forests) often have a multi-fold cross-validation as part of the model building
  - Sometimes you will need to partition the data into "training" and "test" samples prior to model estimation
- Sometimes a mediocre propensity model can yield meaningful treatment groups through appropriate classification.

Applications to Survey Design and Operations

ssrs
research. refined.

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Predictive Modeling for Sample Design

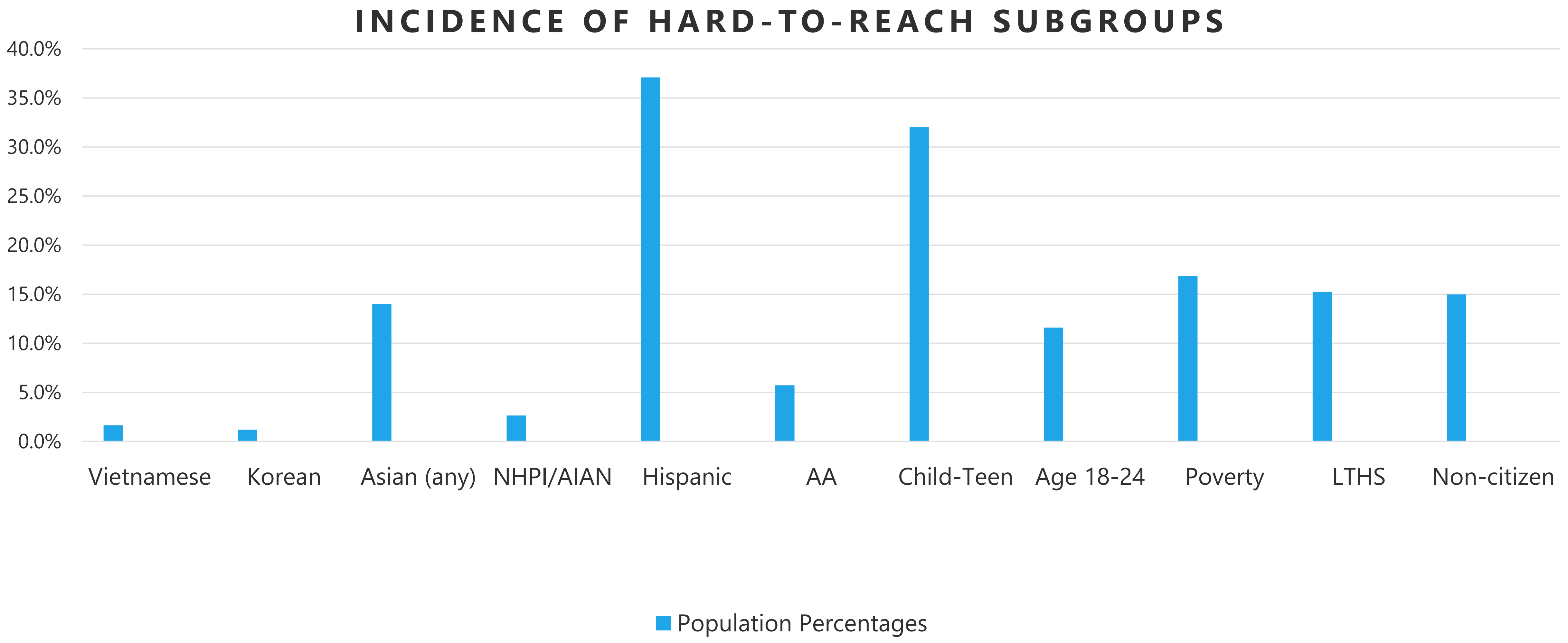
- Used to deviate from random sampling in order to reach more of a targeted group
- May use modeling to alter selection rates for different sample strata (e.g., oversample likely respondents with targeted characteristics)
  - Could be to reach more eligible respondents
  - Could be to compensate for differential response rates by subgroup
- Some designs may use predictive modeling to deviate from probability samples due to limited resources (e.g., only sample cases predicted to be in a targeted group)

Example: California Health Interview Survey (CHIS)

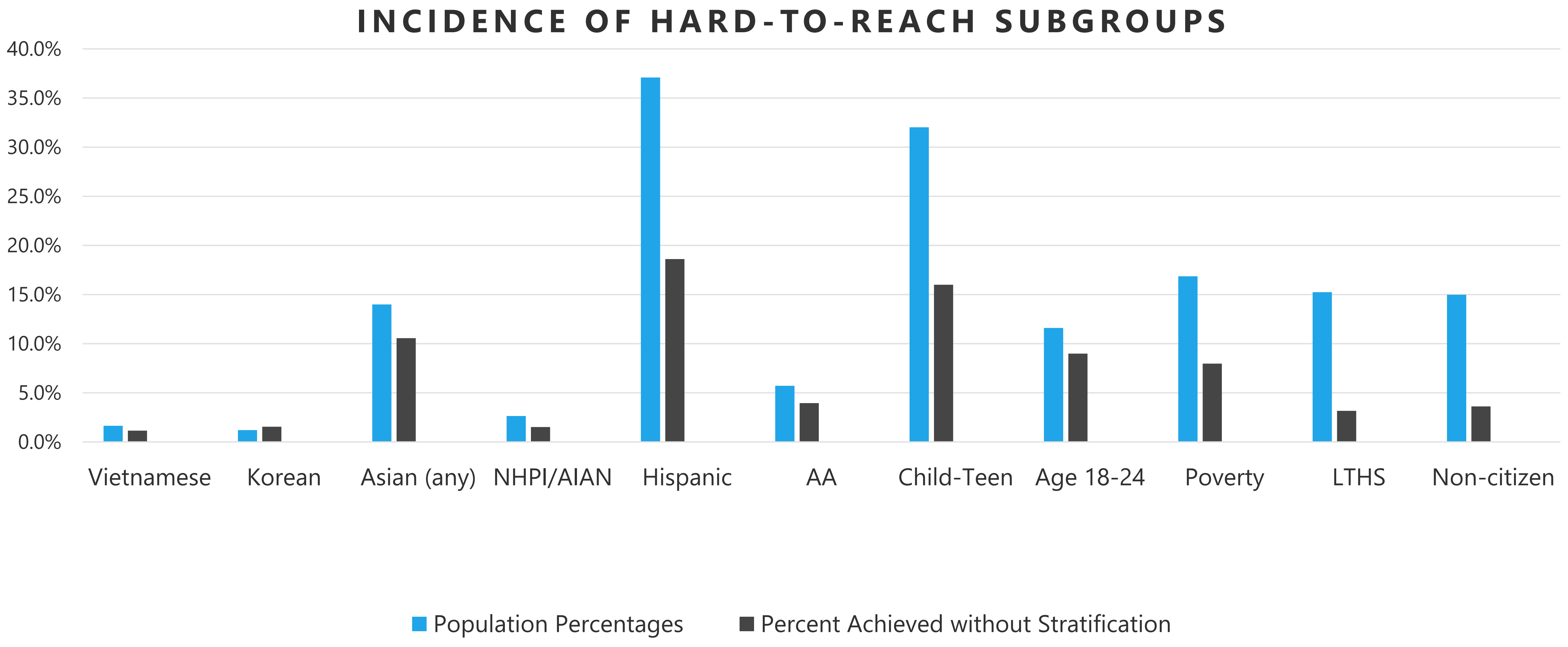
- 40,000 interviews across 2 years
- Representative of California population
- Target sample sizes for all counties, plus sub-counties and health districts in Los Angeles and San Diego
- Must produce reliable estimates for important demographic subgroups:

  - ✓ Vietnamese
  - ✓ Korean
  - ✓ Other Asian
  - ✓ Hispanic
  - ✓ Black/African-American
  - ✓ Native Hawaiian/Pacific Islanders
  - ✓ Households with children/teens
  - ✓ Young adults (age 18-24)
  - ✓ Households in poverty
  - ✓ Adults with less than HS educ.
  - ✓ Non-U.S. citizens

The Challenge



The Challenge



Model Estimation

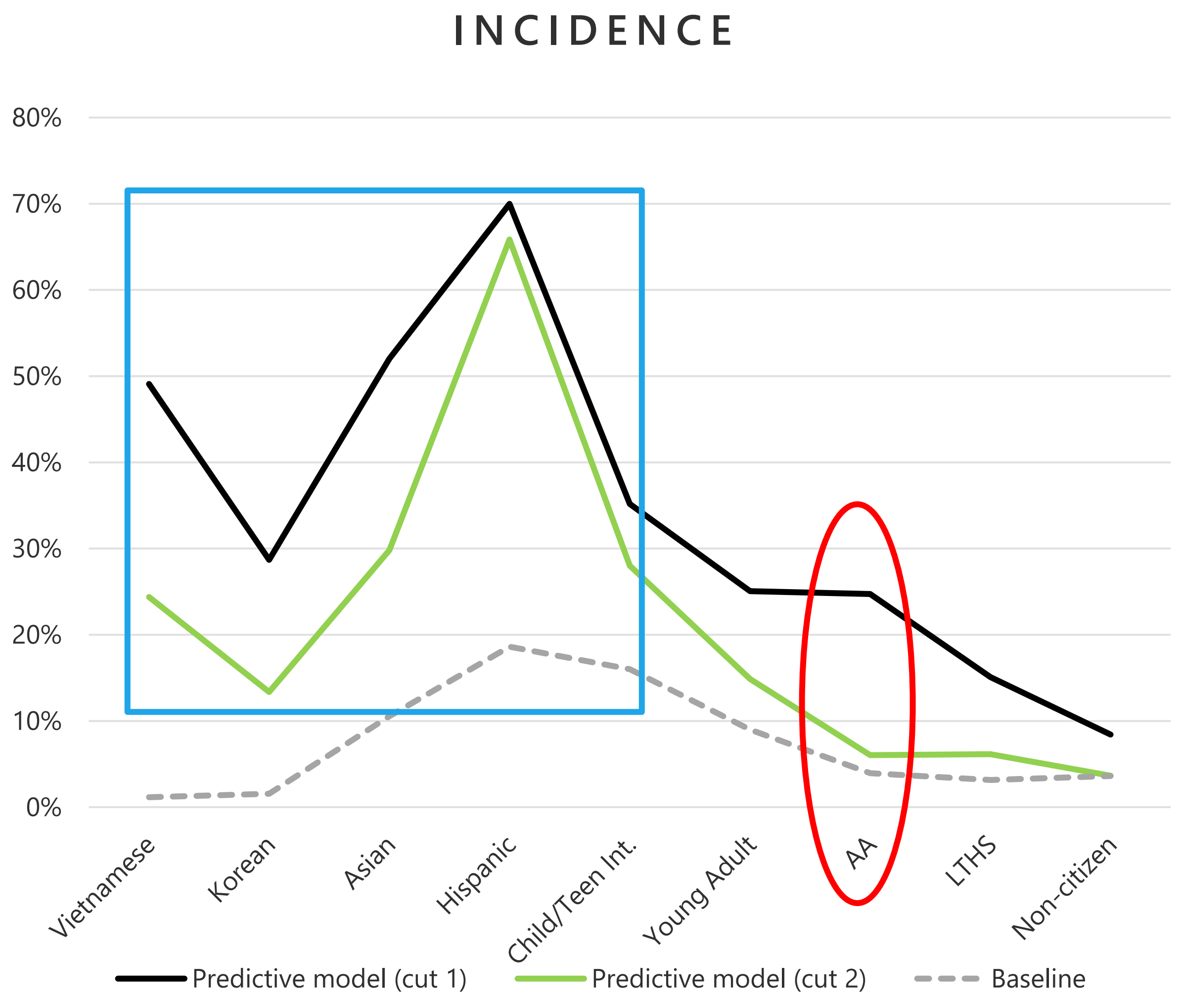
- Auxiliary data linked (by address) to the CHIS: 2019 and 2020 data
- Sources include
  - Voter and consumer data
  - Address characteristics (e.g., mail delivery type, building classification)
  - Surnames list flags
  - Census block-level characteristics (e.g., % non-white, % in poverty, Low Response Score)
- Random Forest models estimated using 2019 data for each characteristic of interest
- Predictions applied to 2020 data to test model accuracy

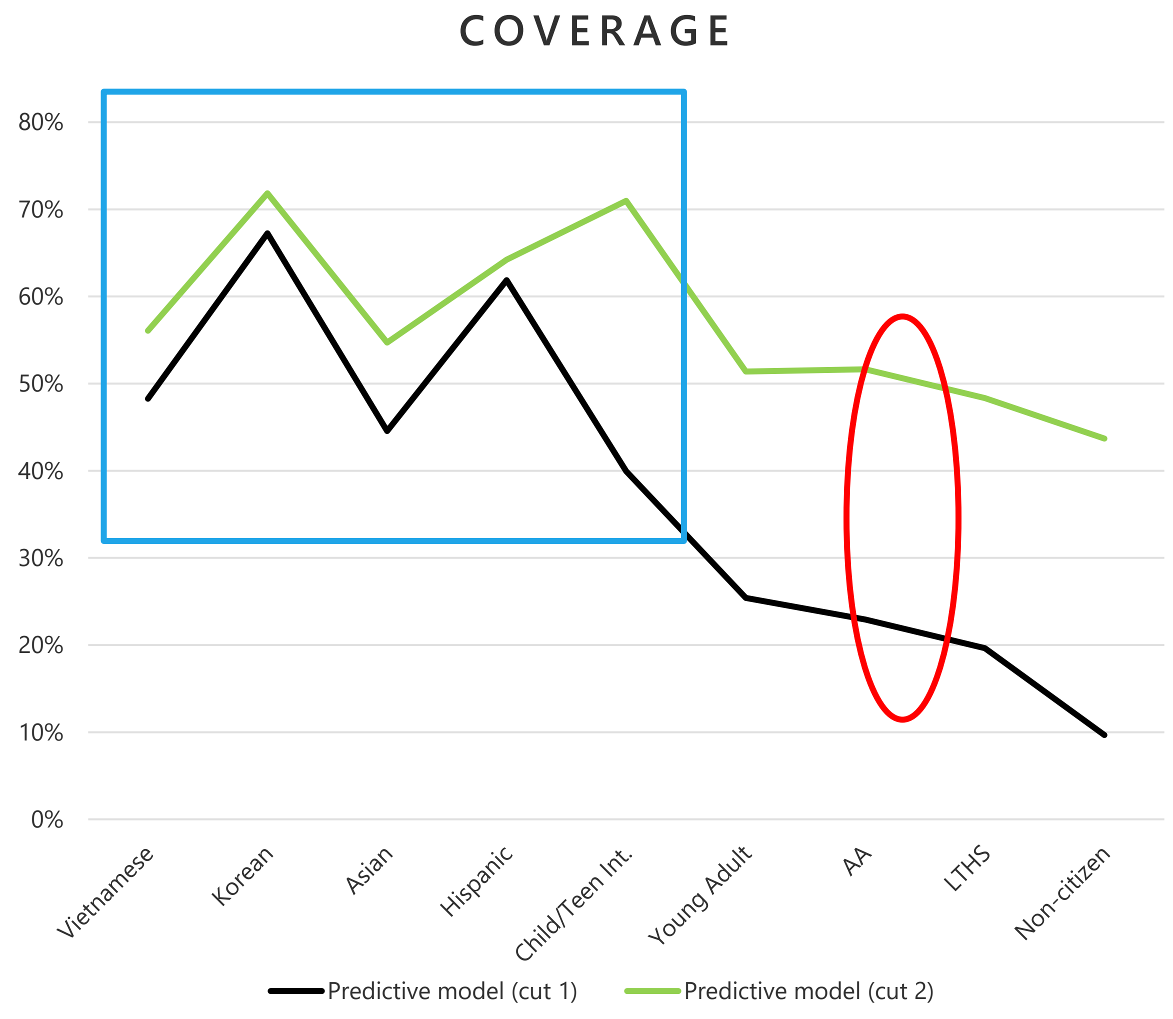
PREVIOUS DATA

TRAINING SET

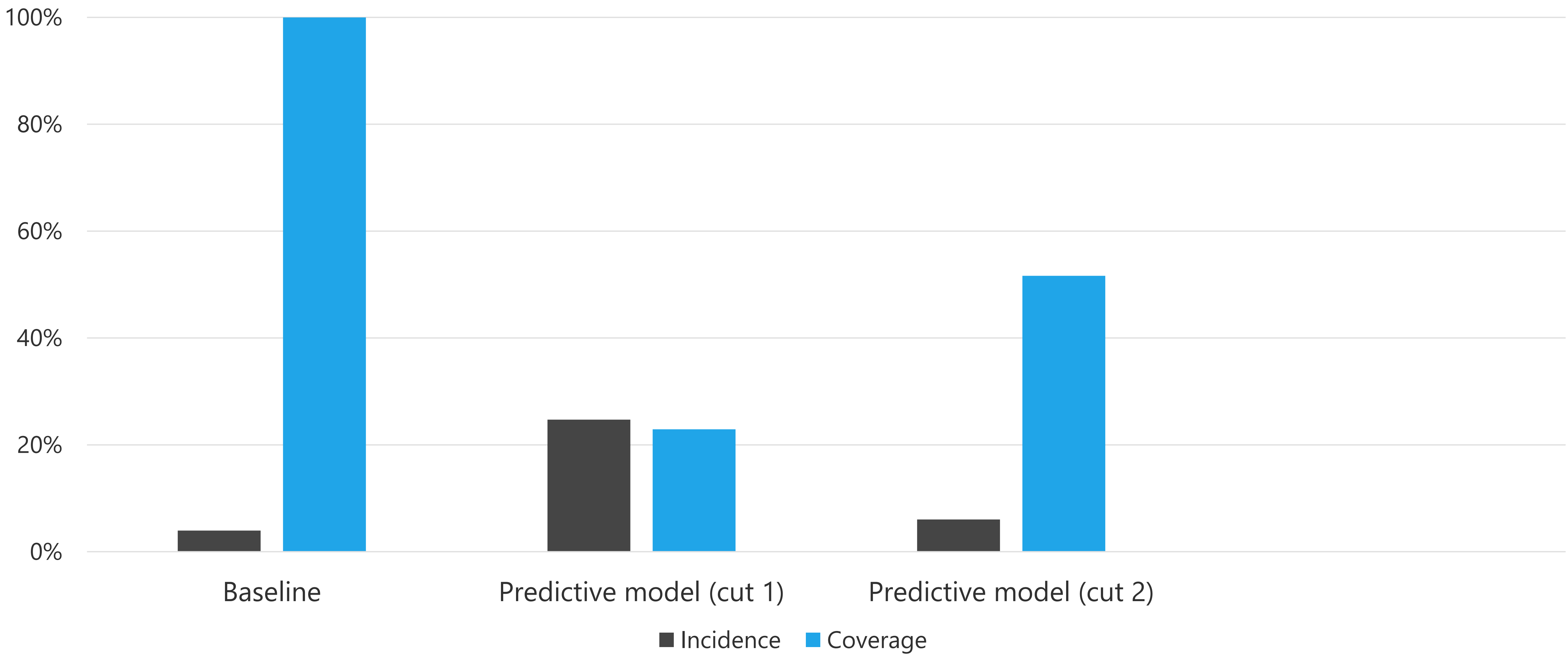
TEST SET

Model Estimation

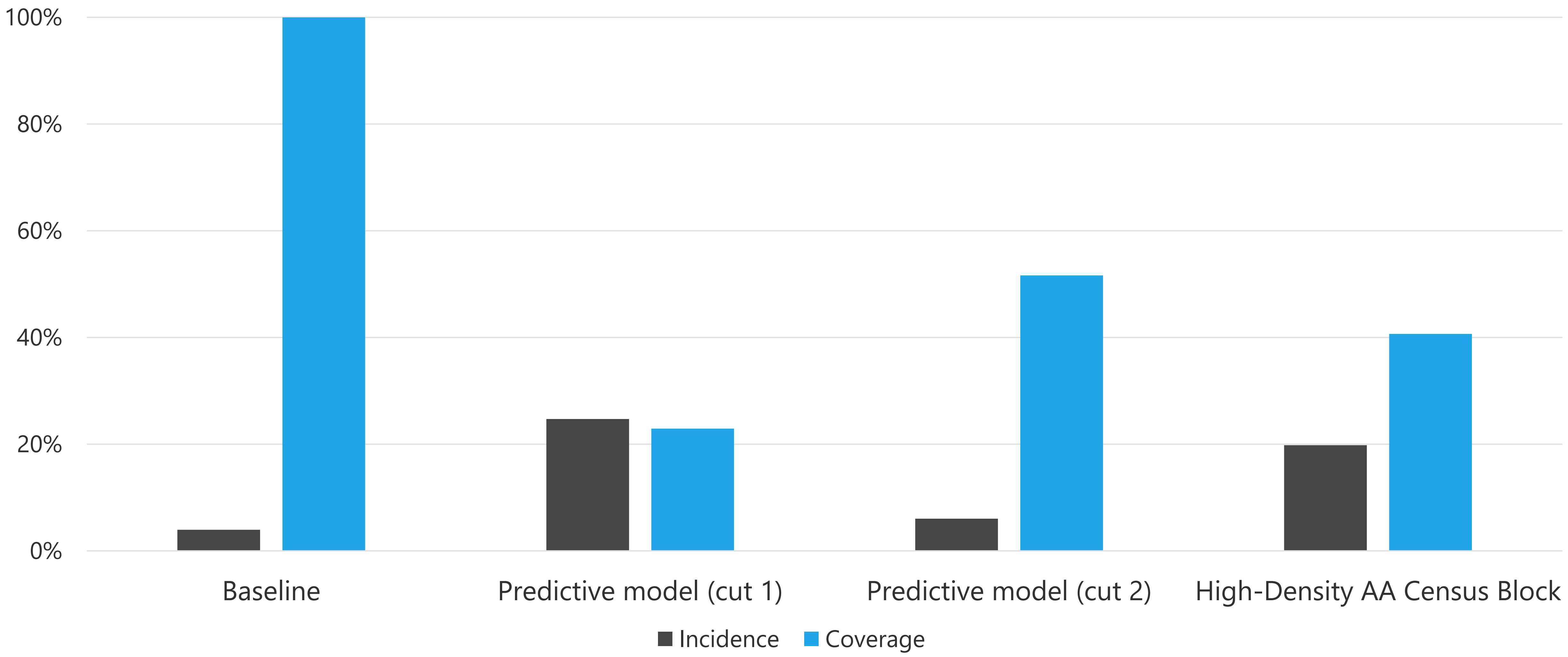




So how can we target African-American respondents?



So how can we target African-American respondents?



Next Steps

- Applying these predictions to sample designs depends on the goal
  - May simply want to identify a subpopulation from which to sample
    - Think about coverage!
  - Disproportionate sampling rates necessitates **stratification**
    - May create a single strata for target cases
    - May develop prioritization

For CHIS

- Strata for sampling require mutually exclusive hierarchy
- The more overlap across strata, the less effective the lower-ranked strata are for sampling (i.e., reduces incidence)
- Tried to put narrowest non-overlapping strata higher up
- Required several iterations, analysis, & redesign

Final Stratification

1. Vietnamese surname flag
2. Korean surname flag
3. Modeled Asian-language interview
4. Modeled Spanish-language interview
5. Hispanic surname flag
6. Other high-density non-English BG
7. Asian (other) surname flag or Asian model
8. Other high-density African-American BG
9. Modeled HH with children
10. Over 65 age flag or model
11. Residual – Matched to auxiliary data
12. Residual – Not matched to auxiliary data

And Then…

- Need to evaluate different sets of sampling fractions across strata
- For CHIS, we evaluated on three metrics:
  - Incidence improvement
  - Design Effect
  - Yield

Strata
Vietnamese
Korean
Asian-lang. Int.
Spanish-lang.
Hispanic
Non-Eng. BG
Asian
AA BG
HH w/ Children
Age 65+
Residual – No Match
Residual - Match

Avg. Incidence improvement
Design Effect (UWE)
Est. Yield

And Then…

- Need to evaluate different sets of sampling fractions across strata
- For CHIS, we evaluated on three metrics:
  - Incidence improvement
  - Design Effect
  - Yield

Strata	Allocation 1 (Aggressive)
Vietnamese	5
Korean	4.5
Asian-lang. Int.	2
Spanish-lang.	5
Hispanic	3.5
Non-Eng. BG	2
Asian	2
AA BG	5
HH w/ Children	4
Age 65+	0.5
Residual – No Match	0.5
Residual - Match	1

Avg. Incidence improvement	37.2%
Design Effect (UWE)	2.69
Est. Yield	17.0

And Then…

- Need to evaluate different sets of sampling fractions across strata
- For CHIS, we evaluated on three metrics:
  - Incidence improvement
  - Design Effect
  - Yield

Strata	Allocation 1 (Aggressive)	Allocation 2 (Moderate)
Vietnamese	5	5
Korean	4.5	4.5
Asian-lang. Int.	2	2
Spanish-lang.	5	3
Hispanic	3.5	2.5
Non-Eng. BG	2	2
Asian	2	2
AA BG	5	3
HH w/ Children	4	3
Age 65+	0.5	0.75
Residual – No Match	0.5	0.75
Residual - Match	1	1

Avg. Incidence improvement	37.2%	35.4%
Design Effect (UWE)	2.69	1.69
Est. Yield	17.0	16.6

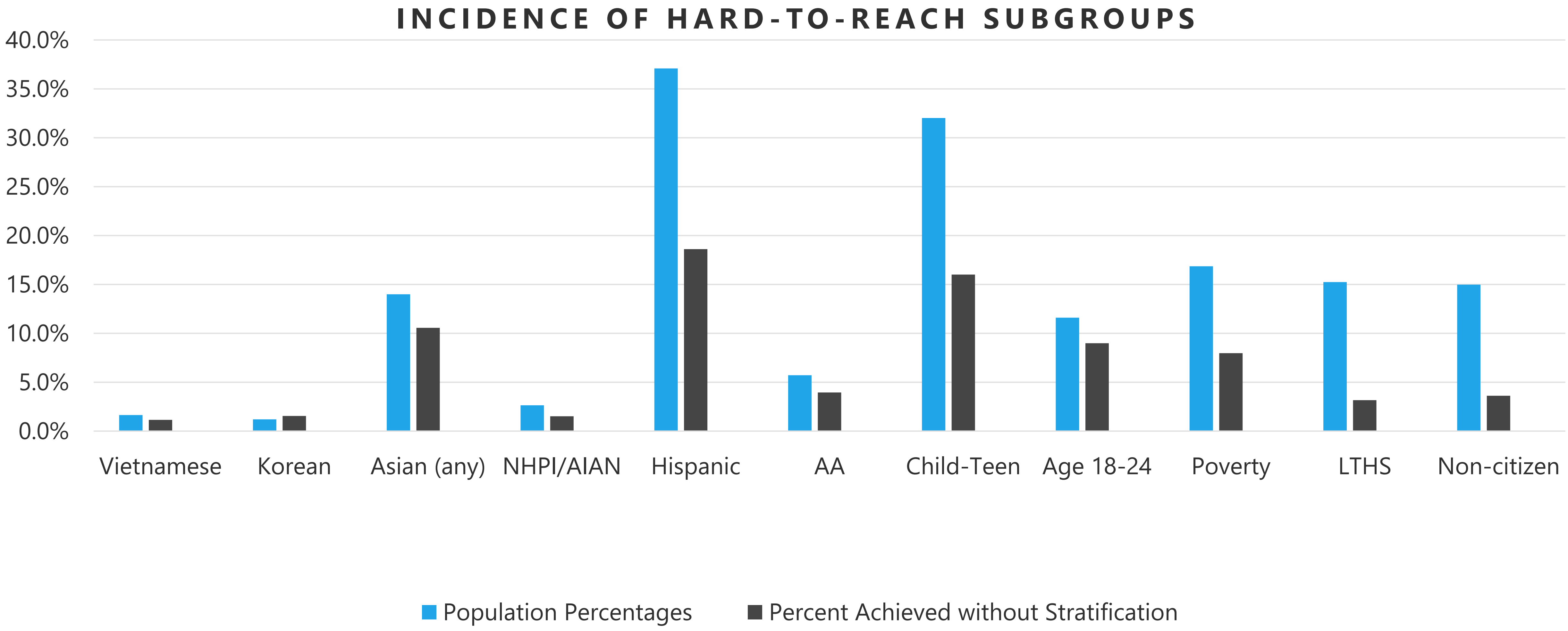
And Then…

- Need to evaluate different sets of sampling fractions across strata
- For CHIS, we evaluated on three metrics:
  - Incidence improvement
  - Design Effect
  - Yield

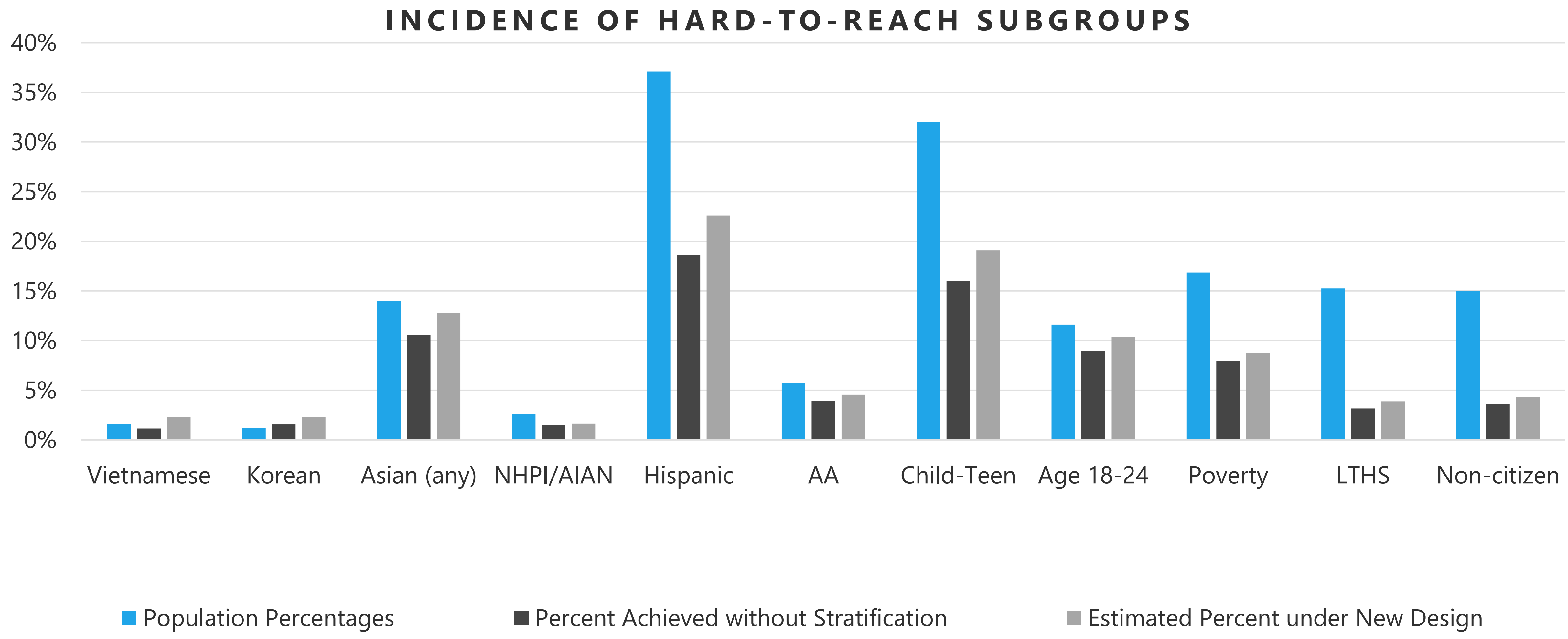
Strata	Allocation 1 (Aggressive)	Allocation 2 (Moderate)	Allocation 3 (Final)
Vietnamese	5	5	4
Korean	4.5	4.5	3.5
Asian-lang. Int.	2	2	2
Spanish-lang.	5	3	1.75
Hispanic	3.5	2.5	0.8
Non-Eng. BG	2	2	1.5
Asian	2	2	1.1
AA BG	5	3	5
HH w/ Children	4	3	2
Age 65+	0.5	0.75	0.75
Residual – No Match	0.5	0.75	0.65
Residual - Match	1	1	1

Avg. Incidence improvement	37.2%	35.4%	**33.2%**
Design Effect (UWE)	2.69	1.69	**1.42**
Est. Yield	17.0	16.6	**16.2**

The Challenge



The Result


INCIDENCE OF HARD-TO-REACH SUBGROUPS
40%
35%
30%
25%
20%
15%
10%
5%
0%
Vietnamese
Korean
Asian (any)
NHPI/AIAN
Hispanic
AA
Child-Teen
Age 18-24
Poverty
LTHS
Non-citizen
Population Percentages
Percent Achieved without Stratification
Estimated Percent under New Design

Predictive Modeling for Operational Enhancement

- Often, we do not want to treat all sample cases with the same protocol, e.g.,
  - Different language requirements for different sampled cases
  - Different message saliency
  - Different likely mode preference

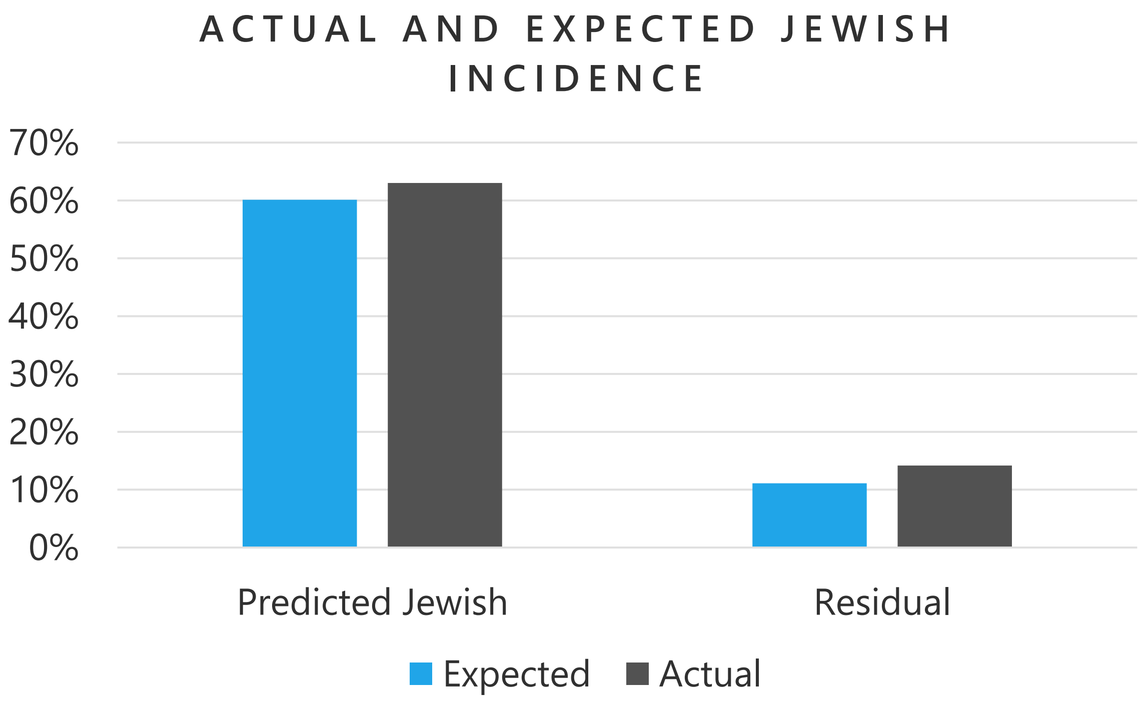
Example: New York Jewish Community Survey*

- Semi-annual survey of 8 counties in New York (5 boroughs + 3 neighboring counties)
- Two-fold Purpose
  1. Collect information on demographics, religious practice, and cultural experience of the area's Jewish population
  2. Generate population estimates for the number of Jews in the service area
- Modeling was necessary to maximize the number of full-survey responses collected from Jewish respondents
- Also, to manipulate operational strategies to appeal to different target groups
  - Identification as a Jewish population study
  - Inclusion of hard copy questionnaires
  - Differential incentives

*Sponsored by the United Jewish Federation of New York

Predictive Model

- No ABS data available for training the model for the pilot study
  - Initial model built using aggregated RDD data from several years of SSRS telephone Omnibus
  - After the pilot, model was re-trained using pilot ABS data (continual improvement)
- Outcome was self-identification as Jewish
- Independent variables appended from commercial databases and voter lists
- Model built and applied through Random Forests



Operational Assignment by Predictive Strata

TABLE 1: LANGUAGE OF OUTREACH MATERIALS BY STRATA AND GEOGRAPHY

	English	English/ Russian	English/ Yiddish
Predicted Jewish	• Bronx • Kings-Bensonhurst • Kings Bay-Residual Nassau • New York • Queens, • Richmond • Suffolk • Westchester	Kings: • Coney Island	Kings: • Williamsburg • Borough Park • Flatbush
Residual	• Bronx • Kings Nassau • New York • Queens • Richmond • Suffolk • Westchester		

TABLE 2: INCENTIVE BY STRATA

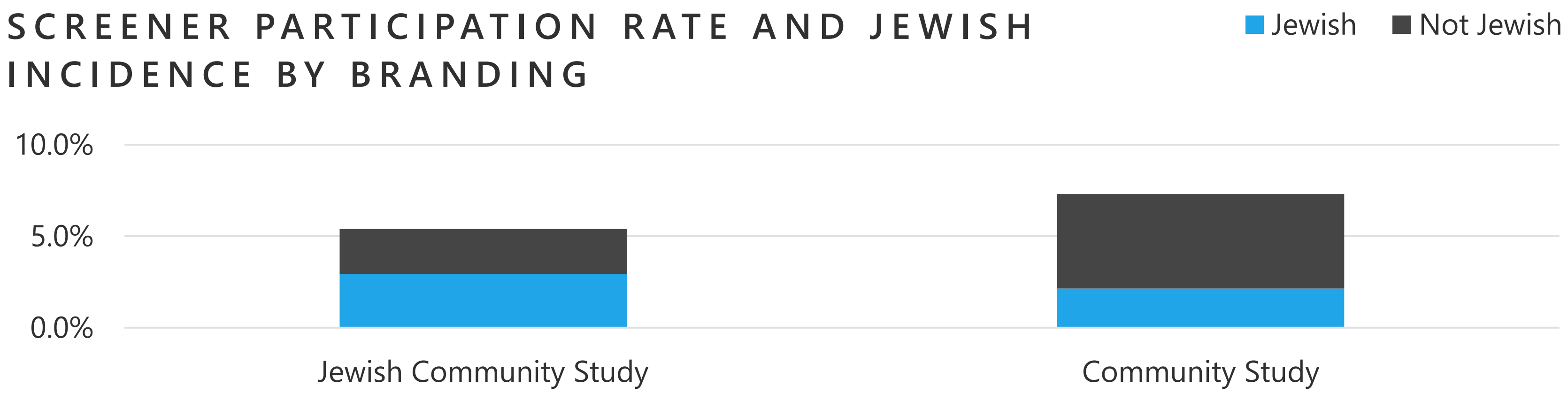
	Non-contingent Incentive	Telephone Outreach
Predicted Jewish	$2	Yes
Residual – High Jewish	$2	No
Residual – Low Jewish	$0	No

TABLE 3: FINAL OUTREACH MATERIALS BY STRATA AND GEOGRAPHY

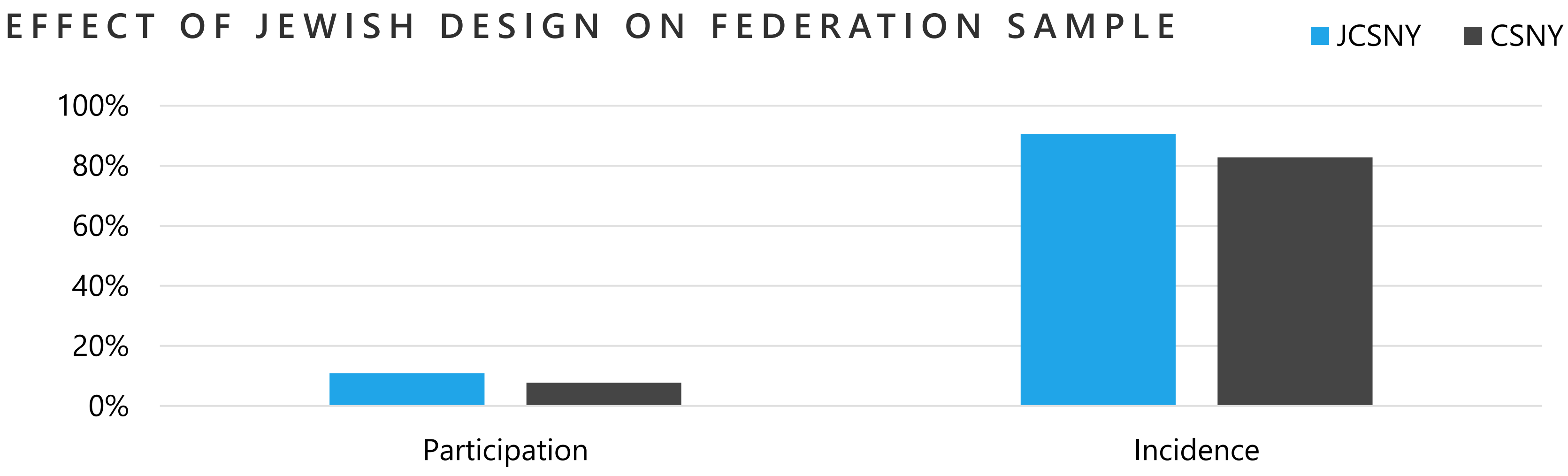
	English Hard Copy	Russian Hard Copy	Yiddish Hard Copy	English Letter
Predicted Jewish	• Bronx • Kings-Bensonhurst • Kings Bay- Residual • Nassau • New York • Queens • Richmond • Suffolk • Westchester	Kings: • Coney Island	Kings: • Williamsburg • Borough Park • Flatbush	
Residual				• Bronx • Kings • Nassau • New York • Queens • Richmond • Suffolk • Westchester

Another Lever: Branding

JEWISH COMMUNITY STUDY of NEW YORK



COMMUNITY STUDY of NEW YORK



Predictive Modeling for Cost Reduction

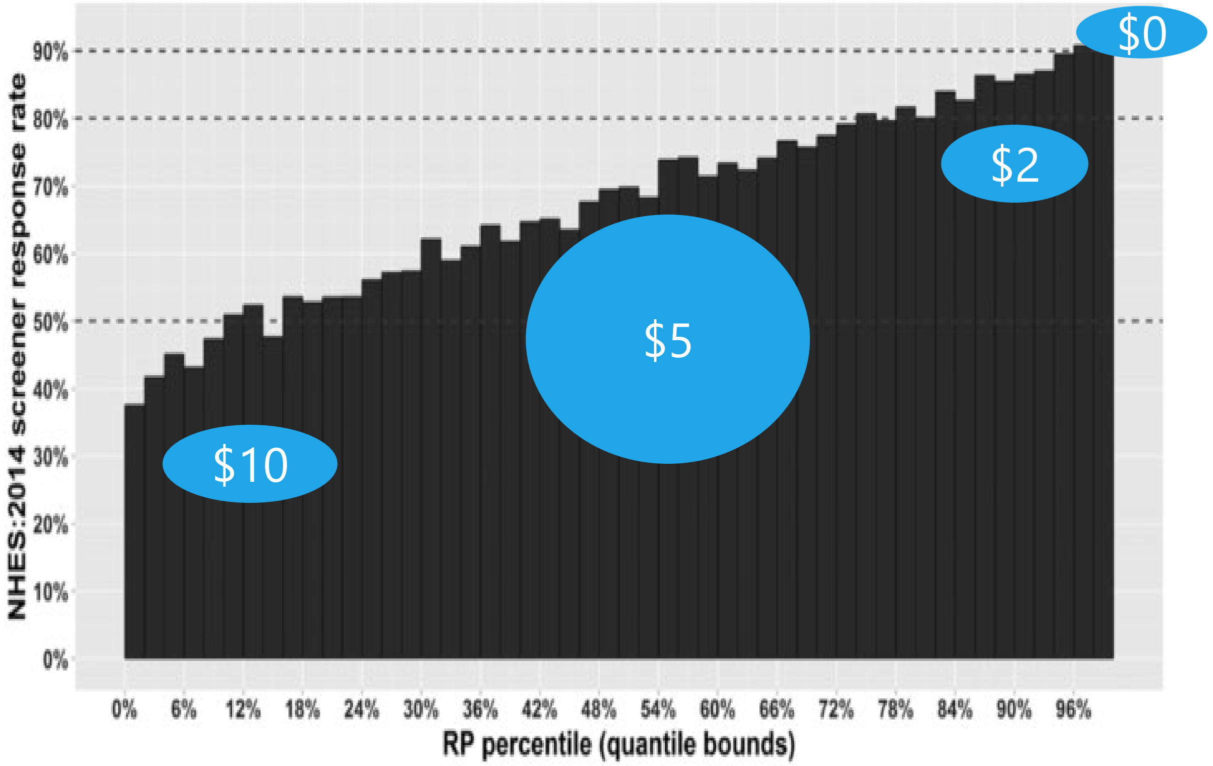
- Responsive/adaptive designs used to manipulate and manage cost and labor allocation is common
- Predictive modeling may be able to enhance or fine-tune those efforts, e.g.,
  - Differential allocation of incentives
  - Strategic use of higher-cost mailings (i.e., FedEx, Priority Mail)
  - Data-driven labor prioritization
- This methodology may provide a mechanism for reducing cost without introducing bias or even decreasing bias

Example: Tailored Use of Higher-cost Treatments

- Research conducted by Jackson, McPhee, and Lavrakas (2020)
- National Household Education Survey (NHES), 2016
  - Cross-sectional, two-phase, ABS survey
  - Majority of data collected via self-administered hard-copy questionnaire*
  - Standard screener incentive since 2012 was $5
- 171,000 addresses randomly assigned to one of 3 treatments
  - $5-only: Received standard $5 non-contingent incentive regardless of predicted response propensity (RP)
  - $2-only: Received $2 non-contingent incentive regardless of predicted RP
  - Tailored Incentive: Received either $10, $5, $2, or $0 assigned based on predicted RP

**Current administrations of the NHES use push-to-web methodology first. The 2016 administration piloted web-push*

Predictive Model



- Data: NHES Feasibility Study, 2014
  - Split into training & test samples
- Model: Logistic Regression
- Dependent variable: Screener response
- Independent variables:
  - address characteristics available on CDS frame
  - address-level demographics appended to the sample file by ABS vendor
  - geographic indicators
  - demographic estimates from the ACS, appended from the Census Planning Database (incl. the LRS [Erdman and Bates 2017])

Screener Response Rates, by Incentive Condition and RP Cohort

RP Cohort	Incentive Condition			*t* statistics	
	Tailored	$5-Only	$2-Only	$5-only vs. tailored	$2-only vs. tailored
Low RP	42%	40%	38%	-1.7	-3.0*
Medium RP	61%	61%	57%	-0.3	-4.7*
High RP	77%	79%	77%	4.2*	0.6
Very high RP	82%	87%	87%	5.2*	2.8*
Total	**64%**	**65%**	**62%**	**2.1***	**-3.8***

Incentive Cost per Complete, by Incentive Condition and RP Cohort

RP Cohort	Incentive Condition		
	Tailored	$5-Only	$2-Only
Overall	$7.80	$8.07	$3.38
Low RP	$25.43	$13.16	$5.64
Medium RP	$8.42	$8.45	$3.58
High RP	$2.62	$6.31	$2.59
Very high RP	$0.00	$5.67	$2.29

Other Levers

MAIL TYPE (FEDEX/ PRIORITY MAIL)

ENVELOPE SIZE/ COLOR/ QUALITY

INCLUSION OF PAPER QUESTIONNAIRES

POST-INCENTIVES

TELEPHONE (OR IN-PERSON) FOLLOW-UP

EXTRA MAILINGS

Sensitivity Modeling

- Tailored interventions based on response propensity have no guarantee that lower-RP cases will be sensitive to the intervention
- Sensitivity modeling attempts to predict a sample unit's increased likelihood to respond contingent on receiving a particular treatment
- Ability to predict sensitivity may add precision to targeting

Example: Response Mode Sensitivity

- Research conducted by Jackson, Medway, and Megra (2021)
- National Household Education Survey (NHES), 2016 and 2019 waves
  - NHES:2016 included a randomized push-to-web experiment, remaining sample received mailed, self-administered hard-copy (training data)
  - NHES:2019 Standard protocol is mail push-to web, followed by self-administered hard-copy option (test data)
- Focused on two research questions:
  1. Can auxiliary data available on ABS frames identify addresses whose likelihood of response is sensitive to whether they are initially offered web or paper?
  2. Does tailoring the offered mode based on predicted mode-sensitivity improve response rates, data collection costs, and/or nonresponse bias in a household-level survey?

Treatment Sensitivity Model

$$\ln\left(\frac{P[R=1]}{1-P[R=1]}\right) = \beta_0 + \alpha T + \sum_k \beta_k x_k + \sum_k \gamma_k x_k T$$

- *R* is an indicator for phase 1 response (regardless of treatment)
- β_0 is a constant
- *T* is a treatment indicator
- and the x_ks are a subset of the candidate covariates (final covariates selected from random forest prediction model)
- All predictor variables are <u>interacted with treatment indicator</u>
- After estimating logistic regression, response propensity contingent on the treatment ($\hat{\rho}_t$) is predicted by setting *T* = 1 and applying coefficients and similarly the control condition applies *T* = 0.

Treatment Sensitivity Model…continued

$$\hat{S}_{t,c} = \hat{\rho}_t - \hat{\rho}_c$$

- $\hat{S}_{t,c}$ is the predicted sensitivity or treatment effect (Lamont et. al., 2016)
- $\hat{\rho}_t$ is the predicted RP score conditional on receiving the treatment protocol
- $\hat{\rho}_c$ is the predicted RP score conditional on receiving the control protocol

NHES MODE SENSITIVITY

- Dependent variable = response to the screener
  - Two models: any response & **early response**
- Independent variables similar to incentive model
  - Added tract-level Federal Communications Commission (FCC) estimates of residential Internet penetration

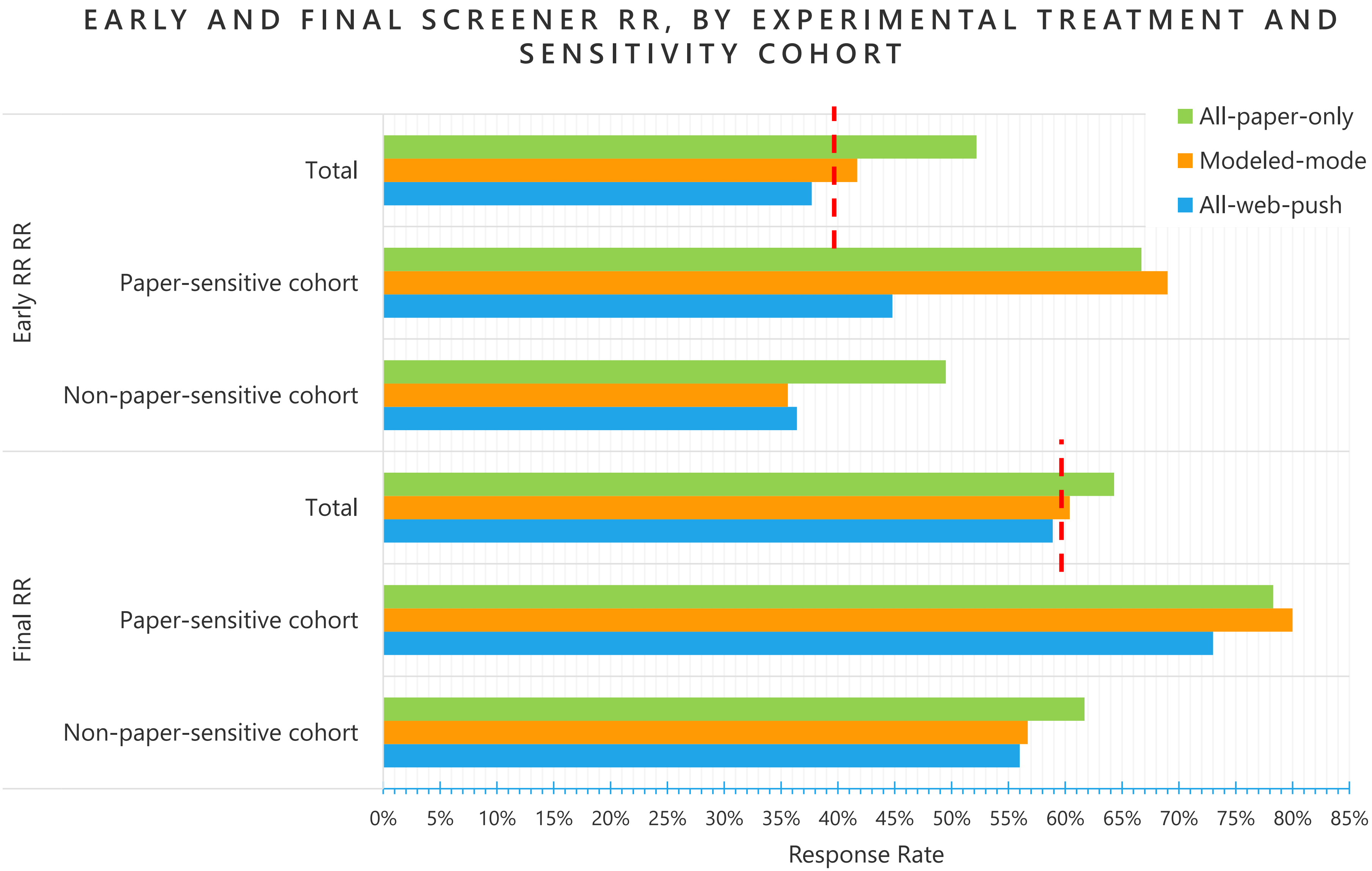
Experimental Design

- “All-web-push” condition received 2 mailings web-push, 2 with hard-copy (HC)
- “All-paper” condition received 4 mailings with self-administered HC, no web
- “Modeled mode” addresses with predicted mode-sensitivity scores above the 85th percentile (the “paper-sensitive cohort”) received the paper-only protocol. The remaining addresses (the “non-paper-sensitive cohort”) received the web-push protocol
  - Most addresses had a positive sensitivity to paper-only, so the “paper sensitive” cohort are those with a particularly large sensitivity

Results

Strongest predictors of mode sensitivity:

- Internet penetration (proxy for access)
- Educational attainment
- Age (person 65+ at the address)

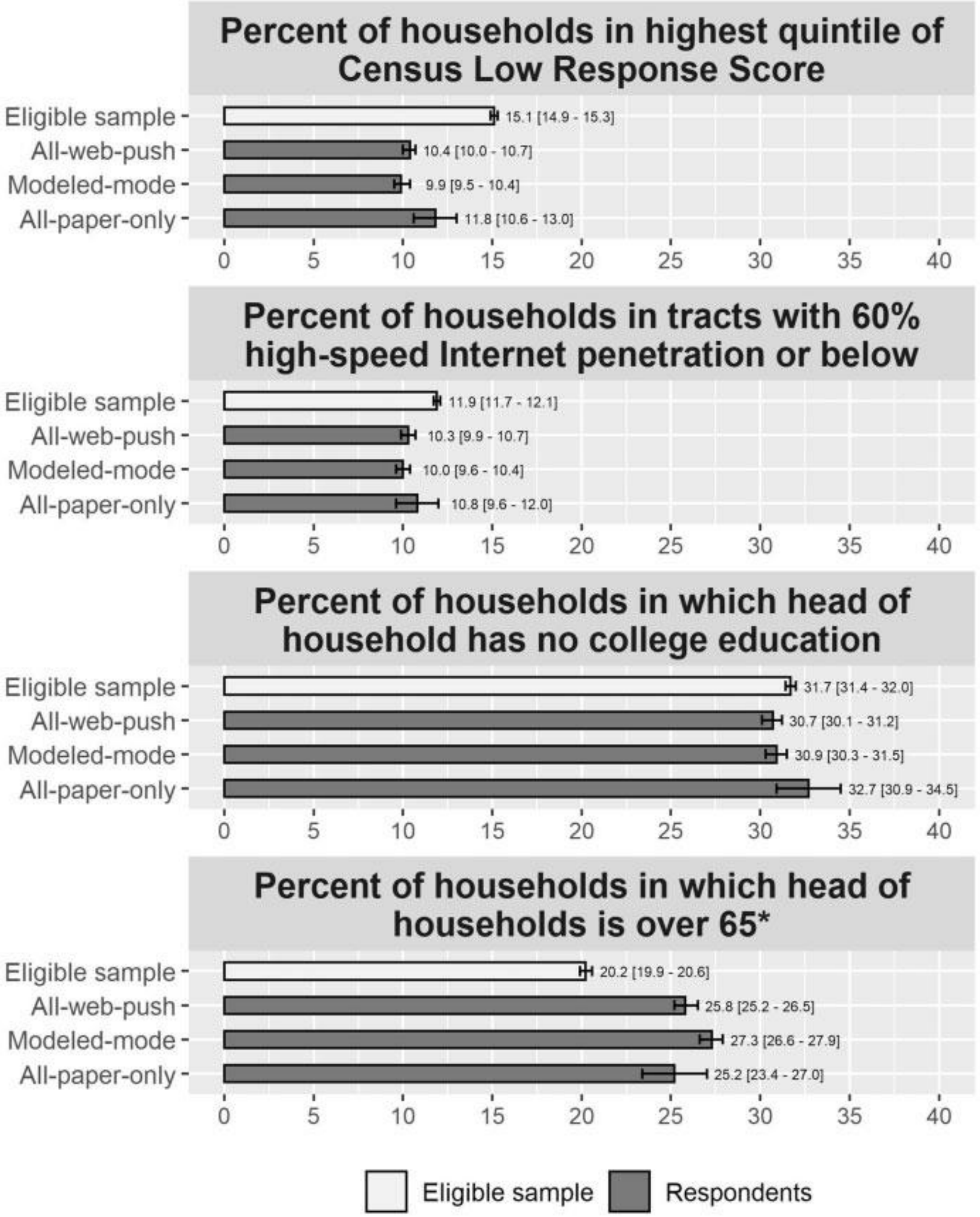


Estimated Cost Measures by Treatment

	TREATMENT		
	All-web-push	Modeled-mode	All-paper-only
Total First-class mailings per response	5.9	5.7	5.1
Total FedEx mailings per response	1.2	1.1	0.9
Percent of First-class mailings with HC	15.8%	22.5%	67.1%
Percent responding by web	59.2%	48.3%	0.0%
Percent responding by inbound phone	7.3%	5.0%	1.1%

Monitor Representation

- Modeling for operational efficiency ideally has minimal impact on sample composition
- Responsive and adaptive interventions should be carefully monitored to avoid introducing bias
  - Distribution of frame data for eligible sample compared to respondents
  - Distribution of benchmarkable variables (e.g., demographics)
  - Use of R-indicators to monitor balance



Applications of Predictive Models for Weighting

- Predictive modeling is not a new tool for weighting, e.g.,
  - Response propensity adjustments for nonresponse (Valliant, Dever, and Kreuter, 2018; Little, 2021)
  - Propensity modeling for weighting non-probability surveys
  - Propensity weighting for "pseudo-RCT" analyses
- Application here is the use of predictive models to identify potential weighting classes in order to account for disproportionate nonresponse related to **key substantive characteristics**

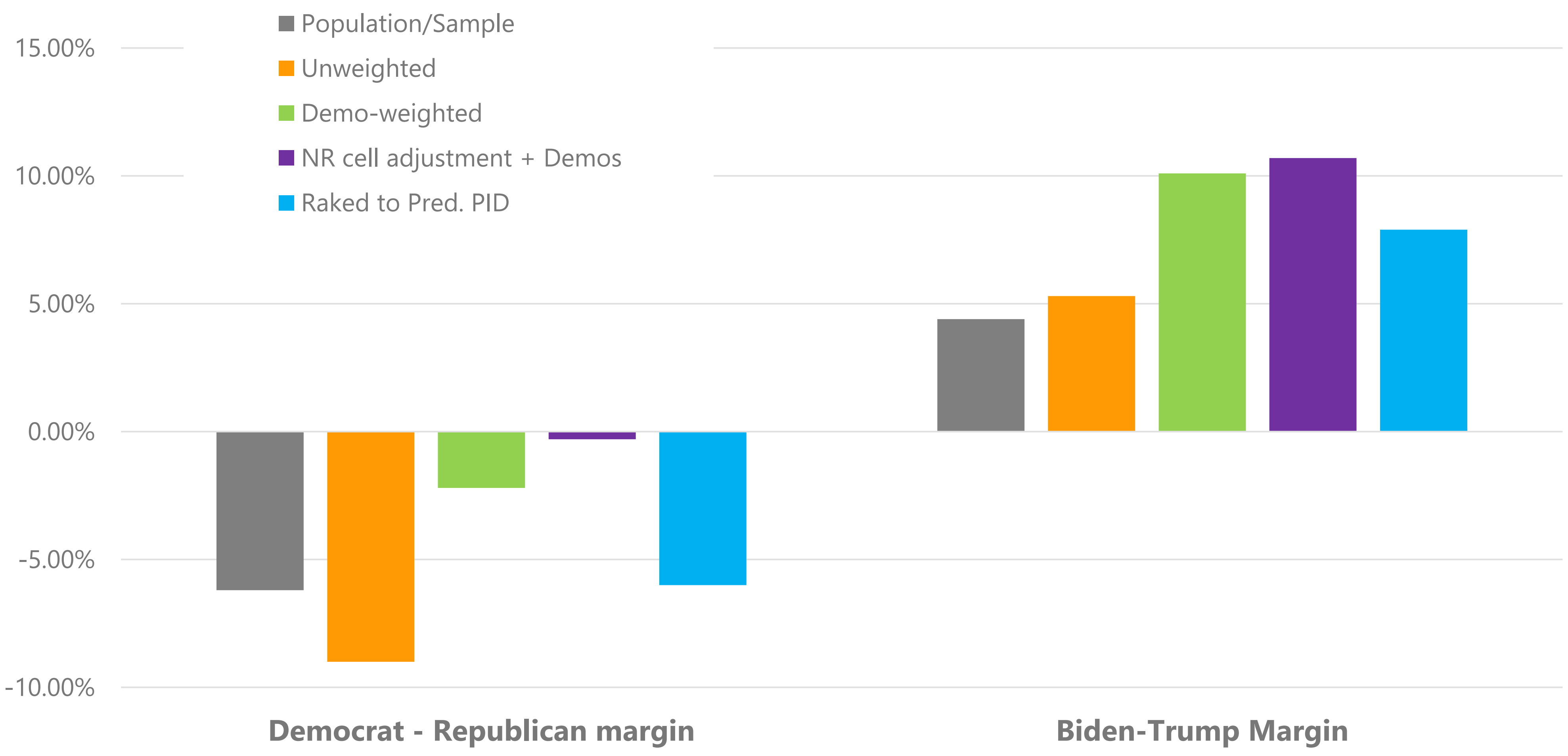
Example: Nonresponse Weighting on Predicted Political Party Identification

- 2021 ABS poll (N=10,000) conducted by the Washington Post & George Mason University to measure political attitudes, party identification, and opinions on current events
- Main purpose was to evaluate the use of predictive modeling to estimate party identification (PID) for sampled addresses prior to data collection
- Predictive model built on respondent data from unrelated large national ABS study
- Dependent variable = PID
- Independent variables appended from voter file and consumer data matched to the address including PID where available and modeled political ideology from consumer database

Partisan Nonresponse?

- Availability of predicted PID for much of the sample (not just respondents) allowed for an assessment of potential differential nonresponse by PID
- Nonresponse adjusted weights were computed and compared to demographic adjustments alone to evaluate efficacy of including predicted partisanship in weighting
- Two weighting options:
  - Nonresponse adjustments within predicted PID groups
  - Sample raking to predicted PID

Effects of Sample-based Party Weighting on Partisan Composition of Respondents and Recalled Vote Distribution



Summary and Considerations

ssrs
research. refined.

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Applications of Predictive Models to Survey Research

- Predictive modeling can be a flexible tool for enhancing and refining survey methods
- Applications for ABS include (but surely not limited to):
  - Targeted sampling
  - Material customization
  - Incentive allocation
  - Mode assignment
  - Treatment sensitivity
  - Weighting

- Lots of other applications to consider (especially outside ABS)
  - E.g., model-based case prioritization/effort allocation
- Should continue refining model estimation and testing alternative applications

Considerations

- Model building and refinement can be labor intensive
  - Customization and tailoring treatments requires additional labor to execute
  - Incremental benefits may not be worth the additional cost
- Ideally, tailored designs would enhance representation and minimize bias by offering respondents most salient materials, incentives, and protocols
- However, using models, especially to strategically limit costs may impact sample representativeness
  - May be able to use these models to reduce cost while minimizing the impact on bias
  - Should be monitored closely throughout data collections
- Additional research is needed into the effectiveness of sensitivity models and robustness of models to changes in study population and operational methods

Q & A

ssrs
research. refined.

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

References

ssrs
research. refined.

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

- Agiesta, J., et. al. (2020) *Task Force on 2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls,* American Association for Public Opinion Research (AAPOR).
- Dillman, D. A., J. D. Smyth, and L. M. Christian (2014). *Internet, Phone, Mail, and Mixed Mode Surveys: The Tailored Design Method,* Hoboken, NJ: Wiley & Sons Inc.
- Groves, R. M., and S. Heeringa (2006), "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 169, 439–457.
- Groves, R. M.,et. al., (2005) *Plan and Operation of Cycle 6 of the National Survey of Family Growth*. Hyattsville: National Center for Health Statistics.
- Groves, R. M., E. Singer, and A. Corning (2000). "Leverage-Saliency Theory of Survey Participation," *Public Opinion Quarterly*, 64, 299–308.
- Jackson, M.T., McPhee, C.B., and Lavrakas, P.J., "Using Response Propensity Modeling to Allocate Noncontingent Incentives in an Address-Based Sample: Evidence from a National Experiment", *Journal of Survey Statistics and Methodology*, Volume 8, Issue 2, April 2020, Pages 385–411.
- Jackson, M.T., Medway, R.L., Megra, M.W., "Can Appended Auxiliary Data Be Used to Tailor the Offered Response Mode in Cross-Sectional Studies? Evidence from an Address-Based Sample", *Journal of Survey Statistics and Methodology*, 2021.
- Lamont, A., M. D. Lyons, T. Jaki, E. Stuart, D. J. Feaster, K. Tharmaratnam, D. Oberski, H. Ishwaran, D. K. Wilson, and M. L. V. Horn (2016). "Identification of Predicted Individual Treatment Effects in Randomized Clinical Trials," *Statistical Methods in Medical Research*, 27, 142–157.
- Lavrakas, P.J., Jackson, M. and McPhee, C. (2018). "The Use of Response Propensity Modeling (RPM) for Allocating Differential Survey Recruitment Strategies: Purpose, Rationale, and Implementation." *Survey Practice* 11 (2).
- Link, M. W., and A. T. Burks (2013). "Leveraging Auxiliary Data, Differential Incentives, and Survey Mode to Target Hard-to-Reach Groups in an Address-Based Sample Design," *Public Opinion Quarterly*, 77, 696–713.
- Little, R.J. (2021) "A Note about the Definition of Response Propensity for Survey Nonresponse," *Journal of Survey Statistics and Methodology*.
- Mercer, A., Lau, A., and Kennedy, C.. 2018. "For Weighting Online Opt-In Samples, What Matters Most?" Pew Research Center. https://www.pewresearch.org/methods/2018/01/26/ for-weighting-online-opt-in-samples-what-matters-most
- Olson, K., et al. (2019). *Report of the AAPOR Task Force on Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys*, American Association for Public Opinion Research (AAPOR).
- Olson, K., J. D. Smyth, and H. M. Wood (2012). "Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates? An Experimental Examination," *Public Opinion Quarterly*, 76, 611–635.
- Peytchev, A., R. K. Baxter, and L. R. Carley-Baxter (2009), "Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error," *Public Opinion Quarterly*, 73, 785–806.
- Valliant, R, Dever, J.A., and Kreuter, F. 2018. *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer.
- Wagner, J. (2013), "Adaptive Contact Strategies in Telephone and Face-to-Face Surveys," *Survey Research Methods*, 7, 45–55.
- West, B., J. Wagner, F. Hubbard, and H. Gu (2015), "The Utility of Alternative Commercial Data Sources for Survey Operations and Estimation: Evidence from the National Survey of Family Growth," *Journal of Survey Statistics and Methodology*, 3, 240–264.

Cameron McPhee

CHIEF METHODOLOGIST

cmcphee@ssrs.com

@cameronbrook

ssrs

research. refined.

AAPOR

AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH