



■

Assessing the Effectiveness of Using Propensity Scores to Target Small and Hard-to-Reach Populations

MARKET
DECISIONS
RESEARCH

Better Data
Better Insights
Better Outcomes

THE 12TH HEALTH SURVEY RESEARCH CONFERENCE 2025

Brian Robertson, PhD, VP of Research
John Charles, MS, Senior Analyst
Katie Klinko, MS Research Analyst

Acknowledgement

Market Decisions Research would like to thank the Vermont Department of Health and HealthSource RI for allowing us to use their survey data in this presentation.



Our Presentation



Project Background

The Challenge

Our Methodology and Approach

Propensity Scoring

Adventures in Data Collection

Analysis and Results

Conclusions and Future Research

Prior Methods Background

Traditional media/press release recruitment

- Widely known and reputable distributors with large audiences
- Competition for public's attention with other news stories and can be high cost to maintain frequent advertising

Online/social media/web-based recruitment

- Large audience, low-cost, efficient, and flexible
- Tradeoff between cost-effectiveness of targeted advertisements and representativeness of sample obtained

Incentives/paid recruitment

- Increases response rates while also increasing costs
- Higher potential for fraud and/or deceitful responses when provided online

Multi-mode recruitment

- Direct organizational outreach is combined with other targeted recruitment methods (push-to-web mail, and phone outreach, and incentivized online panels, etc.) to maximize responses
- Can become high-cost when utilizing several methods of outreach

Prior Research Background

- Lee et al. (2022) utilized propensity score matching to examine whether hospitalization experiences affect older adults' health and daily life engagement. They matched each case to its nearest and most applicable control case on the variables of age, sex, education, income, region, and living status.
- Ferri-García and Rueda (2020) compare propensity estimation by logistic regression against propensity estimation by machine learning (ML) algorithms. They found using ML for propensity score adjustments removed selection bias more effectively than standard logistic regression; however, the ML method is limited by selection mechanisms and data dimensionality.
- DuGoff, Schuler, and Stuart (2014) used 2008 Medical Expenditure Panel Survey data to explore various propensity score methods in estimating treatment effects. They found propensity score models which accounted for survey weights achieved more unbiased estimates of treatment effects, which are then able to be generalized to the original overall population.

Understanding Propensity Scoring

- Propensity scoring is a statistical technique used to estimate the likelihood of an individual having a specific characteristic or outcome based on observed variables. It is commonly applied in:
 - Observational studies to control for confounding by balancing groups.
 - Survey research to adjust for nonresponse bias or improve representativeness.
 - Targeted interventions, such as prioritizing specific populations for outreach.
- In the context of our approach, propensity scores are used to identify and flag cases in a sample that are likely part of hard-to-reach populations, such as smokers or uninsured individuals. This approach allows us to focus survey efforts on individuals who are more likely to exhibit the desired characteristics.

Why Use Propensity Scores for Targeting?

- Hard-to-reach populations often exhibit low response rates or are underrepresented in traditional sampling frames. This creates challenges in obtaining accurate and reliable data. Propensity scores address these issues by:
 - **Prioritizing Efforts:** Flagging individuals who are more likely to belong to the population of interest. Resources are directed toward individuals with the highest likelihood of meeting target criteria.
 - **Maximizing Efficiency and Cost-Effectiveness:** Reducing outreach and resources spent on individuals unlikely to meet the target criteria.
 - **Improving Representation:** Enhancing representation of key populations, resulting in more accurate survey results.
- By leveraging available demographic data and predictive modeling, propensity scores provide a systematic and scalable method for targeting.

Estimating Propensity Scores

- To estimate propensity scores, we use a logistic regression model and data from a reference or source population dataset (e.g., BRFSS, ACS, etc.). The probability of an outcome $P(Y = 1|X)$ is predicted based on covariates X , including demographic and socioeconomic variables k . **The propensity score formula is:**

$$P(Y = 1|X) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}}$$

Where:

- $Y = 1$ is the outcome of interest (e.g., being a smoker or uninsured).
- X_1, X_2, \dots, X_k are the predictor variables (demographics, socioeconomic factors).
- β_0 is the intercept, and $\beta_1, \beta_2, \dots, \beta_k$ are the coefficients.

Estimating Propensity Scores

Variable Selection:

- To identify the most predictive variables for estimating propensity scores, a systematic approach combining correlation analysis and decision tree modeling was employed in the reference dataset.
- 1. **Correlation analysis** was conducted to evaluate the relationship between each potential predictor variable (X) and the binary outcome (Y) of interest.
- 2. **Decision tree models** were used to identify key predictors and their interaction effects. A decision tree iteratively splits the data based on thresholds in predictor variables to maximize the purity of outcome groups ($Y=1$ or $Y=0$). The process involved:
 - Recursive Partitioning: Splitting the data at nodes to minimize impurity.
 - Variable Importance: Ranking predictors by their contribution to reducing impurity.

Estimating Propensity Scores

Final Predictor Selection:

- Predictor variables included in the logistic regression model were those that:
 - Showed significant correlation or association with the outcome ($p < 0.05$).
 - Ranked highly in decision tree variable importance scores.
 - Demonstrated logical and practical relevance to the outcome.
- For example, age and education emerged as the strongest predictors for identifying uninsured individuals, while gender, age, education and income levels were key for predicting smokers. This combined approach ensured that the selected variables maximized predictive accuracy and interpretability.

Estimating Propensity Scores

Fitting the Model:

- IBM SPSS Statistics was used to run this analysis using the NOMREG procedure. NOMREG is used to model relationships between a categorical dependent variable with two or more categories (nominal or ordinal) and one or more independent variables (predictors).
- The NOMREG procedure uses the maximum likelihood estimation (MLE) method to estimate the regression coefficients (β). MLE maximizes the likelihood function:

$$L(\beta) = \prod_{i=1}^n (P(Y_i = 1|X_i))^{Y_i} (1 - P(Y_i = 1|X_i))^{1-Y_i}$$

- Goodness-of-fit was assessed using pseudo- R^2 .
- The next table shows an example of the final output of this analysis.

Propensity Score Table Example

Gender	Age Categories	Income	Education Level	Marital Status	Children in household	Predicted Probability of Outcome of Interest	Priority
Male	18 to 34	Less than \$10K	Less than high school	Not married	Yes	0.701	High
Female	18 to 34	\$20K-\$25K	Less than high school	Not married	Yes	0.616	High
Male	35 to 44	\$50K-\$75K	High school diploma	Not married	No	0.522	High
Male	35 to 44	\$75K	High school diploma	Not married	Yes	0.382	High
Male	45 to 54	\$35K-\$50K	Less than high school	Married	No	0.315	High
Male	45 to 54	\$35K-\$50K	Some college	Married	No	0.309	High
Female	18 to 34	\$50K-\$75K	College degree	Not married	Yes	0.098	Medium
Female	65 or older	\$75K	College degree	Married	Yes	0.013	Low
Female	65 or older	\$75K	College degree	Married	Yes	0.008	Low

Final Output:

The final output of the analysis is a table of propensity scores assigned to each unique combination of predictor variable categories.

Estimating Propensity Scores

Applying Propensity Scores to the Sample File:

- Once propensity scores were estimated, the next step was to transfer these probabilities to the sample file:
 - **Matching:** Aligning propensity scores from the source population dataset with sample cases based on shared demographic variables.
 - **Flagging Cases:** Assigning a flag to cases in the sample with high scores of having the target characteristic.
 - **Thresholding:** Defining probability thresholds to categorize cases as high priority or low priority.
- For example, cases over a desired propensity score threshold for smoking were flagged for targeted outreach.

Workflow Overview

- The implementation process involved four main steps:
 - **Data Preparation:** Standardizing the source population data and sample datasets to ensure comparability.
 - **Propensity Estimation:** Calculating probabilities for each outcome of interest using logistic regression.
 - **Probability Transfer:** Matching and transferring probabilities to the sample file.
 - **Case Flagging:** Identifying high-priority cases using predefined thresholds.
- This workflow ensured that the most relevant cases were flagged for targeted survey efforts.

Test Case

The Vermont Adult Tobacco Survey

The Vermont Adult Tobacco Survey is a survey of Vermont adults.

Data are used to evaluate the effectiveness of Vermont Tobacco Control Program efforts to reduce smoking and increase awareness and knowledge of smoking-related issues.

Summary

Vermont has conducted the Adult Tobacco Survey annually from 2001 to 2008, every other year from 2010 to 2016, and in 2022.

The goal is to complete 1,600 to 2,000 respondents each year. Half with Vermonters who smoke, and half with Vermonters who do not smoke.

The Challenge

A primary focus is to gather data from those using or who have recently used tobacco products: How do you find and survey current smokers and e-cigarette user and recent quitters?

First, these groups represent a small percentage of the population.

Group	% of Vermont Residents (2020 BRFSS)
Current Smokers	12.6%
Current E-Cigarette Users	3.7%
Recent Quitters (Quit <5 years ago)	7.2%

Second, members of these groups are generally less likely to respond to surveys on tobacco use, especially one administered by **a Department of Health**

Methods

The Sampling Frame

A Random Address Base Sample (ABS) of Vermont addresses

An ABS sample targeting Vermont households more likely to include smokers using information about lifestyles, characteristics, and behaviors of the household residents to predict the likely presence of smoking behavior.

An online survey panel of Vermont residents.

Additionally, sample was obtained through the distribution of public advertisements targeted towards current or recent tobacco and/or e-cigarette users.

Methods

Data Collection

Data collection included an online survey in combination with a mailed survey

- Letter targeting sample of randomly selected Vermont households
- Survey booklet mailing to households that were less likely to have access to the internet
- **Letter targeting sample of Vermont households more likely to include smokers based on propensity scoring**
- Email invitations to smokers that recently participated in a research focus group.
- Public Advertising and Flyers
- Online Panel
- Final Follow-ups Email invitations

2888 completed surveys including 696 current and recent tobacco users

Propensity Scoring

Calculating Propensity Score

2020 Vermont Behavioral Risk Factor Surveillance Survey (BRFSS) data was used to estimate the propensity of being a tobacco user based on geodemographics

- Age
- Gender
- Income
- Education Level
- Rurality
- Housing Status
- Marital Status
- Number of Children in Household

Propensity Scoring

Calculating Propensity Score

- The regression model computed a propensity score (predicted probability) of being a tobacco user
- Higher scores were associated with a higher likelihood that person was a current tobacco user
- Once calculated from the BRFSS, the scores were appended to records in the sample file for the Vermont ATS by matching geodemographics

Propensity Score Table

Gender	Age Categories	Income	Education Level	Rurality	Housing Status	Marital Status	Number of children in household	Predicted Probability of Smoking
Male	35 to 44	Less than \$10K	Less than high school	Urban	Renter	Not married	2 children	0.689
Female	18 to 34	\$20K-\$25K	Less than high school	Rural	Renter	Not married	1 child	0.566
Male	35 to 44	\$50K-\$75K	High school diploma	Rural	Renter	Not married	2 children	0.482
Male	35 to 44	\$75K	High school diploma	Rural	Owner	Not married	1 child	0.315
Male	45 to 54	\$35K-\$50K	Less than high school	Urban	Owner	Married	2 children	0.315
Male	45 to 54	\$35K-\$50K	Some college	Rural	Renter	Married	None	0.259
Female	18 to 34	\$50K-\$75K	Some college	Rural	Renter	Not married	1 child	0.259
Female	65 or older	\$75K	College degree	Urban	Owner	Married	2 children	0.013
Male	65 or older	\$75K	College degree	Urban	Owner	Married	None	0.012
Female	65 or older	\$75K	College degree	Urban	Owner	Married	None	0.011

Analysis and Results

How well do propensity scores work in identifying smokers/e-cigarette users and recent quitters?

Again, looking only at the surveys completed during the initial push to web and mail surveys:

- We ran a simple crosstabulation of completed surveys
 - Smoking/e-cigarette use status by propensity score category

And it did show that using propensity scores could be an effective method to identify current tobacco users and recent quitters

Analysis and Results

Summary of Results Broken out by Propensity Score

Smoking and E-Cigarette Use Status					
Propensity Score	Never Used	Formerly Used	Recently Quit (<5 yrs. ago)	Currently Use	% Current Users and Recent Quitters
Less than .1	63%	30%	4%	3%	7%
.1-.19	60%	30%	4%	6%	10%
.2-.29	49%	35%	10%	7%	16%
.3 or greater	58%	23%	4%	15%	19%
Total	60%	30%	4%	5%	9%

The percentage of current smokers/e-cigarette users increased with increasing propensity scores

Analysis and Results

Study Limitation

We lost the ability to fully evaluate the effectiveness of propensity scores as a tool to target current users and recent quitters:

- The scores weren't used as a criteria for selection or targeting of respondents except for one mailing.

But because we appended propensity scores to the ABS probability sample, we could at least evaluate the effectiveness of propensity scores during the initial push-to-web and mail surveys

Based on the results, we developed a more formal test of using propensity scores to reach (an even more challenging) group!

Test Case

The Rhode Island Health Information Survey

The Rhode Island Health Information is A survey of Rhode Island Households and Gather information on all people residing in the household.

The survey aims to provide the state with a better understanding of Rhode Islanders' health coverage and health insurance needs. Answers to the survey will help to guide state policy and programs related to health care.

Summary

The state of Rhode Island and HealthSource RI have conducted the Health Information Survey bi-annually from 2012 to 2024.

The target Completes varied year to year; 5600 in 2012 to 2016. 5000 in 2018 and 3000 in 2020 to 2024

The Challenge

A primary focus is to gather data from those who currently lack any type of health insurance: How do you find and survey those that lack insurance?

First, these groups represent a small percentage of the population.

Group	% of Rhode Island Residents (2022 RI HIS)
Uninsured	2.9%

Second, uninsured residents tend to be lower income, less educated, and non-white; groups who are less likely to respond to surveys.

Test Case

The Sampling Frame

- The sample design relied on a dual frame listed landline and Random Digit Dialing (RDD) cell phone sampling methodology.
- A statewide sample of listed landline and RDD cell phone numbers was generated at random.
- Address information was appended to each sample record if available
- When available geodemographic information was appended to allow the propensity score analysis

Methods

Data Collection

Data collection for the 2024 RI Health Information Survey used a dual mode data phone and online collection protocol.

- For those without a mailing address, calls began immediately.
 - An initial survey invitation letter was then sent to all households with a mailing address in the sample. Invitation letters were sent out weekly
 - After two weeks, telephone follow-up calls began to households sent the invitation letter. During the calls, respondents were given the option to complete the survey by telephone or to complete the survey online.
 - Once all invitation letters had been sent and attempts were made by telephone, an email invitation was sent to any households that had not previously completed the survey for which an email was available.
 - To increase survey response among key groups (the uninsured and race/ethnic minority groups), a second invitation letter was sent to select non-respondents based on the likelihood the household included uninsured members..
-
- 3,256 Rhode Island Households completed the survey

Propensity Scoring

Calculating Propensity Score

2018 - 2022 Rhode Island Health Information Survey data was used to estimate the propensity a household having a member that was uninsured user based on geodemographics

- Zip Code
- Age of resident
- Annual household Income
- Number of children in the household
- Gender
- Martial Status
- Ethnic Group
- Level of Education
- Use of prepaid cell phone

Propensity Scoring

Calculating Propensity Score

- The regression model computed a propensity score (predicted probability) of a household having one or more residents without health insurance.
- Higher scores were associated with a higher likelihood that there was an uninsured member of the household
- Once calculated from the RI HIS, the scores were appended to records in the sample file for the by matching geodemographics (if available).
- Sample records were assigned a likelihood based on propensity score (low, medium, or high likelihood)
- For this test we appended propensity scores PRIOR to data collection
- However, we did not use the scores to target cases (with one exception) but rather to provide data for analysis

Analysis and Results

How well do propensity scores work in identifying uninsured residents?

In this case we saw a significant increase in the percentage of uninsured residents in the medium and high likelihood groups when compared to the low likelihood group

While the percentages are small it did improve efficiency in identifying uninsured residents by 100% when comparing the medium and high likelihood groups to the low likelihood group.

Analysis and Results

Summary of Results Broken out by Propensity Score

	Propensity Scoring for Presence of Uninsured			
	Not assigned	Low Probability of Household with Uninsured	Medium Probability of Household with Uninsured	High Probability of Household with Uninsured
Type of Insurance	Count	Count	Count	Count
Private health insurance	155	4132	617	952
Medicare	195	453	69	187
Rite Care	27	731	135	508
Military, Veterans, or TRICARE	19	175	44	37
Rite Share	0	0	0	3
HealthSource RI	0	470	104	316
Medicare Advantage plan or supplement	0	192	33	39
No Insurance Coverage	1	68	23	44
Total Persons	281	5644	939	1846
	0.4%	1.2%	2.4%	2.4%

Analysis and Results

Study Limitation

Since we were simply testing the effectiveness of the model, we didn't use the score to specifically target households most likely to include uninsured resident except for one small mailing at the end.

The most significant limitation is the information available for sample records which is limited to demographics and geography that can be paired with the data source that develops the model.

But we are currently using this technique for The Vermont Household Health Insurance Survey to target households that are the more likely to include uninsured members using a slightly different mix of geodemographics.

Conclusions and Future Research

How can we improve the process of identifying and surveying hard to reach populations?

- **Identify a sample source with complete (and additional) information for propensity scoring**
 - Provide complete demographic information for all sample records
 - Not all records contained demographic information
 - In many cases sample records had only partial demographic information

Conclusions and Future Research

How can we improve the process of identifying and surveying hard to reach populations?

- **Identify sample sources that provide information on other household factors, behaviors and attitudes**
 - Additional correlates
 - Veteran's status, insurance coverage, physical and mental health status, sexual orientation, substance use
 - Loss of employment, presence of children on Medicaid, interruptions in coverage among household members, cost and other barriers to receiving care, high out of pocket health expenses, physical and mental health status

Conclusions and Future Research

How can we improve the process of identifying and surveying hard to reach populations?

- **Use area based meta data**
 - Rather than focusing on the characteristics of the sample, look at the characteristics of the location of the sample
 - Within each geographic area (county, zip code, census tract) look at factors among the population that would suggest a higher likelihood of a hard-to-reach population
 - This would allow the use multiple data sources and use a full range of demographic, information, behavior, and attitudes data to predict likelihood of belong to a hard to reach population.

For More Information

Brian Robertson, Ph.D.

brianr@marketdecisions.com

John Charles, MS

jcharles@marketdecisions.com

Katie Klinko, MS

kklinko@marketdecisions.com

Thank you!

For More Information

Batterham, Philip J. “Recruitment of Mental Health Survey Participants Using Internet Advertising: Content, Characteristics and Cost Effectiveness.” *International Journal of Methods in Psychiatric Research* 23, no. 2 (2014): 184–91. <https://doi.org/10.1002/mpr.1421>.

DuGoff, Eva H. ; Schuler, Megan ; Stuart, Elizabeth A. “Generalizing Observational Study Results: Applying Propensity Score Methods to Complex Surveys.” *Health Services Research* 49, no. 1 (2014): 284–303. <https://doi.org/10.1111/1475-6773.12090>.

Ferri-García, Ramón ; Rueda, María Del Mar. “Propensity Score Adjustment Using Machine Learning Classification Algorithms to Control Selection Bias in Online Surveys.” *PloS One*, vol. 15, no. 4, 2020, pp. e0231500–e0231500, <https://doi.org/10.1371/journal.pone.0231500>.

Goodrich, Brittney ; Fenton, Marieke ; Penn, Jerrod ; Bovay, John ; Mountain, Travis. “Battling Bots: Experiences and Strategies to Mitigate Fraudulent Responses in Online Surveys.” *Applied Economic Perspectives and Policy*, vol. 45, no. 2, 2023, pp. 762–84, <https://doi.org/10.1002/aepp.13353>.

Lee, Ji Yeon ; Kim, Bora ; Lee, Kyung Hee ; Park, Chang Gi. “Does Hospitalisation Impact the Successful Ageing of Community-dwelling Older Adults?: A Propensity Score Matching Analysis Using the Korean National Survey Data.” *International Journal of Older People Nursing* 17, no. 1 (2022): e12413-n/a. <https://doi.org/10.1111/opn.12413>.

Neundorf, Anja ; Öztürk, Aykut. “How to Improve Representativeness and Cost-Effectiveness in Samples Recruited through Meta: A Comparison of Advertisement Tools.” *PloS One* 18, no. 2 (2023): e0281243-. <https://doi.org/10.1371/journal.pone.0281243>.

Southern, Danielle A ; Lewis, Steven ; Maxwell, Colleen J ; Dunn, James R ; Noseworthy, Tom W ; Corbett, Gail ; Thomas, Karen ; Ghali, William A. “Sampling ‘hard-to-Reach’ Populations in Health Research: Yield from a Study Targeting Americans Living in Canada.” *BMC Medical Research Methodology* 8, no. 1 (2008): 57–57. <https://doi.org/10.1186/1471-2288-8-57>.