

# Address-Based Sampling for Jewish Community Studies

---

**Zachary H. Seeskin, Ph.D., NORC at the University of Chicago**

**David Dutwin, Ph.D., NORC at the University of Chicago**

**Leonard Saxe, Ph.D., Cohen Center for Modern Jewish Studies, Brandeis University**

AAPOR Annual Conference 2022

May 11, 2022

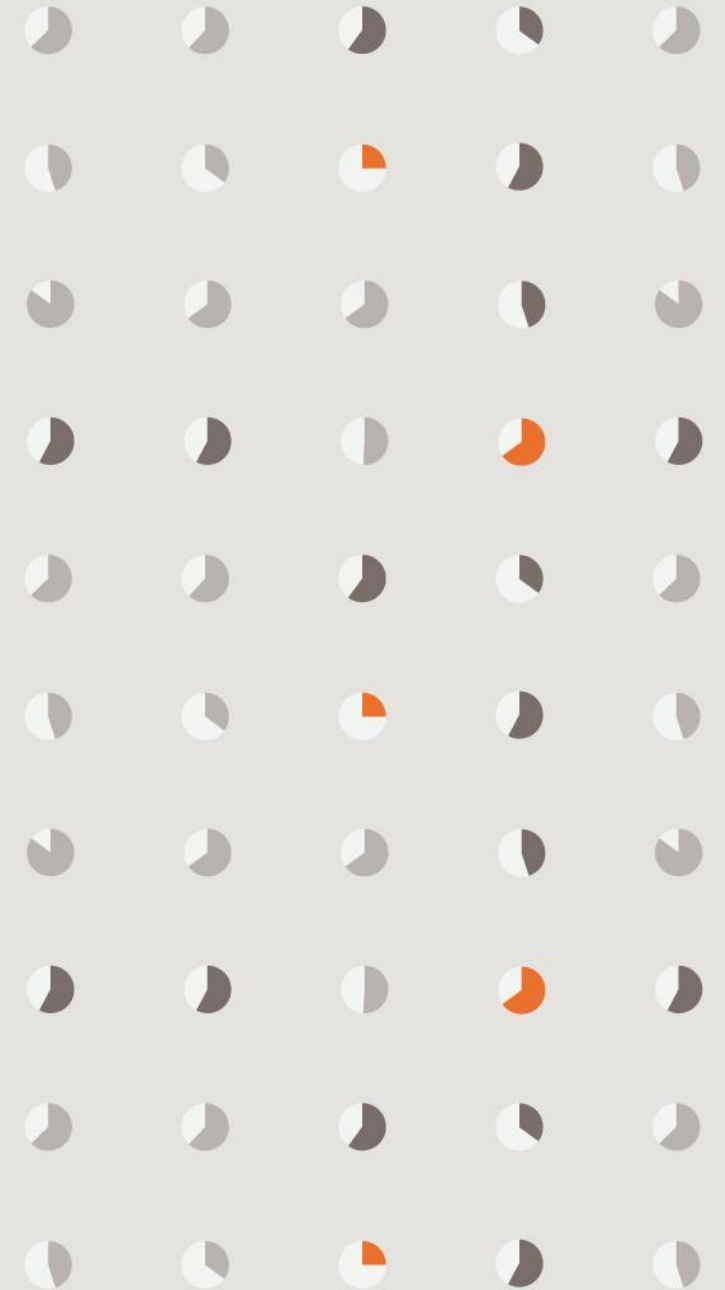
## Acknowledgements

**From NORC:** Stephen Cohen, Ned English, Julie Banks, Evan Herring-Nathan, Chang Zhao, Patrick Coyle, Peter Herman, Katie Archambeau, Andrea Malpica, Becki Curtis, Frankie Duda, Shannon Nelson, Margrethe Montgomery, Steven Pedlow, Michael Yang

**From Brandeis:** Janet Aronson, Matt Boxer, Matthew Feinberg

# Overview

- 1 Motivation
- 2 NORC's Approach to Sampling for Jewish Community Studies
- 3 Methods
- 4 Outcomes
- 5 Concluding Thoughts



Jewish community studies provide estimates of the size and characteristics of Jewish populations

- Essential because the census and federal surveys do not identify individuals by religion
- Studies yield information to plan and assess programs
- Includes focus on subgroups: **Families, Young adults, Denominational groups, Interfaith households, Non-synagogue members, Financially struggling, Jews of Color, LGBTQ Jews, Jews of different ethnic backgrounds, Geographic regions**



## Jewish Chicago: Who We Are

A 2020 POPULATION STUDY



While new to the practice of Jewish community studies, address-based sampling (ABS) provides the strong advantages of known probabilities of sampling and high population coverage of the population at reasonable cost.

- Random digit dialing no longer a gold standard methodology (Lavrakas et al. 2017)
  - Increasing survey costs
  - Population coverage concerns
- Sampling outside of Jewish organization lists and from general population is critical for reaching less engaged members of Jewish community
- Two prominent recent examples:

**2020 Metropolitan Chicago Jewish Population Study**

**2021 Study of Jewish LA**

Fundamental challenge is to survey a rare population and provide reliable estimates on a reasonable budget.

Three strategies were used in large-scale studies of Chicago and Los Angeles:

1. Thorough development of a frame of membership and participant lists from Jewish organizations
2. Stratification of non-organization list households by developing geographic measures of Jewish incidence
3. Use of predictive modeling or vendor data to identify likely Jewish households not on Jewish organization lists

Jewish community study sampling designs segment the sampling frame into three groups of strata. Available data are used to assign area households from a U.S. Postal Service list to a sampling stratum.

Sampling Source	Sampling Frame	Jewish Incidence	Cost per Complete	Jewish Engagement Level
Deduplicated Jewish Organization Lists	Jewish organization participant addresses	Very High	Low	A Range
Remaining Households Predicted as Likely Jewish	USPS Computerized Delivery Sequence File	Moderate	Moderate	Lower
Remaining Households <u>Not</u> Predicted as Likely Jewish, Stratified by Geographic Measures of Jewish Incidence	USPS Computerized Delivery Sequence File	Low	High	Lower

**Note:** 2020 Chicago study used vendor identification of likely Jewish households rather than a predictive model.

Robust Jewish organization list collection is important to ensure cost-effectiveness and to support reaching subpopulations of the Jewish community.

- In-depth undertaking to collect and curate lists from a range of Jewish organizations lists allows for diverse segments of the community to be included in the organization list frame
- Comprehensive organization list gathering and thought for sample design is essential for reaching small subgroups
- 40+ organizations represented for 2020 Chicago study; 60+ organizations represented for 2021 Los Angeles study



Multiple measures related to Jewish incidence developed at the census block group level to stratify geographic areas by likely Jewish incidence.

The percentage of block group residential addresses that are:

1. On organization lists.
2. Identified or predicted as likely Jewish using vendor data.
3. Have surnames from vendor data identified as distinctive Jewish names.

Majority of households have a non-zero probability of being sampled, while we oversample high density Jewish block groups.

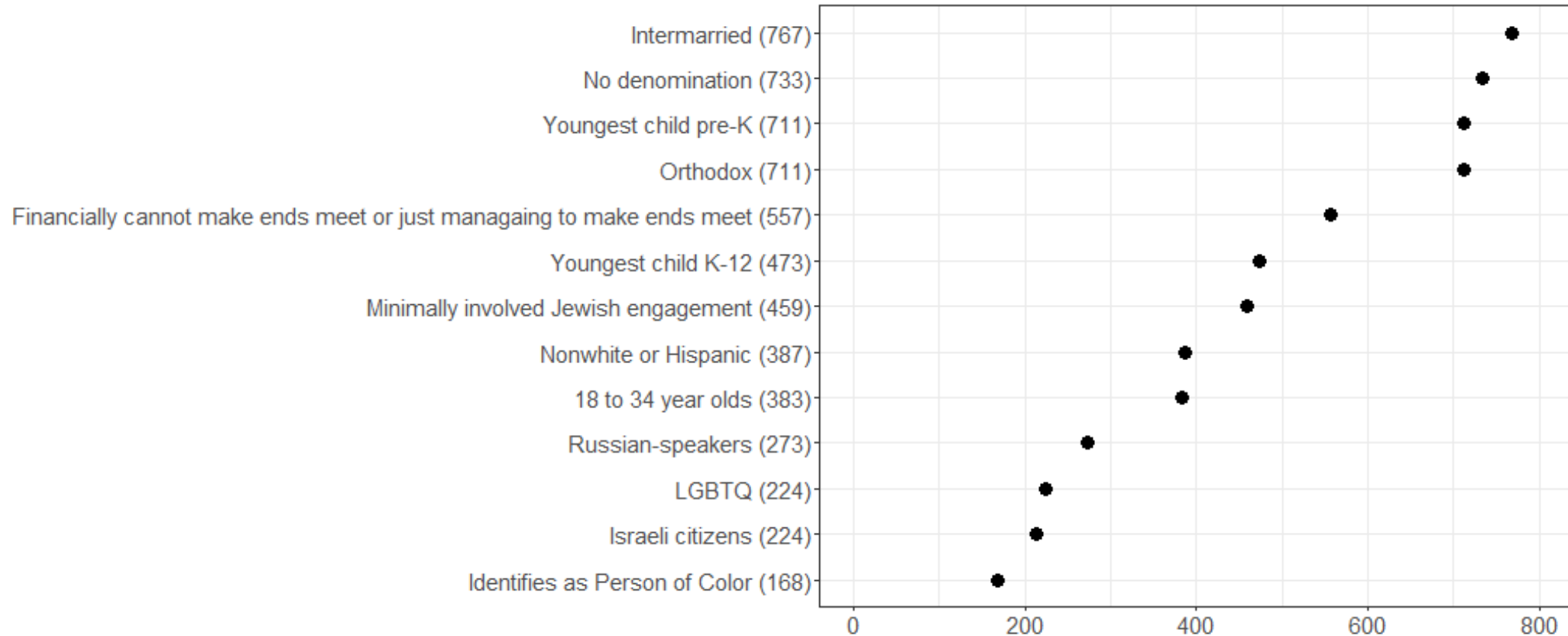
Predictive modeling facilitates reaching households not on Jewish organization lists in cost-effective and efficient manner.

For 2021 Los Angeles study:

- Households from a first-stage sample matched with more than 1,200 variables available from a vendor database and block group-level data from the 2020 U.S. Census Bureau Planning Database
- Gradient boosted machines used to estimate the probability of being Jewish for each adult matched with vendor data
- If a household had an adult with a sufficiently high propensity score for being likely Jewish, household assigned to 'Likely Jewish' stratum

Chicago and LA methodologies succeeded with obtaining completes to support estimates for hard-to-reach subgroups of the Jewish community.

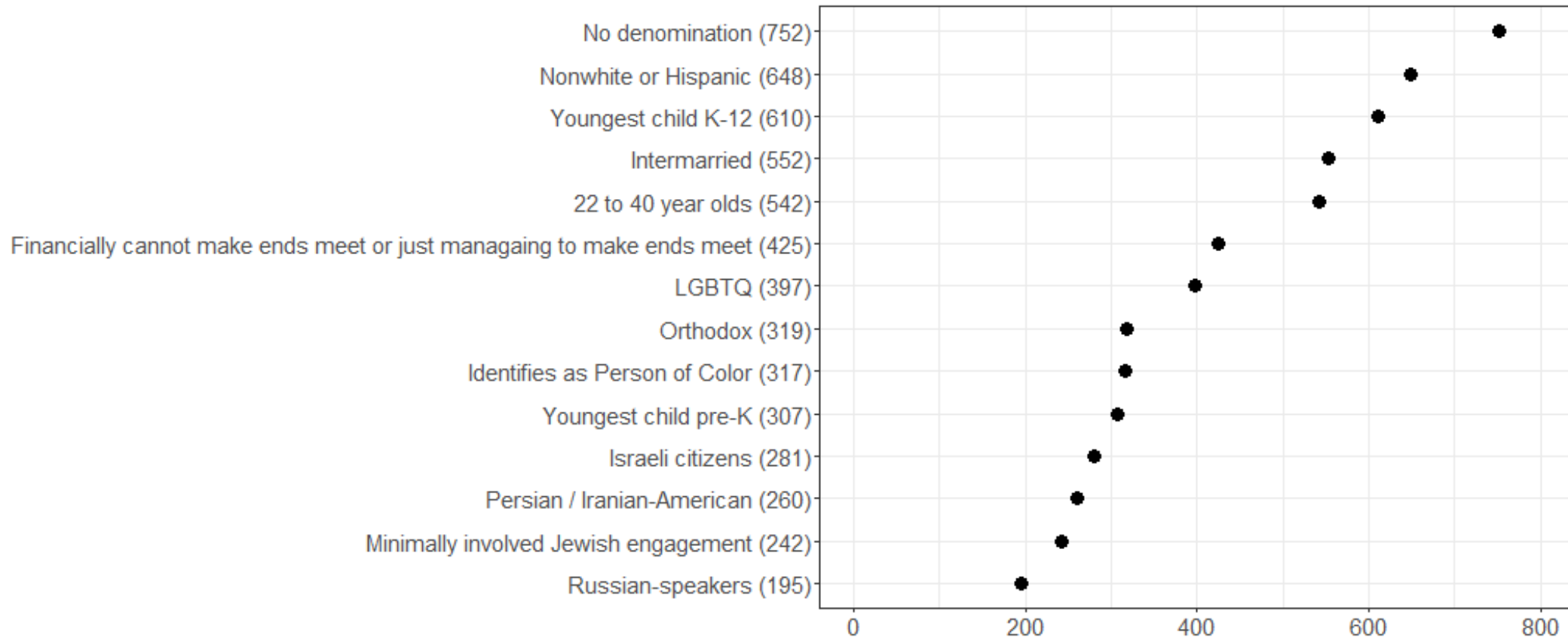
**2020 Chicago Study Completes for Key Subgroups out of 3,877 Total Completes**  
Number of Subgroup Completes in Parentheses



Chicago and LA methodologies succeeded with obtaining completes to support estimates for hard-to-reach subgroups of the Jewish community.

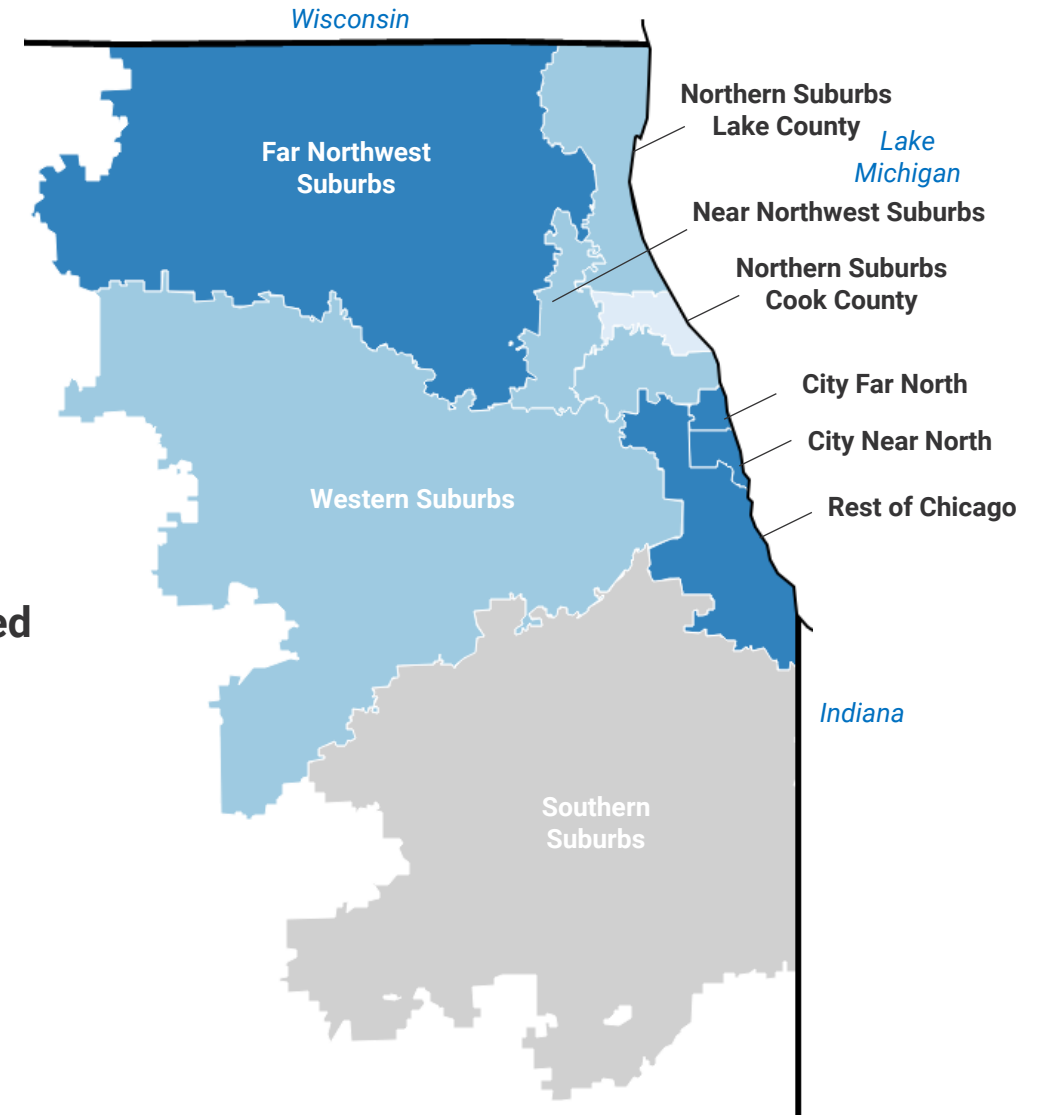
### 2021 LA Study Completes for Key Subgroups out of 3,012 Total Completes

Number of Subgroup Completes in Parentheses

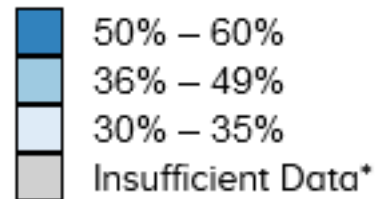


The 2020 Chicago methodology in particular supported estimates of the characteristics of the Jewish population for nine different geographic subregions of the Metropolitan Chicago area:

<https://2020populationstudy.juf.org/>

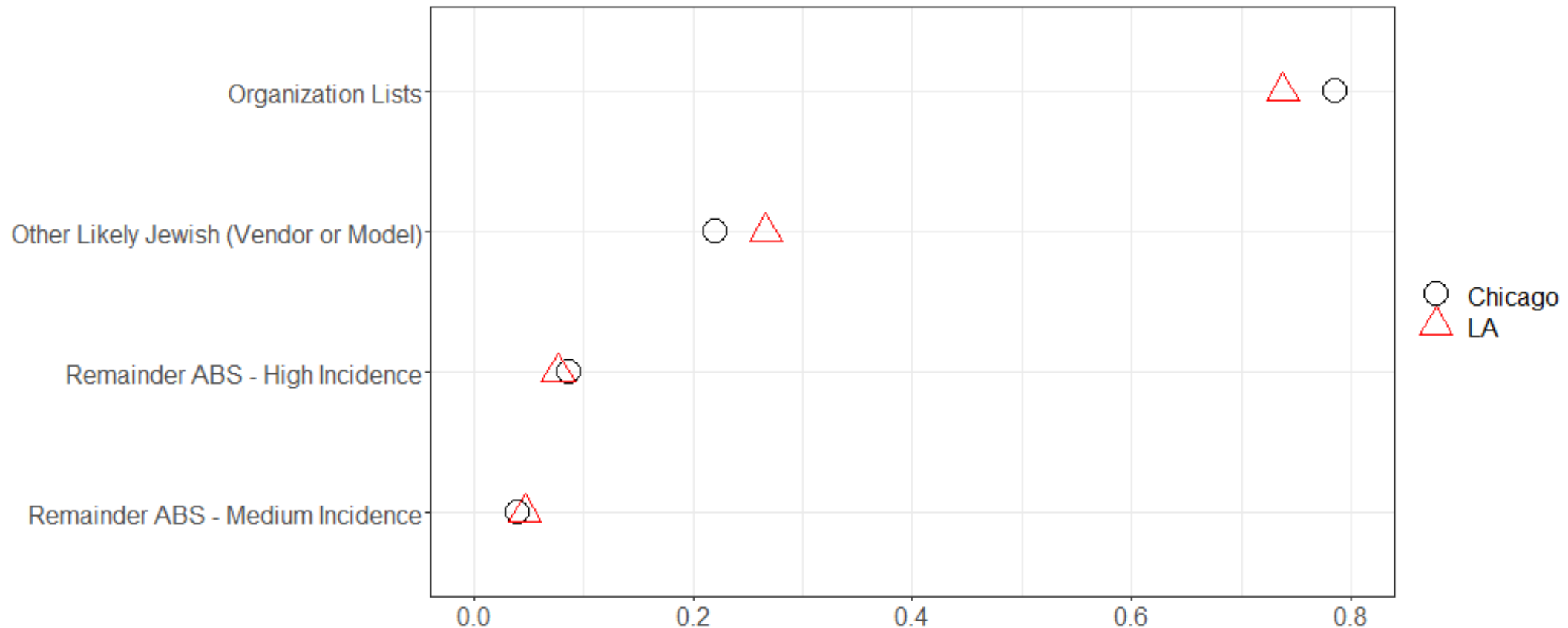


**% of Jewish households with financial, health, or social service need**



Employing predictive modeling and oversampling high incidence Jewish areas benefited eligibility rates for these studies.

### 2020 Chicago and 2021 LA Studies: Eligibility Rates by Sampling Source



Address-based sampling has a promising future as a methodology for Jewish community studies, but it must be employed with care. Many lessons apply for studies of rare populations.

- Comprehensive development of an organization list frame critical for representing diverse segments of the community and supporting estimates for small subgroups.
- Stratification using measures of Jewish incidence balances needs for high population coverage and maintaining budget.
- Predictive models for likelihood of being Jewish are a proven success to cost-effectively include households not on Jewish organization lists.

- Aronson, J.K., Boxer, M., Saxe, L., Seeskin, Z. H., Cohen, S. (2021). 2020 Jewish Chicago: Who We Are: Population Study. <https://www.jewishdatabank.org/databank/search-results/study/1127>
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
- Dutwin, D. (2020). Feedback Loop: Using Surveys to Build and Assess Registration-Based Sample Religious Flags for Survey Research. *Big Data Meets Survey Science: A Collection of Innovative Methods*, 535-559.
- English, N., Li, Y., Mayfield, A., & Frasier, A. (2014). The use of targeted lists to enhance sampling efficiency in address-based sample designs: Age, race, and other qualities. *2014 Proceedings of the American Statistical Association, Survey Research Methods*. Alexandria, VA: American Statistical Association.
- Lavrakas, P., Benson, G., Blumberg, S., Buskirk, T., Cervantes, I. F., Christian, L., Dutwin, D., Fahimi, M., Fienberg, H., Guterbock, T., Keeter, S., Kelly, J., Kennedy, C., Peytchev, A., Piekarski, L., & Shuttles, C. (2017). The future of U.S. general population telephone survey research. AAPOR Task Force Report.
- Olson, K., Smyth, J. D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., Mathiowetz, N. A., McCarthy, J. S., O'Brien, E., Opsomer, J. D., Steiger, D., Sterrett, D., Su, J., Suzer-Gurtekin, Z. T., Turakhia, C., & Wagner, J. (2021). Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR Task Force Report. *Journal of Survey Statistics and Methodology*, 9(3), 381-411.



A LANDMARK NORC STUDY  
SINCE 1972



# Thank you.

**Zachary H. Seeskin**  
Senior Statistician  
[Seeskin-Zachary@norc.org](mailto:Seeskin-Zachary@norc.org)

---

 Research You Can Trust™

---

 **NORC** at the  
University of  
Chicago