The Collaborative Multiracial Post-Election Survey (CMPS) Oversamples

INTRODUCTION TO THE COLLABORATIVE MULTIRACIAL POST-ELECTION SURVEY (CMPS) OVERSAMPLES

Lorrie Frasure, *University of California, Los Angeles, USA*
Janelle Wong, *University of Maryland, College Park, USA*
Edward D. Vargas, *Arizona State University, USA*
Matt A. Barreto, *University of California, Los Angeles, USA*

DOI:10.1017/S1049096524000659

This article introduces the contributions to the *PS Spotlight: The Collaborative Multiracial Post-Election Survey (CMPS) Oversamples.* Each feature in this issue uses data from the 2020 CMPS to help us understand the strengths and limitations of survey oversamples and to discuss best practices for users of these data. The CMPS has changed the way data are collected and shared in the social sciences. It is a nonpartisan, cooperative, multiracial/ethnic, multilingual, post-presidential election online survey conducted in the United States. It was developed by political scientists in 2008. Beginning in 2016, these co-principal investigators (PIs) led an innovative cooperative funding strategy that broadened the scope of access to high-quality national survey data with large samples of racial/ethnic and underrepresented groups in the United States. See Barreto et al. (2018) for a comprehensive description of the CMPS survey design and methodology.

The 2020 CMPS featured six unique oversamples of minoritized populations in the United States. For many reasons, the need to oversample minoritized populations is more pressing than ever. Nevertheless, the experiences of these groups often are understudied and underexplored in social science research because traditional sampling methods cannot yield sufficient data for statistical analyses. The CMPS's oversampling approach provides scholars with the sample size needed to conduct in-depth analyses. This can result in more comprehensive and inclusive findings, allowing researchers to understand more fully the attitudes, behaviors, and needs of underrepresented communities.

For the 2020 CMPS, we invited nine scholars to serve as Oversample Directors. They are experts in race, ethnicity, identity, and politics for the six unique oversamples, and include the following:

- Afro-Latinos: Danielle Pilar Clealand (University of Texas at Austin)
- Black Immigrants: Christina Greer (Fordham University) and Candis Watts Smith (Duke University)
- LGBTQ+: Andrew Flores (American University)
- Middle Eastern and North African (MENA)/Muslim Americans: Karam Dana (University of Washington Bothell) and Nazita Lajevardi (Michigan State University)
- Native Americans: Raymond Foxworth (University of Colorado Boulder) and Laura Evans (University of California, Riverside)
- Native Hawaiians: Ngoc Phan (Hawai'i Pacific University)

Random probability sampling techniques of relatively small racial, ethnic, and other identity groups may not yield a sufficient sample size to allow researchers to conduct in-depth analysis of those groups. Survey oversamples allow researchers to select respondents so that some groups comprise a larger share of the survey sample than in the overall population (Mercer 2016). After data collection is complete, sampling weights are applied to the data to align the population with their actual proportion. The larger sample allows researchers to conduct more in-depth analysis of an individual group (Mercer 2016). The practice of oversampling groups that comprise a small proportion of the general population is expensive. Researchers often cannot afford the costs to carry out this sampling technique. Therefore, cooperative funding strategies remain important in building partnerships for scholars to collect survey oversamples.

This cooperative research model is user-content driven, whereas survey items included on the CMPS are generated through survey-question contributions from a national consortium called the CMPS Scholars Research Network. To date, this consortium includes almost 250 researchers from almost 100 accredited colleges and universities, across multiple academic disciplines, including the social sciences, psychology, public policy, public health, education, law, and other fields. With support from the National Science Foundation, the 2020 CMPS expanded in both size and content. It included a total of 14,977 completed interviews in the primary sample of Asian American, Black, Latinx, and White respondents. Table 1 presents the overall sample sizes of additional completed interviews with adults across various oversample groups from hard-to-reach populations. This includes Afro-Latinos, Black Immigrants, Native Americans, Native Hawaiians, Muslims (including MENA), and people who identify as LGBTQ+. The full dataset included 17,545 adult interviews. The invitation and the survey were available to respondents in English, Spanish, Chinese (simplified and traditional), Korean, Vietnamese, Arabic, Urdu, Farsi, and Haitian Creole.

Because of the primary interest in the 2020 election, the project began with a large sample of registered voters from online sources that were pre-matched to the voter file. In addition, the data included a sample of nonregistered adults, including noncitizens. It is important to note that the 2020 CMPS also included a pilot sample of 16- and 17-year-old youth (N=1,457). The combined dataset was collected online in a respondent self-administered format from April 2 to October 4, 2021.

Understanding Oversample Categories and How Respondents Self-Identify

Oversample categories in table 1 are *not* mutually exclusive. The categories are based on respondent self-categorization and

© The Author(s), 2025. Published by Cambridge University Press on behalf of American Political Science Association.

https://doi.org/10.1017/S1049096524000659 Published online by Cambridge University Press

self-reporting. For example, respondents can be in the LGBTQ+ oversample as well as the American Indian oversample; the Afro-Latino oversample as well as the Black Immigrant oversample;. or the American Indian sample as well as the Hispanic/Latino oversample.

For many reasons, the need to oversample minoritized populations is more pressing than ever.

Throughout this project, the intention of CMPS co-PIs was to collect the most diverse sample of respondents in the United States as possible and to allow respondents to self-identify their race, ethnicity, and immigration status. The 2020 US Census revealed that the multiracial population (i.e., respondents who indicated two or more race categories) increased by 127% from 2010 to 2020. For 50 years in the US Census, Hispanic and Latino respondents were almost by definition "multiracial" or "multiethnic" because they answered both race and ethnicity questions. The outcome is that in both the real world and our data, respondents do not necessarily fit into clear or mutually exclusive categories. However, it also is the case that many respondents identify as single-race or monoracial.

For our overall project, we used several decisions in our initial categorization of respondents by race and ethnicity to meet minimum sample-size thresholds. Of course, end users can choose how to subset the data for analysis of different racial or ethnic groups. Therefore, we encourage individual researchers to be transparent in how they classify respondents in their analysis.

Overview of Featured Articles

This section presents the six insightful articles authored by the CMPS Oversample Directors. Ngoc Phan and Leilani DeLude's discussion underscores the distinction among Native Hawaiians, Asian Americans, and Pacific Islanders in their racial, social, and legal position in American society. They emphasize factors that render recruitment difficult, including distrust of outsiders, contextual factors shaping self-identification, and geographic limitations of islands and the continental United States. They conclude by reviewing key findings from the 2020 CMPS Native Hawaiians oversample.

Danielle Pilar Clealand discusses the nuanced process of identifying Afro-Latinos, including measuring racial identity and blackness in a mixed-race population. The primary difficulty in data collection is the tension between participants' self-identification and the racial frameworks of the United States, Latin America, and social sciences. Clealand describes the sociodemographic characteristics of the sample, such as whether respondents identify primarily as Latino or black and how that matters for identity politics.

Raymond Foxworth and Laura Evans (with Cheryl Ellenwood) examine the Native American oversample. Because there are 574 federally recognized sovereign tribes in the United States, the 2020 CMPS recognized the importance of capturing these communities' unique political perspective. They describe the role of Native American identification, exploring participants' sociodemographic characteristics and political attitudes, and conclude by discussing the challenges in obtaining Native American samples and future considerations.

Christina Greer and Candis Watts Smith note that the black population in the United States is ethnically diverse, from immigrant communities to second-generation Americans: 12% identifies as immigrants. They emphasize the need for black politics to expand to a more global perspective and the opportunities in the CMPS to do so. Challenges include differences in immigrant identification and country of origin that shape self-identification (e.g., a Black Immigrant or an Afro-Latino).

Nazita Lajevardi and Karam Dana highlight the importance of investigating MENA and Muslim perspectives in political science, given the increasing levels of Islamophobia and discrimination against these communities. Through the 2020 CMPS, they sought to address this need. They encountered several complications in recruiting participants due to distrust—particularly fear of surveillance and measurement issues for religious identification. Their best practices include a screener question about a respondent's religion, a question about the faith group in which a participant was raised, development of a community sample for snowball sampling, and utilization of nonvoter lists.

In his article, Andrew Flores describes the need for large datasets in the study of LGBTQ+ politics, challenges in data collection, general findings, and takeaways for the 2024 CMPS. Given the lack of US Census data about the LGBTQ+ community, it is difficult to follow best statistical practices for stratification and weights. Moreover, the terms in the LGBTQ+ fluctuate changing accessibility terms in survey items. For future studies, he recommends including questions about sexual orientation and gender identity.

Conclusion

Since 2016, through research collaborations, conferences, workshops, and writing retreats, we have convened a diverse and

Table 1

Overall Sample Sizes of the Adult Oversample

	Afro-Latino	Black Immigrant	MENA Muslim	American Indian	Hawaiians Pacific-Islander	LGBTQ+
Registered	519	648	315	915	194	885
Nonregistered	659	831	469	1,041	232	858
Oversample Totals	1,178	1,479	784	1,956	426	1,743

Source: 2020 Collaborative Multiracial Post-Election Survey Oversamples.

https://doi.org/10.1017/S1049096524000659 Published online by Cambridge University Press

multidisciplinary group of researchers in varying stages of their academic career. Using the collaborative, inclusive model of resource-sharing that we developed in 2016 and 2020, the 2024 CMPS will continue to expand research and professional-development opportunities for faculty; undergraduate and graduate students; and postdoctoral scholars from large research institutions, smaller liberal arts colleges, Historically Black Colleges and Universities, Tribal Colleges and Universities, and Hispanic Serving Institutions. This inclusive research and data-collection model will continue to highlight the voices of underrepresented groups in society and politics and also foster community among scholars in the social sciences and beyond.

ACKNOWLEDGMENT

The authors acknowledge the support of the National Science Foundation Award No. 1918510.

DATA AVAILABILITY STATEMENT

The 2020 Collaborative Multiracial Post-Election Survey is available at the Inter-University Consortium for Political and Social Research (https://doi.org/10.3886/ICPSR39096.v1).

CONFLICTS OF INTEREST

The authors declare that there are no ethical issues or conflicts of interest in this research. ■

REFERENCES

Barreto, Matt A., Lorrie Frasure-Yokley, Edward D. Vargas, and Janelle Wong. 2018. "Best Practices in Collecting Online Data with Asian, Black, Latino, and White Respondents: Evidence from the 2016 Collaborative Multiracial Post-Election Survey." *Politics, Groups, and Identities* 6 (1): 171–80. DOI:10.1080/21565503.2017.1419433.

Mercer, Andrew. 2016. "Oversampling Is Used to Study Small Groups, Not Bias Poll Results." Washington, DC: Pew Research Center. www.pewresearch.org/fact-tank/2016/10/25/oversampling-is-used-to-study-small-groups-not-bias-poll-results.

INTRODUCING THE 2020 CMPS MENA/MUSLIM SAMPLE

Karam Dana, *University of Washington Bothell, USA*
Nazita Lajevardi, *Michigan State University, USA*

DOI:10.1017/S1049096524000180

Why Study American Muslims?

For many years, American Muslims have experienced heightened scrutiny due to widespread societal apprehensions about Islam. Research has documented the discrimination against Muslims by the public, elites, and masses (Lajevardi 2020; Oskooii, Dana, and Barreto 2021) as well as the impact of this hostile environment on Muslims themselves (Dana et al. 2019; Sediqe 2020). Overall, findings indicate that whereas American Muslims face challenges akin to other racial and immigrant groups, they also grapple with gaining acceptance and being recognized for their positive contributions to America's diverse society.

Within this context, Muslims face unique challenges that raise particular questions about their democratic inclusion. Moreover, the diversity and heterogeneity of the US Muslim population across numerous markers—including race, denomination, nativity, age, socioeconomic status, sexual and gender identity, and religiosity—all matter for shaping whether and how intensely they have encountered such challenges, as well as their responses. A comprehensive understanding of their experiences, and how these compare to those of other American groups, requires extensive samples of both Muslim and non-Muslim individuals, along with detailed questioning on a broad range of sociopolitical issues.

Fortunately, the 2020 Collaborative Multiracial Postelection Survey (CMPS)—a national survey of voters and non-voters that oversamples racial and ethnic minorities on political and social issues—provides a unique opportunity to delve deeper into the experiences of US Muslims compared to other minoritized groups. The 2020 CMPS consisted of large samples (~4,000) of white, Latino, Black, and Asian respondents. Importantly, given the salience of other socially stigmatized communities, the 2020 CMPS also included oversamples of LGBTQ+ Americans, Native Americans, Native Hawaiians, Afro Latinos, African Immigrants, and Muslim Americans. With the CMPS, for the first time, researchers can thoroughly investigate Muslim experiences in the areas of politics, identity, housing, employment, law enforcement, immigration, media, education, and environmental concerns.

Complications of Collecting Survey Data on Muslims

Studying US Muslims presents unique challenges, which distinguishes it from research on other marginalized groups. Typically, researchers use demographic details such as group size, composition, and location when they study marginalized populations (e.g., racial and ethnic minorities). This rich dataset is notably lacking for Muslims, making study design difficult. The following three main issues hinder comprehensive survey data collection about US Muslims:

1. The US Census does not query religious identification. Although certain groups, such as Hispanics and Asian Americans, may be inferred from census-recorded surnames, there is no analogous method for Muslims due to the absence of a reference list (Barreto and Dana 2019). This data gap limits accurate estimations of religious populations nationwide (Barreto and Dana 2019).
2. Muslims are incredibly diverse in their racial backgrounds. Although a substantial proportion have roots in the Middle Eastern/North African (MENA) region, using MENA populations for rough estimations is problematic because the US Census currently groups MENA Americans under the "White" category, which obfuscates specific national origins (d'Urso 2022).
3. Alarmingly, some US Muslims perceive survey-recruitment efforts as a potential surveillance method, which heightens their anxiety, lowers response rates, and possibly biases responses due to social desirability (Calfano, Lajevardi, and Michelson 2019).

Findings From The 2020 CMPS

This section describes our sampling challenges and provides a preview of our sample.

Sampling Challenges

Achieving a fully representative Muslim sample is challenging due to inherent difficulties. However, we used benchmarks from the

https://doi.org/10.1017/S1049096524000659 Published online by Cambridge University Press