Political Science Research and Methods (2025), **13**, 781–797
doi:10.1017/psrm.2025.18

ORIGINAL ARTICLE

Survey sampling in the Global South using Facebook advertisements

Leah R. Rosenzweig[1,2], Parrish Bergquist[3], Katherine Hoffmann Pham[4], Francesco Rampazzo[5] and Matto Mildenberger[6]

[1]Center for Global Development, Washington, DC, USA; [2]Development Innovation Lab, University of Chicago, Chicago, IL, USA; [3]Political Science Department, University of Pennsylvania, Philadelphia, PA, USA; [4]Department of Technology, Operations, and Statistics, NYU Stern School of Business, New York, NY, USA; [5]Department of Sociology, Nuffield College, University of Oxford Nuffield College, Oxford, England, UK and [6]Department of Political Science, University of California Santa Barbara, Santa Barbara, CA, USA
Corresponding author: Parrish Bergquist; Email: pberg@upenn.edu

First two authors listed are joint first authors whose names have been randomly ordered. Contributing co-authors' names are listed in random order.
This article was originally published with an incorrect author affiliation. The error has been corrected and the online HTML and PDF versions updated.

(Received 8 May 2024; revised 20 September 2024; accepted 29 September 2024)

Abstract
Survey research in the Global South has traditionally required large budgets and lengthy fieldwork. The expansion of digital connectivity presents an opportunity for researchers to engage global subject pools and study settings where in-person contact is challenging. This paper evaluates Facebook advertisements as a tool to recruit diverse survey samples in the Global South. Using Facebook's advertising platform, we quota-sample respondents in Mexico, Kenya, and Indonesia and assess how well these samples perform on a range of survey indicators, identify sources of bias, replicate a canonical experiment, and highlight trade-offs for researchers to consider. This method can quickly and cheaply recruit respondents, but these samples tend to be more educated than corresponding national populations. Weighting ameliorates sample imbalances. This method generates comparable data to a commercial online sample for a fraction of the cost. Our analysis demonstrates the potential of Facebook advertisements to cost-effectively conduct research in diverse settings.

Keywords: Facebook; Global South; Meta; public opinion; social media; survey error; survey samping

Survey research in the Global South traditionally requires local enumerators to conduct face-to-face surveys with respondents, necessitating large budgets and lengthy fieldwork. Historically, there have been few alternatives to in-person recruitment due to poor electricity coverage and limited phone or internet connectivity. However, today much of the world's population is digitally accessible. By the end of 2022, an estimated 68% of the world's population had a mobile phone subscription, and approximately 55% had access to mobile internet (Rotondi *et al.*, 2020; GSMA, 2023). This connectivity presents an opportunity for public opinion researchers seeking to collect data in settings where on-the-ground, resource-intensive models of research are costly or challenging, such as due to natural disasters, violent conflicts, or pandemics.

© The Author(s), 2025. Published by Cambridge University Press on behalf of EPS Academic Ltd. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

This paper evaluates one emerging method to recruit online samples: social media advertisements. The analysis builds on prior work finding that low-cost online platforms allow scholars to generate valid survey-experimental results (Berinsky *et al.*, 2012; Mullinix *et al.*, 2015; Coppock and McClellan, 2019). We take this research program a step further, asking if Facebook can be used to generate high-quality samples for broader survey research purposes. According to Meta, Facebook's parent company, Facebook had over 3 billion monthly active users as of June 2023 (Meta, 2023), more than a third of the global population. As such, it is the most widely used social media platform around the world (Ortiz-Espina, 2019). Given this massive user base, the platform offers researchers access to nationally, culturally, and demographically diverse global populations. Scholars are already using the platform to quickly and cheaply recruit diverse populations (Ryan, 2012; Kapp *et al.*, 2013; Broockman and Green, 2014; Ramo *et al.*, 2014; Samuels and Zucco Jr, 2014; Bond and Messing, 2015; Hirano *et al.*, 2015; Jäger, 2017; Pötzschke and Braun, 2017; Bicalho *et al.*, 2020; Grow *et al.*, 2020; Rosenzweig, 2021; Finkel *et al.*, 2023; Grewal, 2023; Kilavuz *et al.*, 2023; Noh *et al.*, 2023; Jacobson, 2024; Offer-Westort *et al.*, 2024), but survey methodologists caution against an uncritical adoption of new recruitment methods without considering possible limitations on survey sample quality (e.g., Ansolabehere and Schaffner, 2010; Berinsky *et al.*, 2012; Ternovski and Orr, 2022). To this end, we provide a systematic assessment of the opportunities and drawbacks of Facebook advertisements as a recruitment tool in the Global South.

While most research has used Facebook to recruit convenience samples or specific target populations, an emerging literature has assessed whether the platform can also recruit nationally representative samples. As Boas *et al.* 2020 demonstrate in India and the United States, Facebook's high penetration and diverse user base make the platform an attractive tool for recruiting nationally representative samples. Zhang *et al.* 2020 demonstrate this potential by using targeted Facebook advertising to recover an approximately nationally representative sample in the United States. Neundorf and Öztürk (2023) show how different targeting strategies influence the cost and representativeness of samples recruited in the UK, Turkey, Spain, and Czechia. This research underscores Facebook's potential for survey sampling, but researchers are still left with a geographically and conceptually incomplete understanding of the platform's benefits and drawbacks.

Here, we expand the geographic scope and conceptual underpinnings of Facebook's potential as a recruitment tool for survey respondents. To provide the conceptual scaffolding for our analysis, we apply the Total Survey Error framework (Deming, 1944; Ansolabehere and Schaffner, 2010; Groves and Lyberg, 2010). We follow prior work (Berinsky *et al.*, 2012) in using a canonical survey experiment to show that Facebook passes the (low) bar of providing a platform for conducting valid survey experiments. To explore the tool's use for broader public opinion research applications, we compare the demographic composition and estimates of political attitudes from Facebook-recruited samples in Kenya ($n = 1{,}528$), Mexico ($n = 5{,}168$), and Indonesia ($n = 3{,}277$) with estimates derived from nationally representative benchmark data sets in each country. We also compare our Facebook sample in Indonesia with an online sample recruited by a survey firm, which is the likely alternative for most researchers. Our analysis of survey error highlights the sources of potential bias in Facebook survey samples, exposes which sources of bias researchers can control, and illuminates how researchers can reduce bias. Whereas most existing studies have focused on the U.S. and Europe, we evaluate the quality of Facebook-recruited samples in low and middle-income countries with lower levels of internet access, literacy, and Facebook marketing investment.

We show that quota sampling, which is also useful in other survey modalities, can help to overcome the bias inherent to using the Facebook user base as a sampling frame for national populations. Practically, we find that the Facebook platform allows for quick and cheap recruitment. However, our analysis also shows that researchers should tailor their use of this sampling method to both the type of research question asked and the target population of interest, and we show specific steps that researchers can take to reduce survey bias. More broadly, many other survey recruitment and implementation modalities face challenges of bias, noise, and nonresponse that are similar to those we

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

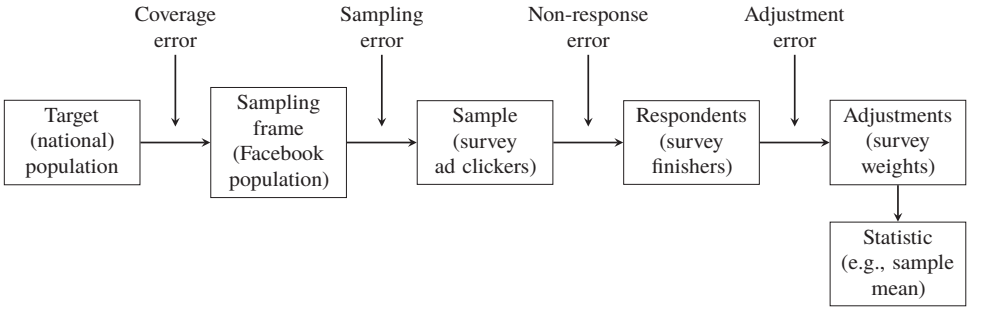


Figure 1. Components of representation error in Facebook samples.

highlight with Facebook. As such, we hope this paper is broadly useful to researchers considering the trade-offs involved in using other new platforms.

1. Conceptual framework: defining and measuring sources of survey error in public opinion samples

Our definition of survey quality builds on prior assessments of cost-quality trade-offs in public opinion samples. Scholars have assessed the quality of samples recruited through Amazon's Mechanical Turk (MTurk) (Berinsky *et al.*, 2012; Huff and Tingley, 2015), Prime Panels (Litman *et al.*, 2017), Lucid (now Cint) (Coppock and McClellan, 2019; Ternovski and Orr, 2022), Google Consumer Surveys (Santoso *et al.*, 2016), and Facebook Advertising (Kosinski *et al.*, 2015; Jäger, 2017; Boas *et al.*, 2020; Zhang *et al.*, 2020; Neundorf and Öztürk, 2023). Much of this work focuses on the quality of U.S. online samples, typically by comparing sample demographic statistics with the U.S. Census (Berinsky *et al.*, 2012; Huff and Tingley, 2015; Coppock and McClellan, 2019; Zhang *et al.*, 2020). These assessments are relatively straightforward in the U.S., where high-quality census data are frequently updated and many probability and quota samples are available as benchmarks.

We extend this work to three settings in the Global South, while also conceptually disaggregating and empirically assessing the types of errors that threaten the validity of conclusions drawn from Facebook surveys. We define error as unobserved disturbances that influence a statistical quantity of interest. Such error is problematic if it causes survey estimates to differ systematically from true population parameters. These systematic differences reflect bias in our estimates.

We use the Total Survey Error framework, first developed in the 1940s (Deming, 1944) and used by modern survey researchers (e.g.,Ansolabehere and Schaffner, 2010; Groves and Lyberg, 2010; Groves *et al.*, 2011; Lyberg and Weisberg, 2016), to define the distinct sources of error that threaten the external validity of conclusions derived from Facebook-recruited surveys. Since the framework's introduction, survey methodologists have defined two groups of quality indicators for survey statistics: measurement error[1] and representation error (Groves and Lyberg, 2010; Groves *et al.*, 2011). Here we focus on the latter. Errors of representation are systematic or random imperfections in the relationship between a target population, sampling frame, and sampling units, and they threaten external validity.

Figure 1, adapted from Groves *et al.* 2011 and Groves and Lyberg (2010), shows the inferential steps required to draw population conclusions from a Facebook-recruited survey sample. First, coverage error is defined as the gap between a target population and a sampling frame. In our case, the

[1]Measurement error addresses gaps between the constructs that a survey is designed to assess, the measures used in a survey, and responses to those measures. Measurement error is primarily a matter of questionnaire design and is not platform-specific. Still, we do examine one type—respondent attention—in the SI to assess if this component of measurement error might affect Facebook sample quality differently than on other platforms.

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

target population is a country's national adult population, and the sampling frame is the national adult population with a Facebook account in that country. The existence of adult residents without Facebook accounts will generate "undercoverage." Coverage bias is quantified as the difference between the mean value of a descriptive statistic in the national population and the Facebook population. Second, sampling error arises because not all individuals in the sampling frame are surveyed. Random sampling variation occurs since many different sets of individuals could be drawn from the sampling frame, simply by chance. Sampling bias arises when individuals in the sampling frame have unequal chances of being included in the sample. Third, unit nonresponse error arises when some sampled individuals fail to record meaningful and complete survey responses. Unit nonresponse is the gap between people who click on the advertisement for our survey and those who finish the survey. Nonresponse bias arises if certain types of people are more likely to finish the survey than others, and if there is a relationship between the likelihood of finishing a survey and survey responses. Finally, adjustment error arises when researchers weight data to give greater representation to cases that are underrepresented in the sample. These weights are used to reduce coverage, sampling, and nonresponse bias, but they can also increase these biases (Bailey, 2024).

We examine total survey error and disaggregate its components with two sets of empirical analyses. Our analyses build on prior work that compares survey-derived statistics to external benchmarks, such as a "gold-standard" survey, census, or re-interview data (Ansolabehere and Schaffner, 2010; Groves and Lyberg, 2010; Berinsky *et al.*, 2012; Kennedy *et al.*, 2018; Coppock and McClellan, 2019; Zhang *et al.*, 2020; Holliday *et al.*, 2021). We compare, first, the demographic compositions of our samples and, second, the public opinions measured in our surveys with those derived from other high-quality samples. Our disaggregated assessment of survey error allows researchers to gauge Facebook's suitability for specific research applications and highlights ways to minimize biases.

2. Data collection

2.1. Case selection and benchmark data sets

We fielded surveys in Mexico, Kenya, and Indonesia. As detailed in the SI, we selected case countries from three different continents that are neither best nor worst cases in terms of Facebook usage, where recent and accurate census data are available, where recent nationally representative survey data are available, and where high literacy rates and mobile phone access ensure broad accessibility of an online survey. We compare Facebook-recruited samples against high-quality data benchmarks in each case country. First, we use national censuses (KNBS, 2010; INEGI, 2015; BPS Statistics Indonesia, 2020). Second, we use well-respected, in-person, nationally representative surveys: the Latin American Public Opinion Project (LAPOP) Americas Barometer (LAPOP, 2018–2019), the Afrobarometer (Afrobarometer, 2019), and the Asian Barometer (Asian Barometer, 2019). In Indonesia, we also compare our Facebook sample with an original survey we fielded with Dynata, a survey firm that recruits respondents through its online panel. We make the comparison with Dynata to provide insights about the relative quality of Facebook samples with the most viable alternative for many researchers.

2.2. Quota sampling and survey weights

To minimize sampling bias, we use a stratified sampling approach designed to mimic the demographic-geographic stratified sampling approaches used by our in-person benchmark surveys. Within each geographic stratum, we designed target cells based on the demographic characteristics used in the benchmark surveys: gender in Kenya, and gender and age in Mexico and Indonesia. We then attempted to correct observed or expected imbalances by targeting additional respondents within underrepresented categories, including education (in Mexico and Kenya) and age (in Kenya). For all samples, we used iterative proportional fitting, or raking, to create weights for respondents who completed the survey. Our weights are designed to reflect the distribution of the national

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

populations (as measured in the national census) across gender, education, age cohort, and geography. The SI includes more details about our sampling and weighting protocols.

2.3. Survey instruments

To direct respondents to the surveys, we created Facebook pages representing our survey campaigns and placed ads from these pages targeting each sampling stratum. After clicking on a Facebook ad, respondents were sent to a Qualtrics-hosted survey fielded in the languages appropriate for each country. In all three surveys, we collected information on demographics and attitudes, which we used to compare Facebook samples against national census and benchmark survey data. The Kenya survey replicated questions from the Kenyan Census and the 2019 Afrobarometer survey. The Mexico survey replicated questions from the Mexican Census and the 2019 LAPOP survey. The Indonesian survey replicated survey questions from the 2019 Asian Barometer and from our original survey fielded with Dynata in October 2021.

3. Comparing demographics

To evaluate the quality of the statistical summaries derived from Facebook samples, we first examine total survey error by comparing the demographic characteristics of our samples with benchmark surveys and national census data. Our quantity of interest is the weighted sample mean survey response.

Figure 2 plots the distribution of demographic characteristics, compared with benchmark surveys and national census data. In all three countries, the weighted Facebook samples (filled, dark blue squares) differ most from the census (gold crosses) with respect to age and education. In Kenya and Indonesia, Facebook survey samples are younger than the national population and the barometer samples (red triangles). By contrast, the Mexican Facebook sample (similar to LAPOP) contains a greater share of respondents over 50, compared with the national population.

In general, Facebook samples are more educated than national populations and the barometer surveys. However, the Indonesian Facebook sample is less biased on education than the commercial online sample. The commercial online sample also underrepresents lower education respondents more severely than our Facebook sample.

We find little gender bias across all three countries. The distribution of religions in Kenya and Indonesia is also fairly representative, although some bias is evident in Mexico. Marital status in Mexico and Indonesia, and tribe in Kenya, also show minimal bias. However, there is substantial bias toward urban populations in Kenya.

3.1. Coverage error

To identify the specific sources of these biases, we begin with coverage bias: the mismatch between a target population and sampling frame. Here, we compare the national population (described by the census and shown as gold crosses in Figure 2) with the Facebook population (filled, dark blue squares in Figure 2) in each country. We approximate the demographic distribution of the full Facebook population using `pySocialWatcher`, a Python package. Around the time of data collection for each Facebook survey, we retrieved from the Facebook Marketing API the estimated number of daily active users and monthly active users matching different age, gender, education, and location characteristics. (Araujo *et al.*, 2017).

Coverage bias seems to contribute to the overrepresentation of more highly educated respondents observed in our Facebook samples. In all three countries, individuals who did not complete secondary school are severely underrepresented in the Facebook population, as they are in our Facebook samples. Similarly, those with a college degree are overrepresented in the Facebook population. In Kenya, education-related coverage bias is even larger than overall bias (the difference between the census

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

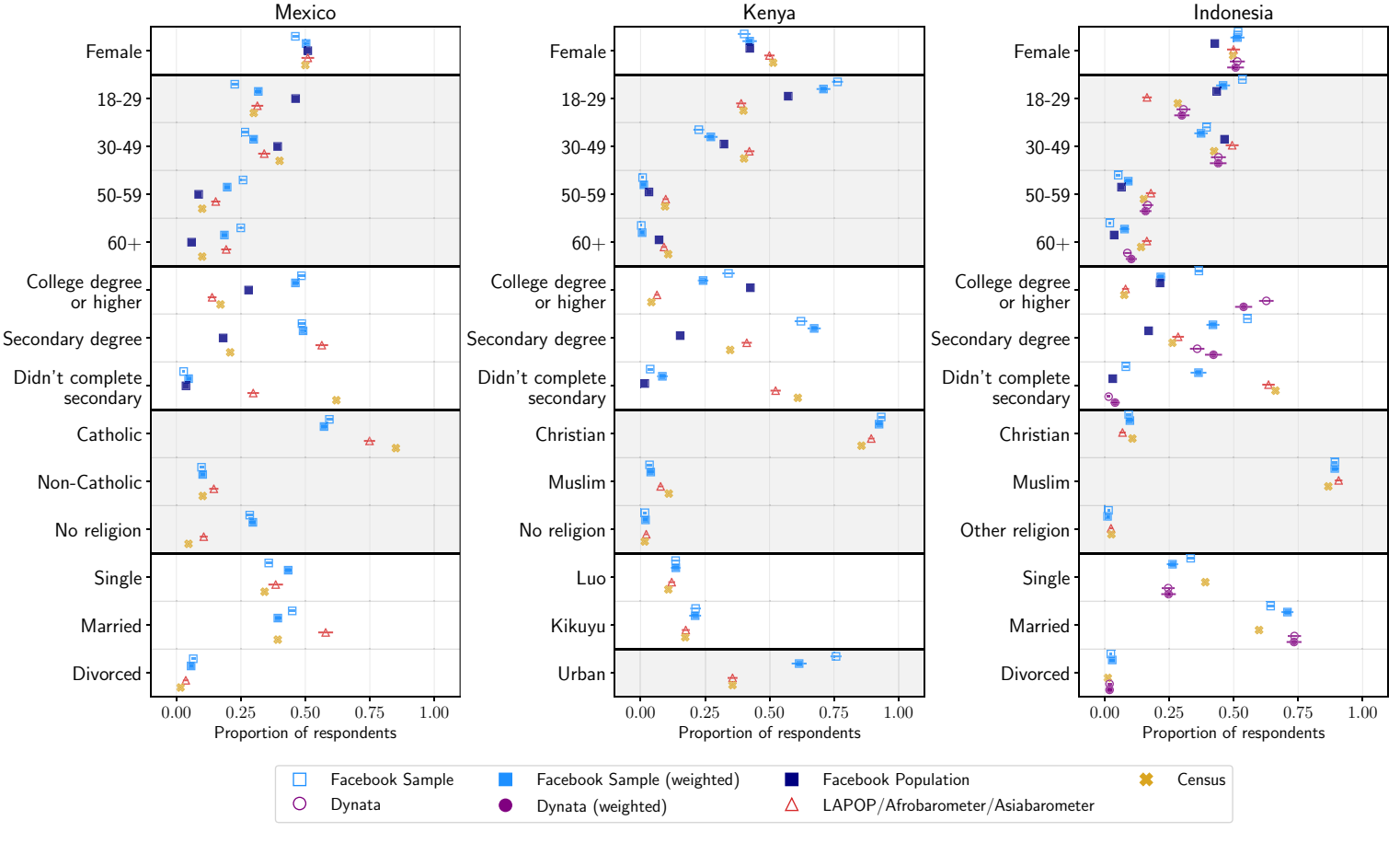


Figure 2. Demographic benchmark comparisons. *Note*: Means are plotted with 95% confidence intervals for demographics reported in the national census (no confidence intervals), nationally representative survey samples (LAPOP/Afrobarometer/Asian Barometer), the Facebook population (no confidence intervals), our original survey recruited with Dynata, and our Facebook samples. Facebook samples are weighted using raking to match the national census on gender, education, age cohort, and geography.

and our weighted Facebook estimates), and quota sampling appears to alleviate this bias. In Mexico, by contrast, college-educated individuals are more overrepresented in the Facebook sample than in the Facebook population. Thus, while coverage bias may account for some of the overrepresentation of more educated respondents in our samples, sampling and nonresponse error are also contributing factors, as we discuss in subsequent sections.

Coverage bias also helps account for the overrepresentation of young people in the Kenyan and Indonesian Facebook samples. In both countries, 18–29-year olds are overrepresented in the Facebook population and older individuals are slightly underrepresented. Still, in the Kenyan and Indonesian Facebook samples, 18–29-year olds are overrepresented even compared with the Facebook population. This suggests that coverage bias only partially accounts for the overrepresentation of young respondents. By contrast, coverage bias cannot account for the overrepresentation of older individuals in the Mexico Facebook sample, since these individuals are underrepresented in the Facebook population in Mexico.

Similar to the cases of Indonesia and Kenya, the sampling frame in Mexico overrepresents young people, but our unweighted Facebook sample approaches the census proportion of young people. We attribute this better balance to our use of more specific quota cells in Mexico.[2] Our stratified sampling approach helped us overcome coverage bias in recruiting our sample in Mexico.

With respect to gender, the Facebook population matches the census in Mexico quite closely, but in Kenya and Indonesia the Facebook populations underrepresent females compared to the national populations. The Kenya Facebook sample (weighted and unweighted) mirrors this underrepresentation. In Indonesia, however, the Facebook sample (weighted and unweighted) closely approximates the gender balance in the national population.

To assess how much of a problem coverage bias might generate, researchers can use the Facebook Marketing API to examine the Facebook population data for their target population of interest before beginning data collection. Stratified quota-based sampling and population-based weighting can help mitigate coverage bias even in cases where the sampling frame may not be representative of the population on a particular dimension.

3.2. Sampling error

While quota-based sampling can help researchers mitigate coverage bias, we need to examine whether the Facebook ad platform introduces sampling bias by limiting researchers' ability to reach particular groups. Sampling error arises because only some of the individuals in the sampling frame (Facebook users in each country) are included in the potential sample (Facebook users who are shown the ad and click on it). Sampling bias could arise at two points in the sampling process. First, ad design may appeal to some individuals more than others. Researchers have control over this step. They can, for example, experiment with the ad design to appeal to different respondent types. Second, the Facebook advertising platform may systematically fail to reach some individuals. While researchers can target certain individuals with recruitment quotas, Facebook's back-end data and algorithms (which are beyond researchers' control) determine whether ads reach targeted individuals.

To assess sampling bias, we focus on individuals who clicked our survey ad ("ad clickers"). We can observe Facebook's back-end demographic information for ad clickers, even if they did not complete the survey, because the survey records the sampling stratum of the ad they clicked on. We use this Facebook-inferred demographic information to check whether we recruited at least one individual from each stratum and, conversely, whether any individual types were systematically excluded. In Kenya and Mexico, we received responses from every quota, although in Kenya we did not reach quota targets in 49 of 66 strata.[3] In Indonesia, we were unable to recruit respondents from 67 (25%)

[2]See SI for more details.

[3]Strata targets were set according to the Afrobarometer's population-derived weights associated with each stratum's geolocation.

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

Table 1. Accuracy of Facebook targeting, as defined by the percent match between Facebook- and self-reported data

Characteristic	Mexico	Kenya	Indonesia
Primary sampling characteristics:			
Gender	99%	91%	76%
Age	87%	44%*	77%
Location◇	67%	73%	55%
Additional criteria:			
Education level	70%	12%‡	

Note: *In Kenya, a match was defined as whether respondents' self-reported age was 32 years old or above, for those who responded to an ad targeting this age group. The self-reported ages of these respondents ranged from 19 to 48 years old, with a mean of 31.
‡In Kenya, a match was defined as whether self-reported educational attainment was secondary school or less, for individuals who responded to an ad targeting respondents with an "unspecified" level of education.
◇For location matches, we aggregate self-reported and targeted location to the largest administrative unit used for targeting. In Kenya, this corresponds to the former provinces used as administrative units, matched from the county respondents say they live in and the locations used for ad targeting. For Indonesian respondents, we checked whether respondents self-reported living in the same province as the targeted province for the ad through which they were recruited. For Mexican respondents, we checked whether respondents self-reported living in the same group of municipalities targeted by the ad through which they were recruited. The "Recruitment and sampling" section of the SI provides additional clarity on the administrative units used for sampling in each country.

of 272 targeted strata. The vast majority of these strata contained men and women over 50 years old, though we also failed to reach women between 30 and 49 years in three provinces. This suggests that in Indonesia, sampling bias contributes to the underrepresentation of older individuals in our sample.

Of course, if Facebook's back-end data are inaccurate, then even perfect success at recruiting people from all survey strata will not ensure that all individuals have a chance of being sampled. To examine this possibility, we compare self-reported demographics with those reported by Facebook. Table 1 provides the percentage of people in each category for which self-reported and Facebook-targeted characteristics match.

The accuracy of Facebook's ad targeting varies across demographics and between countries. Facebook's targeting was remarkably accurate in correctly identifying respondents' gender in Mexico and Kenya (99% and 91% match, respectively) and slightly less accurate in Indonesia (76%). In both Indonesia and Kenya, the 10–20% of respondents recruited from an ad targeting the opposite gender might have resulted from respondents sharing ads with friends. It would not be surprising that greater sharing would have occurred in the context of the incentivized surveys in Kenya and Indonesia, compared to the non-incentivized survey in Mexico. Targeting by age was quite accurate in Mexico (87% match) and Indonesia (77% match), but not in Kenya (44%). Geographic targeting was similarly accurate in Mexico and Kenya but less so in Indonesia. Furthermore, education targeting was not very successful in Kenya, which is in part due to the fact that we attempted to reach less-educated respondents by targeting those with an "unspecified" level of education—meaning they had not reported their level of education on their profile. In contrast, education targeting in Mexico was reasonably accurate. We targeted respondents with no higher than a high school degree, and 70% of our respondents self-reported an education level consistent with this targeting criterion. We report more details on these comparisons in the SI.

These findings indicate that inferred demographics are not always accurate, highlighting the importance of collecting self-reported demographic data. We also recommend that researchers periodically examine the composition of the sample of ad clickers to gauge success at recruiting respondents from each quota (see SI Figure S4 for an example). To increase entrants from underrepresented groups in the sample of ad clickers, researchers could increase spending for these groups' ad sets, modify ads to increase appeal for these groups, or leave ads running for a longer period.

Overall, quota sampling helps us recruit a more diverse sample of respondents. Left to its own devices, Facebook's algorithm would optimize ad targeting to recruit the least expensive sample. This optimization entails recruiting respondents that are most similar to those who have already entered the survey. Running an ad campaign across quota sampling cells weakens this regression toward the most prevalent (or cheapest) respondent types. Still, the errors in Facebook's back-end inference of

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

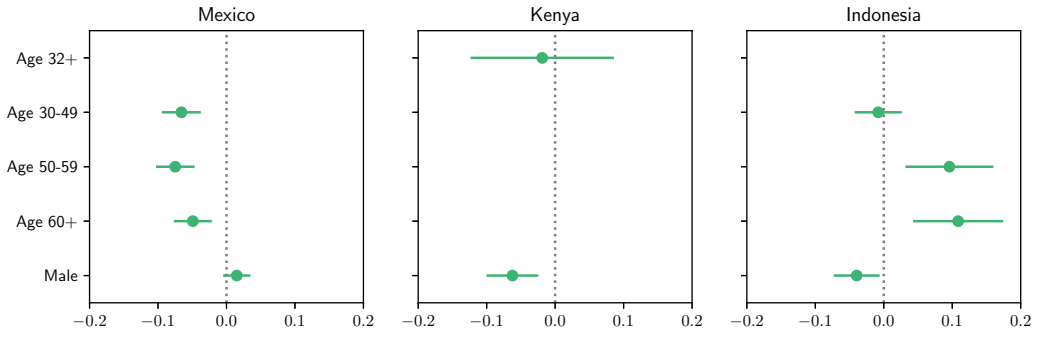


Figure 3. Predictors of nonresponse. *Note*: The figure shows the results from a linear regression of attrition (attrition = 1, completion = 0) on demographic characteristics used by Facebook to target individuals. Omitted categories are Female and 18–29 for the Mexican regressions, Female and 21–29 for the Indonesian regressions, and Female and younger than 32 years old for the Kenyan regressions. 95% confidence intervals are reported using heteroskedasticity-robust standard errors.

demographic attributes, as well as the fact that some people are excluded from the Facebook platform altogether, mean that researchers cannot fully overcome the gap between the Facebook and national populations through quota sampling alone.

3.3. Nonresponse error

Unit nonresponse is common across survey contexts and causes bias when it is nonrandom and correlated with survey answers (Bailey, 2024). We define attrition and nonresponse interchangeably, as entering the survey but failing to provide meaningful and complete responses.[4] Overall, the attrition rate was 46%, 31%, and 25% in Mexico, Indonesia, and Kenya, respectively. For context, attrition from the online sample for the American National Election Study was 14% in 2020 (DeBell *et al.*, 2020, p. 74), and attrition from the Cooperative Election Study was 21% in 2022 (Ansolabehere *et al.*, 2023, p. 11).[5] We examine whether nonresponse is systematic using the same Facebook-assigned demographic information used in the previous section. We first assess whether the likelihood of entering but dropping out of the survey was correlated with users' demographic characteristics (assigned by Facebook). Figure 3 shows that attrition is systematic, but the predictors of attrition vary across samples.

Incorporating age, gender, and education into adjustment weights helps to alleviate nonresponse bias associated with these observable characteristics. We reduce nonresponse bias by upwardly weighting individuals with underrepresented characteristics and down-weighting those with overrepresented characteristics. Researchers could also adjust their sampling strategy on the fly to improve response rates among high-attrition groups. For instance, researchers might increase ad spending or introduce incentives for high-attrition groups. Of course, unobserved sources of nonresponse present a more worrisome problem for descriptive inference, since researchers cannot adjust for unobserved and non-ignorable sources of nonresponse. To account for non-ignorable nonresponse, scholars can use bounds (Manski, 1990), sensitivity analysis (Hartman and Huang, 2024), selection models (McGovern *et al.*, 2018; Gomes *et al.*, 2019), nonresponse weights (Sun *et al.*, 2018), or other methods (Bailey, 2024).

[4]See the description of our sampling approaches in the SI for more details.

[5]We would expect attrition from these canonical surveys to be lower than ours because the samples consist, wholly or partially, of empaneled respondents whose compensation for participating ($40–$200 in the case of the ANES) is substantially higher in value than the compensation we provided (ANES methodology report, page 29).

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

3.4. Adjustment error

We can minimize nonresponse bias by weighting on observable characteristics that predict nonresponse, but weighting could introduce bias if responses within any weighting stratum are unrepresentative of public attitudes within that stratum. Adjustment bias occurs if, within groups defined by the weighting variables, nonresponse is non-ignorable (i.e., correlated with unmeasured characteristics that are also correlated with the outcomes of interest) (Bailey, 2024). If this is the case, up-weighting or down-weighting these unrepresentative individuals will introduce bias into the sample average.

We assess adjustment error by first examining whether weights correct demographic imbalances in the sample. Figure 2 summarizes how the application of weights corrects demographic imbalances by comparing unweighted (hollow, light blue squares) and weighted (filled, light blue squares) Facebook samples with census data (gold crosses). Weighting corrects the slight gender imbalances in all three samples. For age and education, weighting reduces the distance between the Facebook samples and the national populations but does not completely eliminate it. The remaining imbalances in age and education are due to trimming weights at the 95th percentile, to avoid excessively overweighting very rare respondent types. By construction, the Facebook sample almost perfectly matches the census populations when weights are not restricted. We find that our (trimmed) weights ameliorate biases.

We next investigate whether the weights introduce bias on other variables we measure and for which we have benchmark comparisons, but which are not incorporated into the weights (bottom sections, Figure 2). The intuition here is that non-ignorable nonresponse within strata would make itself apparent if weighting draws the sample away from the population distribution on dimensions that are not incorporated into the weights. Weighting minimally affects the distribution of these non-weighting variables (religion, marital status, tribe) except urban bias in the Kenya sample, which is reduced. This suggests a negligible correlation between age, gender, education, and geography and unmeasured factors that are associated with the cultural variables shown in Figure 2. This analysis provides some assurance that weighting does not increase bias for descriptive inferences. Likewise, weighting only slightly impacts public opinion estimates (see Section 5) and tends to move these estimates slightly toward—rather than away from—benchmarks. Weighting does introduce bias into our estimates of experimental effects in Indonesia; we examine this bias in the next section. In general, population-based weighting improves the quality of the statistics derived from our surveys, without introducing additional bias.

4. Replicating classic survey-experimental findings

Most social science researchers are primarily interested in public attitudes or behaviors, rather than demographic summaries. We now turn to these quantities of interest. As an initial check on the face validity of survey results derived from our samples, we use a canonical behavioral experiment—the Tversky and Kahneman (1981) "disease problem" used to test prospect theory. Since other low-cost platforms have been shown to produce reliable replications of this and other survey-experimental findings (Berinsky *et al.*, 2012; Mullinix *et al.*, 2015; Coppock and McClellan, 2019), we include this analysis as a low-bar, benchmark test with which many scholars are already familiar.

This survey experiment asks respondents to "Imagine that your country is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed." Then respondents are randomly assigned to see two program options which are framed either in terms of the number of lives that will be saved (the "save" condition) or the number of lives that will be lost (the "die" condition) under the two options. Within each condition, the difference between the two programs is whether the outcome is stated in certain or probabilistic terms. In expectation, the payoffs of both policies are equal. The finding that has been replicated in many samples is that respondents are more likely to choose the probabilistic (risky) option to avoid losses (in the "die" condition), and they are more likely to choose the certain option to accrue gains (the "save" condition).

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

Table 2. Replication of Tversky and Kahneman (1981) disease problem

	Tversky & Kahneman		Berinsky, Huber, & Lenz		Kenya Facebook sample (weighted)		Indonesia Facebook sample (weighted)	
Options	Save	Die	Save	Die	Save	Die	Save	Die
Certain	72%	22%	74%	38%	61%	36%	55%	49%
Risky	28%	78%	36%	62%	39%	64%	45%	51%

Note: The table shows the proportion of respondents choosing "certain" and "risky" policies to manage a disease, when the policies are framed in terms of lives saved vs. deaths. When the policies are framed in terms of the number of lives saved, a majority of respondents prefer the certain policy. When the policies are framed in terms of the number of people who will die, the majority prefers the risky option.

Table 2 shows the original results from Tversky and Kahneman (1981) among their sample of U.S. students, a replication by Berinsky *et al.* (2012) using a U.S.-based MTurk sample, and results from our Facebook samples in Kenya and Indonesia.[6] Our results are largely consistent with these studies and others that have replicated this experiment across cultural contexts (Ruggeri *et al.*, 2020; Im and Chen, 2022). For instance, Im and Chen (2022) report that the percent of respondents who pick the certain option when framed in a positive (save) vs. negative (die) framing in South Africa is 57% and 37%, in Mexico is 59% and 37%, and in Indonesia is 51% and 33%.

These preferences manifest in the same direction in our samples, with one notable exception: the application of weights skews the result in the "die" condition in Indonesia. Table S3 shows that the unweighted result is consistent with our expectations in the Indonesian sample. However, the result does not hold for respondents over the age of 50 (Table S4), which suggests that the older individuals in our sample are distinctive in how they respond to this experiment. Recall from Section 3.2 that we were unable to recruit respondents over the age of 50 from several of the strata in our sample (sampling bias). Up-weighting these individuals[7] in our estimates of experimental effects pulls the average effect toward the biased (relative to the benchmark) result among older individuals. This reflects adjustment bias. We highlight this finding because it reinforces our suggestion (introduced in Section 3.2) to monitor sampling quotas and dedicate resources to ensure sufficient recruitment within all strata. The finding also cautions against the uncritical use of weights to address underrepresentation of certain groups.

5. Comparing public opinion estimates

We next report descriptions of public opinions and political behaviors. We compare responses from our Facebook samples to identical opinion questions fielded in benchmark surveys (Figure 4). In general, where there are gaps between the Facebook-derived estimates and benchmark samples, Facebook samples report greater political activity. Note that the application of weights tends to move Facebook-derived estimates toward benchmark survey estimates. Applying weights tends to improve the accuracy of the estimates.

Persistent gaps in Figure 4 are likely attributable to some combination of sample composition and survey mode. Representation bias could account for some upward bias in estimates of political activity, since Facebook samples overrepresent highly educated respondents who are likely to be more politically engaged (Verba *et al.*, 1995). Still, survey mode could also play a role here. In an in-person survey context, social desirability bias may deter respondents from reporting activities, such as protesting, that challenge the government. Consistent with this tendency, barometer survey respondents report lower levels of these types of activities and, in Kenya, higher levels of approval of the President. Conversely, evidence suggests that turnout is overreported in surveys (Holbrook and Krosnick, 2010), and this bias in self-reporting may be larger for in-person surveys

[6]The Indonesian data analyzed here are from the pilot version of the survey, fielded in July 2023. The SI provides more information on the pilot wave of the survey.

[7]The average weight for respondents over 50 is 2.34.

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

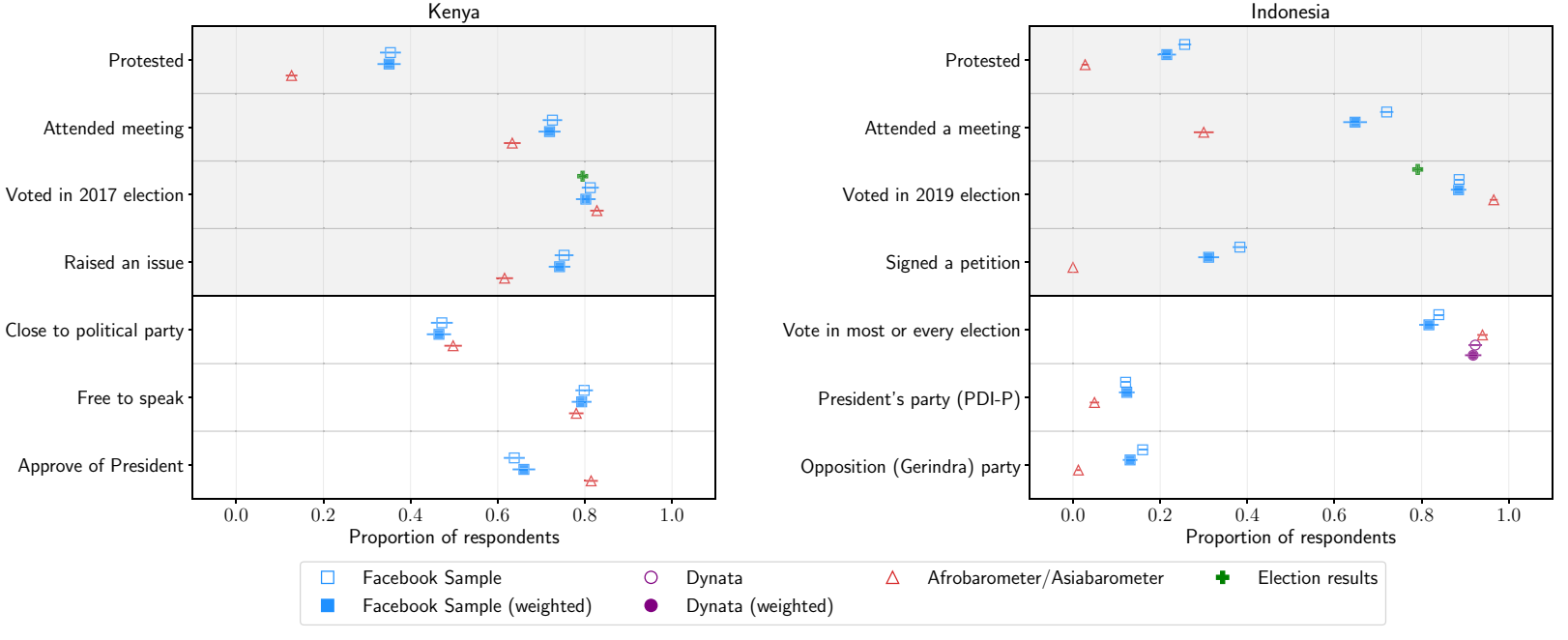


Figure 4. Political attitudes and behaviors. *Note*: This figure shows self-reported behaviors including identifying with a political party, voting, and engaging in activities such as community meetings and protests, from Afrobarometer/Asian Barometer and Facebook samples in Kenya and Indonesia. Estimates for both samples have been weighted using individual weights provided by Afrobarometer/Asian Barometer surveys or the raking procedure described above (for Facebook samples).

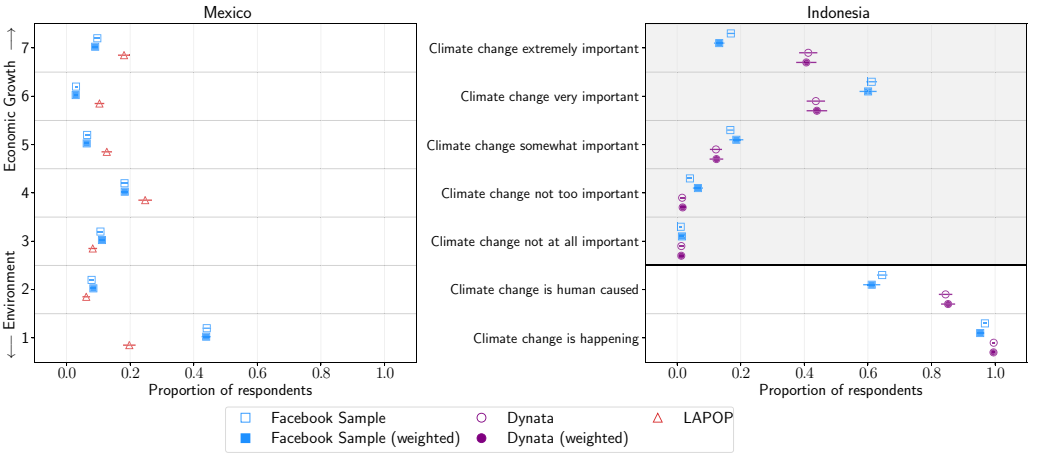


Figure 5. Public opinion estimates. *Note*: This figure shows responses from LAPOP and Facebook samples, to the question of whether environmental protection (1) or economic growth (7) should be given priority. Facebook results are weighted, and LAPOP results are unweighted, consistent with LAPOP documentation.

(Jackman and Spahn, 2019). Accordingly, voter turnout is the only place where Facebook samples report *lower* activity than corresponding barometer estimates. Moreover, the Facebook sample estimates are closer to the official turnout figures reported in each country (indicated by the green crosses in Figure 4).

Consideration of political context reinforces the idea that different estimates of political activity may be in part attributable to mode effects. The Asian Barometer survey was fielded in July 2019 in Indonesia, just a few months after President Joko Widodo's reelection sparked protests and claims of election fraud from the opposition Gerindra party candidate (Suhartono and Victor, 2019). In this context, only 1% of Asian Barometer respondents reported affiliation with the Gerindra party candidate, who received 44% of the vote. In the Facebook sample, many more respondents (16%) reported affiliation with the Gerindra party. Respondents may have been less willing to express their true political views to an enumerator.

To illuminate the platform's utility for assessing policy views, we examine environmental and climate change opinions in Mexico and Indonesia. In Mexico, we replicated the question wording and response options from a LAPOP question regarding whether economic growth or environmental protection should be prioritized. Facebook respondents were more likely than LAPOP respondents to answer that environmental protection should be given the highest priority (44% vs. 20%) and less likely to choose responses at the economic-growth end of the Likert scale (9% vs. 18%) (Figure 5, left panel). In Indonesia, we asked multiple questions about climate change to our Facebook-recruited respondents and to respondents recruited from Dynata's online panel. Facebook respondents were less worried about climate change and less likely to understand that it is human-caused, compared with the Dynata sample (Figure 5, right panel). These differences likely stem from sample composition. The Facebook sample is more highly educated than the LAPOP sample in Mexico but less highly educated than the Dynata sample in Indonesia, and education is positively correlated with environmental concern in global surveys (Lee *et al.*, 2015).

6. Survey costs

Researchers typically weigh survey quality concerns against the relative costs of different survey tools, and we find that Facebook sampling is quite cost-effective. The costs of Facebook sampling include platform advertisements and (optional) incentives paid to respondents. Not including incentives,

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

the mean cost per completed survey was $0.16 in Mexico, $0.85 in Kenya, and $0.91 in Indonesia.[8] Including incentives, the surveys cost an average of $1.03 per completed survey (ranging from $0.16 in Mexico to $1.57 in Indonesia). This is incredibly inexpensive in Mexico. In Kenya or Indonesia, the cost is comparable with the cost of recruiting online convenience samples using platforms such as MTurk or Cint. Of course, MTurk, Cint, or similar platforms do not enable researchers to contact survey respondents in every country, and their user populations are much smaller than Facebook's. Realistically, in-person field surveys or online panels provide the most feasible alternative in most parts of the Global South. The costs of our Facebook samples are substantially cheaper than these alternatives. For instance, the Indonesia Dynata sample ($n = 1{,}130$) cost $5.75 per completed survey. In-person surveys are even more expensive to field.

7. Conclusion

Facebook is the most popular social media platform in the world with 3 billion monthly users as of 2024 (Dixon, 2023). It's extensive user base presents an opportunity for researchers to quickly and reliably recruit subjects globally, including from contexts underrepresented on existing online subject recruitment platforms. Our replication of the "disease problem" experiment suggests that, similar to other low-cost online survey platforms (Berinsky *et al.*, 2012; Mullinix *et al.*, 2015; Coppock and McClellan, 2019), Facebook is a cost-effective tool for recruiting survey-experimental subjects. The platform's near-global reach makes it particularly useful where other online recruitment platforms are unavailable. Our analysis evaluates the quality of Facebook-recruited samples in three countries across the Global South, highlighting some of the advantages and shortcomings of this method and providing practical guidance to researchers.

We have assessed total survey error and its components, to provide practical insights into where and how researchers can minimize bias in estimates derived from Facebook-recruited surveys. Coverage error favors respondents who have achieved higher levels of education, on average, than the general population. This overrepresentation of highly educated respondents is a structural feature of the Facebook platform in our case countries. By extension, researchers should investigate these imbalances in other target countries before deciding to use Facebook as a survey recruitment tool. Researchers can mitigate bias associated with coverage error by targeting recruitment resources toward respondents who are underrepresented on Facebook. In our case, quota sampling through the ad platform allowed recruitment of a more representative sample than leaving the advertising algorithm to its own devices. However, we do find some evidence of sampling bias in Indonesia, and up-weighting the underrepresented individuals introduces adjustment bias in our experimental results from Indonesia. Together, these findings highlight that researchers should monitor entries into the survey from each stratum and adjust their advertising campaign to ensure recruitment in all strata. Still, the back-end demographic data that Facebook uses for targeting are noisy, particularly for education and location. As a result, quota sampling cannot substitute for the use of design weights based on self-reported demographics. In general, weights reduced demographic imbalances and affected descriptive inferences only slightly. Our weighted estimates tended toward benchmark estimates, compared with unweighted estimates. Researchers should also assess nonresponse rates across quota cells during data collection to inform an iterative sampling strategy that maximizes response rates among high-attrition groups.

Practically, Facebook recruitment costs can be quite low but vary by targeting strategy and survey incentives. In Mexico, we recruited respondents without incentives, while in Kenya and Indonesia we provided a modest airtime credit to encourage participation. This incentive may have encouraged

[8]These costs are based on Facebook's count of completed surveys. The SI provides details about the reach of our survey as advertised on Facebook and the cost of reaching different demographic groups. These costs vary because Facebook ads are deployed using a bidding system in which hard-to-reach populations are more expensive to target.

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

greater participation among resource-constrained respondents, but it also led to instances of gaming and link sharing. Future research could explore sampling strategies that leverage Facebook's social nature through snowball sampling.

The effectiveness of Facebook recruitment for social science research in a particular context depends on several factors. The platform will be particularly successful where (1) phone and internet penetration are widespread, (2) literacy rates are high, and (3) recent census (or other benchmark) population data are available to enable weighting. Researchers should also consider how users' interactions with and expectations of Facebook might influence internal validity. While we tested this method in competitive democracies, concerns about government surveillance might influence responses in authoritarian states. Researchers should consider these concerns in the survey design process, and future research could consider how internal validity varies across political contexts. The popularity of Facebook also changes over time, and the platform may not retain the large and broad user base that makes it an attractive recruitment platform. Changes to Meta's advertising policies may also change researchers' ability to target ads. Such changes are beyond researchers' control but within their ability to monitor in order to use the platform wisely. Moreover, our conceptual framework can be applied to other platforms that may supplant Facebook in the future.

Of course, Facebook should not replace gold-standard, in-person field surveys for recruiting nationally representative samples in the Global South. Our Facebook samples tend to be more educated and slightly younger than the general population, skewing descriptive inferences on political engagement and public policy views. Nonetheless, Facebook outperforms a commercial online survey firm, at considerably lower cost. Researchers can refine their sampling approach, model nonresponse bias, and apply weighting to mitigate sample biases. We further recommend that researchers employ benchmark comparisons similar to those we have performed here. Used effectively, Facebook can be a valuable tool for expanding public opinion research.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2025.18. To obtain replication material for this article, https://doi.org/10.7910/DVN/IXILAR.

Acknowledgements. Leah R. Rosenzweig is a Senior Fellow at the Center for Global Development (lrosenzweig@cgdev.org). Parrish Bergquist is Assistant Professor of Political Science, University of Pennsylvania (pberg@upenn.edu). Katherine Hoffmann Pham is AI Advisor at UNICEF's Office of Innovation (note: this work was conducted independently of UNICEF and does not represent the views or interests of UNICEF or any other UN entity). Francesco Rampazzo is Lecturer in Demography, Department of Sociology, Leverhulme Centre for Demographic Science, and Nuffield College, University of Oxford. Matto Mildenberger is Associate Professor of Political Science, University of California Santa Barbara. The authors thank Chris Bail, Matt Salganik, Anne Helby Petersen, Julien Migozzi, and Tina Law for their assistance in shaping this project. We thank Michaël Aklin, André Grow, Peter Howe, Anthony Leiserowitz, Jennifer Marlon, Umberto Mignozzetti, Blair Read, Baobao Zhang, Ingmar Weber, Alessandro Sorichetta, Dennis Feehan, Chris Tausanovitch, researchers at the Busara Center for Behavioral Economics, Afrobarometer, and seminar participants at NYU Abu Dhabi for helpful comments. Thanks to Emma Franzblau, Gabriel de Roche, and Ingmar Sturm for research assistance. We thank Warsama Abdifitah, Kibuchi Eliud, Ahmed Hared, Lilian Ligeyo, Laban Okune, Gustavo Ovando-Montejo, Nelson Ngige, and Eunice Williams for translation assistance. This work was supported by the Summer Institute for Computational Social Science (SICSS), the Russell Sage Foundation, the Alfred P. Sloan Foundation, the Yale Program on Climate Change Communication, and Meta. To view supplementary material for this article, please visit http://dx.doi.org/10.1017/psrm.2014.11.

Competing interests. The authors have none to declare.

References

Afrobarometer (2019) Afrobarometer Wave 8. http://www.afrobarometer.org

Ansolabehere S and Schaffner B (2010) Does survey mode still matter? *Political Analysis* **22**(3), 285–303.

Ansolabehere S, Schaffner B and Shih M (2023) Guide to the 2022 cooperative election study. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi%3A10.7910/DVN/PR4L8P

Araujo M, Mejova Y, Weber I and Benevenuto F (2017) Using Facebook ads audiences for global lifestyle disease surveillance: Promises and limitations. In *Proceedings of the 2017 ACM on Web Science Conference*, WebSci'17. pp. 253–257. Troy, New York, USA: ACM. https://doi.org/10.1145/3091478.3091513

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

Asian Barometer (2019) Asian Barometer Wave 5. http://www.asianbarometer.org/index
Bailey M (2024) *Polling at a Crossroads*. Cambridge: Cambridge University Press.
Berinsky A, Huber G and Lenz G (2012) Evaluating online labor markets for experimental research. *Political Analysis* **20**, 351–368.
Bicalho C, Platas M and Rosenzweig LR (2020) "If we move, it moves with us:" Physical distancing in Africa during COVID-19. https://static1.squarespace.com/static/5410fc5ae4b0b9bdbd0cde82/t/5f0b12f8edb3507adfee3b0f/1594561273957/Social_distancing_africa_BPR2020.pdf
Boas TC, Christenson DP and Glick DM (2020) Recruiting large online samples in the United States and India: Facebook, mechanical turk, and qualtrics. *Political Science Research and Methods* **8**(2), 232–250.
Bond R and Messing S (2015) Quantifying social media's political space: Estimating ideology from publicly revealed preferences on Facebook. *American Political Science Review* **109**, 62–78.
BPS Statistics Indonesia (2020) Comprehensive documentation population census 2020. https://www.bps.go.id/en/publication/2023/12/22/7c6f371f7156b1230377fc96/comprehensive-documentation-population-census-2020.html.
Broockman D and Green D (2014) Do online advertisements increase political candidates' name recognition or favorability? Evidence from randomized field experiments. *Political Behavior* **36**, 263–289.
Coppock A and McClellan O (2019) Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics* **6** https://doi.org/10.1177/2053168018822174.
DeBell M, Amsbary M, Brader T, Brock S, Good C, Kamens J, Maisel N and Pinto S (2020) Methodology report for the ANES time series study. https://electionstudies.org/wp-content/uploads/2022/08/anes_timeseries_2020_methodology_report.pdf
Deming E (1944) On errors in surveys. *American Sociological Review* **9**, 359–369.
Dixon Stacy Jo (2023) Number of monthly active Facebook users worldwide as of 4th quarter 2023. https://www.statista.com/statistics/264810/number-ofmonthly-active-facebook-users-worldwide/ (accessed 20 February 2025).
Finkel SE, Neundorf A and Rascon Ramirez E (2023) Can online civic education induce democratic citizenship? *American Journal of Political Science* **68**(2), 613-630.
Gomes M, Radice R, Camarena Brenes J and Marra G (2019) Copula selection models for non-Gaussian outcomes that are missing not at random. *Statistics in Medicine* **38**, 480–496.
Grewal S (2023) Military repression and restraint in Algeria. *American Political Science Review* **118**(2), 1–16.
Groves R, Fowler Jr F, Couper M, Lepkowski J, Singer E and Tourangeau R (2011) *Survey Methodology*, Vol. 561. Hoboken, NJ: John Wiley & Sons.
Groves R and Lyberg L (2010) Total survey error: Past, present, and future. *Public Opinion Quarterly* **74**, 849–879.
Grow A, Perrotta D, Del Fava E, Cimentada J, Rampazzo F, Gil-Clavel S and Zagheni E (2020) Addressing public health emergencies via Facebook surveys. *Journal of Medical Internet Research* **22**, e20653.
GSMA (2023) The mobile economy 2023. *Technical Report*. https://www.gsma.com/mobileeconomy/wp-content/uploads/2023/03/270223-The-Mobile-Economy-2023.pdf
Hartman E and Huang M (2024) Sensitivity analysis for survey weights. *Political Analysis* **32**, 1–16.
Hirano S, Lenz G, Pinkovskiy M and Snyder JM (2015) Voter learning in state primary elections. *American Journal of Political Science* **59**, 91–108.
Holbrook A and Krosnick J (2010) Social desirability bias in voter turnout reports. *Public Opinion Quarterly* **74**, 37–67.
Holliday D, Reny T, Rossell Hayes A, Rudkin A, Tausanovitch C and Vavreck L (2021) Democracy fund + UCLA nationscape methodology and representativeness assessment. https://www.voterstudygroup.org/uploads/reports/Data/Nationscape-Methodology-Representative-Assessment-2021Dec.pdf
Huff C and Tingley D (2015) 'Who are these people?' Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics* **2**(3), 1–12.
Im H and Chen C (2022) To save or lose? A cross-national examination of the disease risk framing effect and the influence of collectivism. *Journal of Behavioral Decision Making* **35**, e2276.
INEGI (2015) Population. http://en.www.inegi.org.mx/temas/estructura/default.html#Tabulados (accessed 20 August 2019).
Jackman S and Spahn B (2019) Why does the American National Election Study overestimate voter turnout? *Political Analysis* **27**, 193–207.
Jacobson S (2024) Policies can still create new politics: Contemporary causal evidence from New York Preschool parents. https://doi.org/10.31219/osf.io/9nfg7.
Jäger K (2017) The potential of online sampling for studying political activists around the world and across time. *Political Analysis* **25**, 329–343.
Kapp J, Peters C and Oliver D (2013) Research recruitment using Facebook advertising: Big potential, big challenges. *Journal of Cancer Education* **28**, 134–137.
Kennedy Courtney, Hatley Nick, Keeter Scott, Mercer Andrew, Igielnik Ruth and Traylor Frederic (2018) Comparing Survey Sampling Strategies: Random-Digit Dial vs. Voter Files Washington, DC: Pew Research Center. https://www.pewresearch.org/methods/wp-content/uploads/sites/10/2018/10/REPORT-Comparing-RDD-and-RBS-Surveys-Final-05.06.19-2.pdf

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press

Kilavuz T, Grewal S and Kubinec R (2023) Ghosts of the black decade. *Journal of Peace Research* **60**, 9–25.
KNBS (2010) 2019 population and housing census. http://www.knbs.or.ke.
Kosinski M, Matz S, Gosling S, Popov V and Stillwell D (2015) Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist* **70**, 543.
LAPOP (2018–2019) LAPOP Mexico. https://www.vanderbilt.edu/lapop/mexico.php (accessed 16 October 2019).
Lee TM, Markowitz E, Howe P, Ko C-Y and Leiserowitz AA (2015) Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change* **5**, 1014–1020.
Litman L, Robinson J and Abberbock T (2017) Turkprime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods* **49**, 433–442.
Lyberg L and Weisberg H (2016) Total survey error: A paradigm for survey methodology. In Wolf C, Joye D, Smith TW and Fu Y, *The Sage Handbook of Survey Methodology*. Sage Thousand Oaks, CA: Sage, pp. 27–35.
Manski CF (1990) Nonparametric bounds on treatment effects. *American Economic Review* **80**, 319–323.
McGovern M, Canning D and Bärnighausen T (2018) Accounting for non-response bias using participation incentives and survey design. *Economics Letters* **171**, 239–244.
Meta (2023) Meta reports second quarter 2023 results. https://investor.fb.com/investor-news/press-release-details/2023/Meta-Reports-Second-Quarter-2023-Results/default.aspx (accessed 18 August 2023).
Mullinix KJ, Leeper TJ, Druckman JN and Freese J (2015) The generalizability of survey experiments. *Journal of Experimental Political Science* **2**, 109–138.
Neundorf A and Öztürk A (2023) How to improve representativeness and cost-effectiveness in samples recruited through meta: A comparison of advertisement tools. *PLoS One* **18**, e0281243.
Noh Y, Grewal S and Kilavuz T (2023) Regime support and gender quotas in autocracies. *American Political Science Review* **118**(2), 1–18.
Offer-Westort M, Rosenzweig L and Athey S (2024) Battling the Coronavirus 'Infodemic' Among Social Media Users in Kenya and Nigeria. *Nature Human Behavior* **8**, 823-834.
Ortiz-Espina E (2019) The rise of social media. https://ourworldindata.org/rise-of-social-media (accessed 15 August 2024).
Pötzschke S and Braun M (2017) Migrant sampling using Facebook advertisements: A case study of polish migrants in four European countries. *Social Science Computer Review* **35**, 633–653. https://doi.org/10.1177/0894439316666262
Ramo D, Rodriguez T, Chavez K, Sommer M and Prochaska J (2014) Facebook recruitment of young adult smokers for a cessation trial: Methods, metrics, and lessons learned. *Internet Interventions* **1**, 58–64.
Rosenzweig LR and Zhou LR (2021) Team and nation: Sports, nationalism, and attitudes toward refugees. *Comparative Political Studies* **54**, 2123–2154.
Rotondi V, Kashyap R, Pesando LM, Spinelli S and Billari FC (2020) Leveraging Mobile Phones to Attain Sustainable Development. *Proceedings of the National Academy of Sciences* **117**(24), 13413–13420. https://doi.org/10.1073/pnas.1909326117.
Ruggeri K, Alí S, Berge ML, Bertoldo G, Bjørndal LD, Cortijos-Bernabeu A, Davison C, Demić E, Esteban-Serna C, Friedemann M and Gibson SP (2020) Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour* **4**, 622–633.
Ryan T (2012) What makes us click? Demonstrating incentives for angry discourse with digital-age field experiments. *The Journal of Politics* **74**, 1138–1152.
Samuels D and Zucco Jr C (2014) The power of partisanship in Brazil. *American Journal of Political Science* **58**, 212–225.
Santoso LP, Stein R and Stevenson R (2016) Survey experiments with Google consumer surveys. *Political Analysis* **24**, 356–373.
Suhartono M and Victor D (2019) Violence erupts in Indonesia's capital in wake of presidential election returns. *New York Times*.
Sun B, Liu L, Miao W, Wirth K, Robins J and Tchetgen EJT (1965) Semiparametric estimation with data missing not at random using an instrumental variable. *Statistica Sinica* **28**, 1965–1983.
Ternovski J and Orr L (2022) A note on increases in inattentive online survey-takers since 2020. *Journal of Quantitative Description: Digital Media* 2.
Tversky A and Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* **211**, 453–458.
Verba S, Schlozman K and Brady H (1995) *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, MA: Harvard University Press.
Zhang B, Mildenberger M, Howe P, Marlon J, Rosenthal SA and Leiserowitz A (2020) Quota sampling using Facebook advertisements. *Political Science Research and Methods* **8**(3), 558–564.

Cite this article: Rosenzweig LR, Bergquist P, Pham KH, Rampazzo F and Mildenberger M (2025) Survey sampling in the Global South using Facebook advertisements. *Political Science Research and Methods* **13**, 781–797. https://doi.org/10.1017/psrm.2025.18

https://doi.org/10.1017/psrm.2025.18 Published online by Cambridge University Press