The European Political Science Association, 2016 *doi:10.1017/psrm.2016.XX*

Quota Sampling Using Facebook Advertisements *

Baobao Zhang[1], Matto Mildenberger[2], Peter D. Howe[3], Jennifer Marlon[4], Seth Rosenthal[4], and Anthony Leiserowitz[4]

[1]Department of Political Science, Yale University
[2]Department of Political Science, University of California Santa Barbara
[3]Department of Environment and Society, Utah State University
[4]School of Forestry and Environmental Studies, Yale University

Researchers in different social science disciplines have successfully used Facebook to recruit subjects for their studies. However, such convenience samples are not generally representative of the population. We develop and validate a new quota sampling method to recruit respondents using Facebook advertisements, and publish an R package to semi-automate this quota sampling process using the Facebook Marketing API. To test the method, we used Facebook advertisements to quota sample 2432 U.S. respondents for a survey on climate change public opinion. We conducted a contemporaneous nationally representative survey asking identical questions using a high-quality online survey panel whose respondents were recruited using probability sampling. Many results from the Facebook-sampled survey are similar to those from the online panel survey; furthermore, results from the Facebook-sampled survey approximate results from the American Community Survey (ACS) for a set of validation questions. These findings suggest that using Facebook to recruit respondents is a viable option for survey researchers wishing to approximate population-level public opinion.

*We would like to thank Erin Hartman, Devin Caughey, Peter Aronow, Solomon Messing, and M. Kent Jennings for their feedback on earlier drafts of this research. We also appreciate feedback from participants of the following workshops: the New Faces in Political Methodology IX Workshop (Pennsylvania State University), the Oxford Internet Institute Summer Doctoral Programme, and the Southern California Methods Workshop. We also thank the Skoll Global Threats Fund and the MacArthur Foundation for their support of the project. BZ's work on this project was supported by the National Science Foundation Graduate Research Fellowship. Corresponding author: baobao.zhang@yale.edu

Survey researchers face a fundamental cost-bias tradeoff in survey recruitment. Fielding opinion polls on high-quality probability samples entail substantial costs that limit the scale and breadth of research activity. Recognizing the limits of using university-based convenience samples, many researchers have examined whether web-based crowd-sourcing tools like Amazon's Mechanical Turk and Google Consumer Surveys can be cost-effective methods to survey diverse target populations (Berinsky, Huber, and Lenz 2012; Huff and Tingley 2015; Santoso, Stein, and Stevenson 2016). The empirical properties of these internet-based samples remains a subject of active study.

In this article, we evaluate Facebook's potential as a platform for survey recruitment. We propose a quota sampling method using Facebook advertisements to generate public opinion estimates that approximate national averages efficiently. Using a proof-of-concept study on U.S. climate opinions, we demonstrate that researchers can cost-effectively recruit respondents through quota sampling using Facebook advertisements, at a fraction of the cost of hiring an online survey firm (we sampled over 2000 respondents at about $4 per response). Using Facebook for quota sampling also offers a comparative advantage because it gives researchers control of how they recruit subjects and allows researchers to target specific subpopulations.

Our method contributes to an emerging literature that uses Facebook to recruit respondents into research studies. For example, psychologists enrolled over 4 million Facebook users as subjects by creating a Facebook application that allowed users to take psychometric tests (Kosinski et al. 2015). The researchers subsequently linked these tests to users' Facebook profiles to predict personal attributes and personality traits from users' social media behavior (Kosinski, Stillwell, and Graepel 2013; Youyou, Kosinski, and Stillwell 2015). Medical researchers have similarly used Facebook to recruit subjects from specific subpopulations, such as young adults who smoke cigarettes (Ramo and Prochaska 2012; Ramo et al. 2014) or American women aged 35-49 (Kapp, Peters, and Oliver 2013).

Political science researchers have used Facebook to recruit respondents in Brazil for survey experiments (Samuels and Zucco 2013, 2014), survey likely voters in state primaries (Hirano et al. 2015), deliver political advertisement treatments to diverse respondent pools (Ryan 2012; Broockman and Green 2014) and survey political activists in Germany and Thailand (Jäger 2017). In Samuels and Zucco (2013, 2014), researchers raffled off an iPad to one randomly-selected survey-taker. For a cost of $1.86 per valid survey completion, the researchers obtained 3286 responses. Although the Facebook sample was less representative than a national sample from a survey firm, the results of the Facebook-sampled survey experiment were similar, especially after weighting the responses through post-stratification. Likewise, Jäger (2017) used Facebook samples of local party activists to generate accurate estimates of both a German party leadership race and a local leaders survey.

We contribute to this literature by developing a platform-specific quota sampling technique. Our method uses the Facebook Marketing API to semi-automate a quota sampling process. This method allowed us to recruit 2432 respondents at the cost of

$4.05 per completed response. This approach allows quota sampling on conditional strata themselves rather than just ensuring the sample marginals approximate the population marginals. Results from the Facebook-sampled survey approximate results from a survey conducted by a reputable survey research firm, GfK[1], as well as results from the American Community Survey (ACS). To accompany this research note, we are releasing an R package, `fbsample`, that automates the quota sampling process using the Facebook Marketing API.[2]

Methodology

To obtain the most representative sample given a constrained advertisement budget, we employ quota sampling to target a series of strata, or specific demographic groups. Convenience sampling methods that do not set demographic quotas might potentially produce very non-representative samples. Facebook states that "if you optimize for link clicks, your ads are targeted to people in your audience who are most likely to click the ads' links."[3] When a user first launches an advertisement optimized for link clicks, Facebook will send the advertisement to different types of people to learn who is most likely to click on the link. After this "learning phase" (after about 50 link clicks), Facebook targets the advertisements to those predicted to click on the link.[4] For public opinion researchers, Facebook's advertisement delivery optimization might prove problematic because it could lead to the recruited subjects being homogeneous. For instance, if many young white males clicked on the advertised link during the "learning phase," Facebook may target mostly young white males. To avoid this "homogenization" of their sample, public opinion researchers must target a diversity of demographic strata.

The Facebook Marketing API makes it relatively easy to target specific demographic subgroups. Before 2015, advertisers had to set up advertisements that target specific groups manually. However, after Facebook made its Marketing API publicly available in 2015, advertisers gained the ability to write programs to target many pre-specified groups at once. Our current method uses this capability to quota sample many demographically diverse strata simultaneously.

However, efforts to estimate national public opinion from Facebook quota samples

[1]GfK recruits respondents for its online panel through address-based probability-sampling. Organizations that have used GfK's panel include the American National Election Study, the U.S. Federal Reserve, and Pew Research Center.

[2]`fbsample` is available online at `https://github.com/13bzhang/fbsample`.

[3]Facebook, "About the Delivery system: Optimization." `https://www.facebook.com/business/help/355670007911605?helpref=faq_content#`

[4]Facebook, "About the delivery system: Learning phase." `https://www.facebook.com/business/help/112167992830700?helpref=faq_content`

depend on a series of assumptions, some of which may not hold in reality. Define R_i as whether member i of the population takes the survey, $\mathbf{D}_i$ as the vector of respondent i's characteristics that researchers used to construct strata for quota sampling, and $\mathbf{X}_i$ as the vector of respondent i's characteristics not used to construct strata but are measured in the survey. Furthermore, define Y_i as person i's response to a survey question.

ASSUMPTION 1: $Y_i \perp\!\!\!\perp R_i | \mathbf{D}_i = \mathbf{d}_i, \mathbf{X}_i = \mathbf{x}_i, \forall\ \mathbf{d}_i \in \text{Supp}(\mathbf{D}_i)$ and $\forall\ \mathbf{x}_i \in \text{Supp}(\mathbf{X}_i)$

ASSUMPTION 2: $\Pr(R_i = 1 | \mathbf{D}_i = \mathbf{d}_i, \mathbf{X}_i = \mathbf{x}_i) > 0, \forall\ \mathbf{d}_i \in \text{Supp}(\mathbf{D}_i)$ and $\forall\ \mathbf{x}_i \in \text{Supp}(\mathbf{X}_i)$

Assumption 1 implies that conditional on strata and observed respondent characteristics, responses of those who took the survey would be the same as those who did not take the survey. While this assumption might not hold in reality, the degree to which violations of Assumption 1 will bias efforts to use Facebook advertisements to estimate national public opinion is an empirical question that this research note is designed to address.

Assumption 2 asserts that conditional on strata and observable characteristics, each person in the population has a non-zero probability of taking the survey. Of course, not everyone in a given population has a Facebook account or uses Facebook regularly. According to a 2015 Pew study, 72 percent of American adult internet users use Facebook, and 62 percent of the U.S. adult population uses Facebook[5]. Even so, Facebook allows one to reach a far greater proportion of the U.S. adult population than other ways researchers recruit convenience samples. For instance, researchers recruiting respondents using Mechanical Turk can only reach about 7300 workers in any quarter year (Stewart et al. 2015).

Study Design

To demonstrate the validity of the quota sampling method described above, we conducted a study to compare the results of a quota-sampled Facebook survey with a high-quality probability-sampled GfK online panel survey. The two surveys, which shared 25 identical questions, were conducted a few months apart in 2016. Apart from the questions asking about respondent demographics, both surveys focused on climate change attitudes and policy support. Also, to validate the Facebook sample against the ACS, we asked three questions from the ACS about veteran status, home ownership, and country of birth. (These results are reported in the Online Supplementary Information).

The Facebook Marketing API allows researchers to write programs that target several thousand highly specific demographic groups at once. For our quota sampling, we

[5]Duggan, Maeve. "The Demographics of Social Media Users." Pew Research Center, 2015. http://www.pewinternet.org/2015/08/19/the-demographics-of-social-media-users/

generated 544 strata according to demographic characteristics (e.g., gender, age group, race, level of education, and nine U.S. Census regions). We chose these demographic characteristics because crosstabs containing frequencies conditional on these demographics are readily available through the U.S. Census.[6]

Next, we assigned an advertising budget to each stratum according to proportional allocation for most strata; we allocated greater budget to strata that contain very small sub-populations nationally for two reasons. First, Facebook requires advertisers to set a daily budget of $5.00 per advertisement, although it is not required to spend the entire daily budget. Second, since it is more difficult to recruit from small sub-populations, we had to allocate a larger budget to generate the minimum number of respondents.

Our advertisements informed respondents that by taking the survey they could see how their views on climate change compared to other Americans; no rewards were promised. To maximize viewership, we promoted the advertisements on all Facebook platforms (e.g., news feed, right column advertisement space, etc.), except on Instagram.[7] Over a period of two weeks in July 2016, we recruited 2432 respondents who completed the survey from the 7642 Facebook users who clicked on our advertisement. At the end of the recruitment period, we filled or overfilled 218 strata, partially filled 61 strata, and failed to recruit anyone for 157 strata. On average, the Facebook survey cost $4.05 per complete.[8]

Our probability-sampled online panel study was conducted in March 2016. This survey sampled 2459 respondents with 1346 completions for a completion rate of 54.7%.[9] This study used GfK's Knowledge Panel, with a household recruitment rate for our survey of 12.3% (AAPOR Response Rate 3). Overall, our online panel responses cost nearly six

[6]We cannot verify whether Facebook's internal algorithm can accurately infer users' non-self-reported demographic information through content generated by users. However, Kosinski, Stillwell, and Graepel 2013 suggests that one could accurately predict users' age, ethnicity, gender, and other characteristics using users' Facebook Likes. We plan to validate Facebook's demographic labels against self-reported data in the future.

[7]During the respondent recruitment period, we activated the advertisements in the morning and deactivated them after 8 p.m. EST. After each day of recruitment, we removed advertisements for strata whose quotas had been filled. This allowed some strata to be overfilled, wasting some advertisement budget. We discuss potential solutions to this problem in the Online Supplementary Appendix.

[8]We did not offer monetary incentives. Instead, we told respondents that by completing our survey they would discover how their views on climate change compare with views of other Americans. The average cost per link click was $2.16 (SD $1.73); the distribution of costs per link click is reported in the Online Supplementary Information.

[9]GfK dropped cases where more than half the survey questions were left blank and/or where respondents completed the survey in under 7 minutes, generating a final 1317 observations.

times as much as our Facebook user responses.

RESULTS

Results from the Facebook survey are similar to results from the online panel survey and the ACS even without weighting. For questions where the Facebook and online panel survey results differed, the Facebook respondents produced answers indicating greater concern about climate change.

Table 1 displays the summary statistics for the demographic groups in the GfK, Facebook, and ACS surveys. Contrasted with the GfK and ACS samples, the Facebook sample is younger and less white. To make both surveys more nationally representative, we used inverse probability weighting to weight each sample to the March 2016 Current Population Survey (CPS).[10]

We combine each sample and the CPS, then we use logistic regression to estimate the probability of being included in the sample. Covariates used in our propensity score model included gender, age group, level of education, race, geographic region, whether the respondents lived in a metropolitan area, and the interaction between region and the metropolitan indicator variable. The final weights are the inverse of the estimated probabilities normalized such that the sum of each sample's weights equals the sample size.[11]

As Table 1 demonstrates, the weighted demographic summary statistics of the two surveys are very similar to each other and to the ACS estimates. Comparing the political variables, the Facebook sample contains a greater proportion of Independents and a smaller proportion of Republicans than the GfK sample.

The results for the eight questions about climate change are presented in Figure 1. For the unweighted results, the mean difference is 5.1 percentage points with a standard deviation of 4.2 percentage points. Regarding these unweighted results, the Facebook-sampled respondents expressed slightly less skepticism about climate change, greater concern about climate change, and greater support for policies that mitigate climate change. After weighting both the Facebook survey and the online panel survey to the CPS, the mean difference in responses is 4.3 percentage points with a standard deviation of 3.4 percentage points.

One reason that the Facebook-sampled respondents may be more concerned about

[10]We take an agnostic stance on the ideal weighting strategy. Our assessment is thus driven by the empirical properties of our estimates during validation, rather than evaluating the merits of different weighting strategies.

[11]The correlation between the inverse probability weights for the GfK sample and the weights supplied by GfK is 0.64.

TABLE 1 *Summary Statistics of Survey Respondents*

	GfK	Facebook	ACS	GfK	Facebook	ACS
Weighted	No	No	No	Yes	Yes	Yes
Prop. Male	0.50	0.48	0.48	0.48	0.47	0.49
Prop. Age 18-24	0.08	0.19	0.11	0.12	0.12	0.12
Prop. Age 25-34	0.14	0.16	0.15	0.17	0.20	0.18
Prop. Age 35-44	0.14	0.15	0.15	0.17	0.17	0.16
Prop. Age 45-64	0.39	0.29	0.35	0.35	0.33	0.34
Prop. Age 65 and Above	0.26	0.21	0.24	0.19	0.18	0.20
Prop. with College Degree	0.21	0.23	0.18	0.20	0.19	0.18
Prop. with Graduate Degree	0.16	0.16	0.11	0.11	0.12	0.11
Prop. Hispanic	0.11	0.13	0.13	0.15	0.13	0.16
Prop. Black	0.09	0.06	0.10	0.11	0.11	0.12
Prop. White	0.74	0.61	0.78	0.66	0.68	0.74
Prop. Democrat	0.43	0.42		0.44	0.41	
Prop. Republican	0.36	0.27		0.33	0.27	
Prop. Independent	0.10	0.18		0.10	0.17	
Mean Political Ideology	3.05	2.77		3.04	2.76	
Prop. Registered to Vote	0.83	0.81		0.79	0.81	

The GfK survey included 1317 respondents; the Facebook survey included 2432 respondents; the 2016 American Community Survey (ACS) Public Use Microdata Sample include 2,503,750 respondents. In the GfK and Facebook surveys, respondents who indicated they "lean Democrat" are counted as Democrats; likewise, "lean Republican" was counted as Republican. Political ideology is measured using a 5 point scale where 1 is very liberal and 5 is very conservative.

climate change is that our advertisement explicitly contained language about climate change and included a photo of the Earth. Respondents who are pro-climate action might have been more motivated to click on the advertisement than those who are unconcerned about climate change. A breakdown of the survey results by partisanship revealed that Democrats, Republicans, and Independents within the Facebook survey provided similar responses to Democrats, Republicans, and Independents in the online panel survey, respectively. However, respondents in the Facebook survey who said they are uninterested in politics were significantly more pro-climate action than their counterparts in the online panel survey. This difference between the two samples suggests that within each stratum, Facebook users may not be as-if randomly sorting into the recruitment sample. One possible strategy to avoid self-sorting that affects survey results is to create advertisements that do not discuss the content of the survey being advertised, although this strategy will require careful attention to advertisement wording and design.[12]

As a further robustness check, we also compare three results from the Facebook survey with the 2016 ACS One-Year Estimates. The Facebook sample is also able to somewhat approximate the ACS benchmarks, as detailed in our Online Supplementary Information.

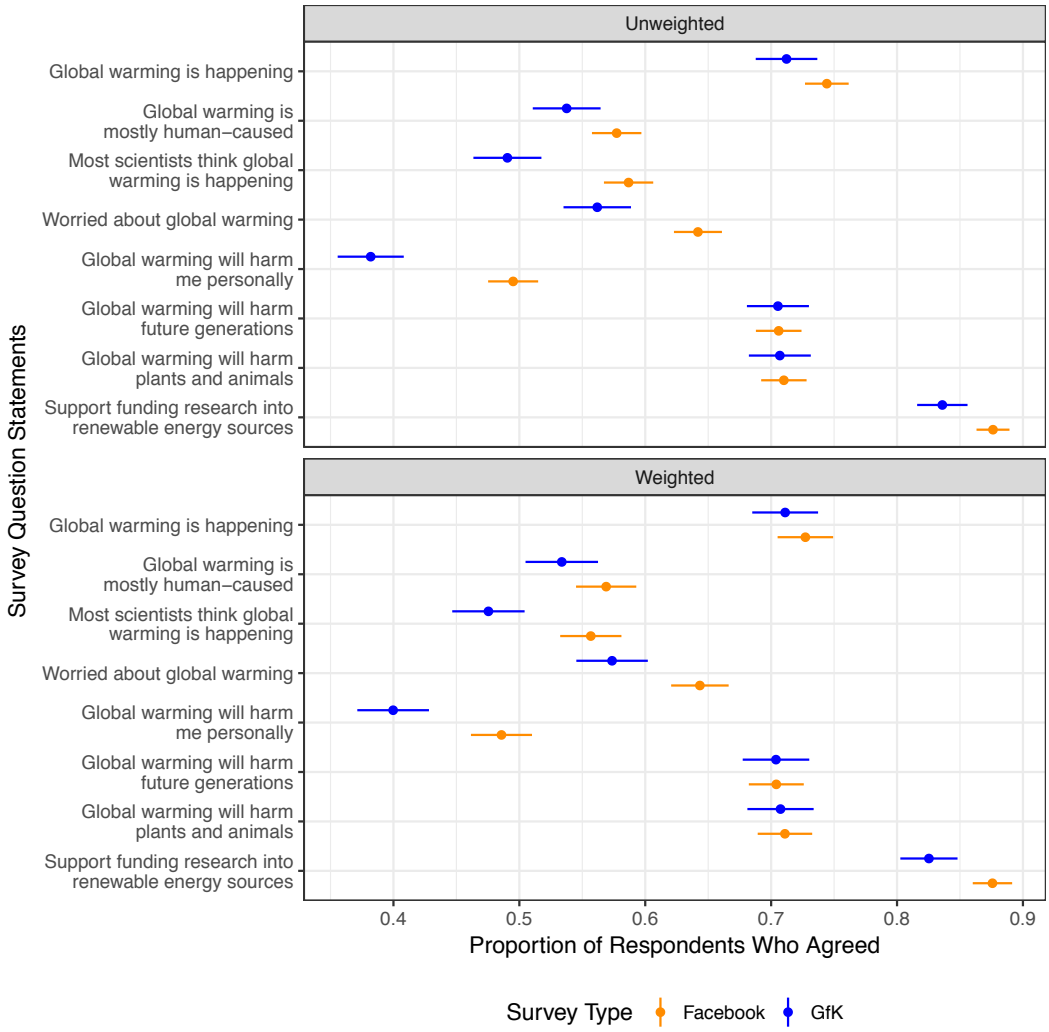
Conclusion

This study produced valuable lessons for best practices in quota sampling using Facebook advertisements; this method can generate results that approximate high quality probability-based national opinion surveys. Further improvements to our proposed method would result from systematic attention to the factors that drive Facebook users to click on advertisements. We highlight some considerations for future researchers. First, researchers might inadvertently recruit particular types of respondents by advertising the content of the survey. This form of self-selection bias may not be eliminated by strata targeting or by conditioning on observable demographic characteristics. To avoid this problem, researchers can advertise their survey using vague language without revealing its core content. Additional design effort may be required, however, to persuade users to click on such advertisements.

Future researchers could also further reduce bias in their estimates by measuring other characteristics of the respondents and using those characteristics to weight their Facebook-sampled survey. In particular, we suggest the utility of obtaining respondent characteristics that are not available through Facebook, such as political party identification,

[12]During the testing phase of this study, we launched advertisements for a generic public opinion survey. These advertisements received very few clicks. We decided to launch instead the advertisements that contained information about climate change to generate a sufficient number of clicks.

Figure 1. Comparing Facebook Survey with GfK Survey: Climate Change Public Opinion



The plots above report the point estimate and the 95 percent confidence interval for each outcome measure. The confidence intervals are calculated from heteroscedasticity-consistent standard errors.

to improve the quality of re-weighting efforts.[13]

More broadly, our findings suggest that using Facebook to recruit respondents is a viable option for survey researchers seeking to approximate public opinion estimates for some populations at significantly lower cost. While our method costs a fraction of recruiting respondents through survey panels that use probability sampling, it is not dramatically cheaper than using other "opt-in" online survey panels. Nevertheless, our method provides greater flexibility to researchers in deciding who to recruit. Furthermore, quota sampling using Facebook advertisements may also be particularly useful for generating targeted samples of geographic or demographic subpopulations, for which national panels may prove to be inadequate or prohibitively expensive. For instance, Sances 2017 has used Facebook advertisements without quota sampling to recruit voters from U.S. municipalities. For future research, we plan to investigate ways to further reduce the cost of this quota sampling method and validate Facebook's targeting demographics against respondents' self-reported demographics.

References

Berinsky, Adam J, Gregory A Huber, and Gabriel S Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk." *Political Analysis* 20 (3): 351–368.

Bond, Robert, and Solomon Messing. 2015. "Quantifying social media's political space: Estimating ideology from publicly revealed preferences on Facebook." *American Political Science Review* 109 (01): 62–78.

Broockman, David E, and Donald P Green. 2014. "Do online advertisements increase political candidates' name recognition or favorability? Evidence from randomized field experiments." *Political Behavior* 36 (2): 263–289.

Hirano, Shigeo, Gabriel S Lenz, Maksim Pinkovskiy, and James M Snyder. 2015. "Voter Learning in State Primary Elections." *American Journal of Political Science* 59 (1): 91–108.

Huff, Connor, and Dustin Tingley. 2015. ""Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents." *Research & Politics* 2 (3): 2053168015604648.

[13]It should be noted that Facebook has made it possible for advertisers to target users by users' political ideology. Users' ideologies are imputed from their social media behavior (Bond and Messing 2015). Therefore, it is feasible for researchers to construct strata that include political ideology as a demographic variable.

Jäger, Kai. 2017. "The potential of online sampling for studying political activists around the world and across time." *Political Analysis:* 1–15.

Kapp, Julie M, Colleen Peters, and Debra Parker Oliver. 2013. "Research recruitment using Facebook advertising: big potential, big challenges." *Journal of Cancer Education* 28 (1): 134–137.

Kosinski, Michal, Sandra C Matz, Samuel D Gosling, Vesselin Popov, and David Stillwell. 2015. "Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines." *American Psychologist* 70 (6): 543.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private traits and attributes are predictable from digital records of human behavior." *Proceedings of the National Academy of Sciences* 110 (15): 5802–5805.

Ramo, Danielle E, and Judith J Prochaska. 2012. "Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use." *Journal of medical Internet research* 14 (1): e28.

Ramo, Danielle E, Theresa MS Rodriguez, Kathryn Chavez, Markus J Sommer, and Judith J Prochaska. 2014. "Facebook recruitment of young adult smokers for a cessation trial: methods, metrics, and lessons learned." *Internet Interventions* 1 (2): 58–64.

Ryan, Timothy J. 2012. "What makes us click? Demonstrating incentives for angry discourse with digital-age field experiments." *The Journal of Politics* 74 (4): 1138–1152.

Samuels, David J., and Cesar Zucco. 2013. *Using Facebook as a subject recruitment tool for survey-experimental research.* Working paper, available at SSRN 2101458.

———. 2014. "The power of partisanship in Brazil: Evidence from survey experiments." *American Journal of Political Science* 58 (1): 212–225.

Sances, Michael W. 2017. "Ideology and Vote Choice in U.S. Mayoral Elections: Evidence from Facebook Surveys." *Political Behavior* In Press.

Santoso, Lie Philip, Robert Stein, and Randy Stevenson. 2016. "Survey Experiments with Google Consumer Surveys: Promise and Pitfalls for Academic Research in Social Science." *Political Analysis* 24 (3): 356–373.

Stewart, Neil, Christoph Ungemach, Adam JL Harris, Daniel M Bartels, Ben R Newell, Gabriele Paolacci, and Jesse Chandler. 2015. "The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers." *Judgment and Decision Making* 10 (5): 479.

Youyou, Wu, Michal Kosinski, and David Stillwell. 2015. “Computer-based personality judgments are more accurate than those made by humans.” *Proceedings of the National Academy of Sciences* 112 (4): 1036–1040.