MENA Newsletter: Pandemic-Safe Research with Online Surveys

Robert Kubinec

Division of Social Sciences

New York University Abu Dhabi

April 8th, 2021

In this article I argue that a possible silver lining of the COVID-19 pandemic, which has required wrenching changes of the MENA research community (Allam, Buttorff, and Shalaby 2020), is that it has increased awareness of the utility of online survey research. Scholarly research in the MENA region, and developing countries more generally, would benefit over the long term from an increased openness to online-only research designs that reach MENA residents where they increasingly interact with one another: on social media. Separate from the pandemic, online-only research has the added advantage of reducing potential harm to respondents and research assistants, whose online activity within social media sites can only be imperfectly observed by domestic authorities.[1]

Online Surveys via Social Media Ads

I draw, in this article, from my own and my colleagues' recent experience implementing surveys using Facebook ad targeting. Facebook is probably the most useful social media platform in the MENA region as its usage vastly exceeds the scope of other popular social media platforms.[2] The lack of censorship of Facebook by MENA countries, even in highly authoritarian regimes, makes it an attractive means to reach residents directly. Facebook usage in the region, reaching 74 percent of all internet users according to the Arab Barometer's most recent data, has yet to peak (Wee and Li 2019, 7). In addition, those who have social media accounts tend to spend a significant percent of their time on the medium, with an astonishing 94 percent spending 2 hours or more on a social media site (Wee and

---

[1] While activists in the MENA region are increasingly targeted by so-called "phishing" campaigns designed to give intelligence services direct access to their devices, without such access it can be difficult for telecom regulators or intelligence services to directly monitor Facebook traffic because it uses encrypted protocols (i.e., https). It is for this reason that China, for example, bans Facebook entirely rather than attempt to screen traffic within the site.

[2] While this is true for most MENA countries, there are exceptions. In Qatar, for example, Facebook usage remains low due of Facebook's requirement that users have profile pictures and real names. However, other Facebook-owned sites like Instagram maintain strong followings, including in Qatar, which can also be used for ad targeting.

Li 2019, 8). These adoption rates mean that social media is close to supplanting television as a primary source for Arabs' news (Wee and Li 2019, 10).

Given the high utilization rate of Facebook in the region, it is no surprise that most online survey research likewise employs Facebook, or one of its subsidiary sites like Instagram, in some way. The major commercial provider of online survey research, YouGov, relies on Facebook ads to recruit users into its panels, from which it draws samples for academic research.[3] Scholars can take one of three approaches in obtaining samples that can be ranked in terms of their involvement in subject recruitment: (1) sign up for a marketing account and target their own ads at Facebook users; (2) recruit a social media marketing firm to target ads on their behalf,[4] or (3) purchase a sample from YouGov.

Recent research has produced important findings from both direct targeting by scholars and via YouGov. In a recent paper with Tessa Thornton, we recruited 1,573 respondents in the West Bank in the spring and summer of 2020 to better understand the relationship between Palestinians and the growing presence of Israeli settlements by employing Facebook ads targeted at different age and gender demographics (Thornton and Kubinec 2020). In a separate project with Andrey Tomashevskiy and Haillie Lee that employed direct ad targeting, we show that Egyptian companies without political connections were much more likely to shut down due to COVID-19 restrictions relative to companies without these connections (Kubinec, Lee, and Tomashevskiy 2020). Williamson et al. (2020) analyzed a massive sample purchased from YouGov to understand in much more detail how support for Islamist political candidates varies across MENA countries. Guiler (2020) employed a sample recruited directly via Facebook ads to probe how Turkish citizens evaluated electoral candidates with and without prison experience. Finally, in ongoing research, Sharan Grewal, Tahir Kilavuz and I use Facebook targeting to closely monitor Algeria's Hirak protest movement, revealing the nature of protester demands and the drivers of movement longevity (Grewal, Kilavuz, and Kubinec 2019).

Addressing Skepticism toward Online Surveys

Despite increasing usage, skepticism remains among the scholarly community about the utility of online survey research, and Facebook ad targeting in particular, as a means of learning about political and social attitudes. It is true that neither Facebook usage nor internet connectivity is randomly distributed within and across MENA countries, with poorer countries like Yemen having less than 50% of the population connected and rural areas showing less connectivity even within countries with high penetration overall (Hoogeven, Rodriguez, and Aziz 2020). In an interesting sociological trend for the region,

[3] Based on a conversation with a YouGov representative at the Middle East Studies Association Conference in the fall of 2019.

[4] This method's primary advantage is both to minimize the time needed to learn Facebook's arcane ad targeting system, and can also help to avoid having the researcher's ads falling under Facebook's misinformation policies, which scrutinize ads that mention social or political issues and do not have any mechanism to differentiate academic research.

men tend to be more likely to use Facebook than women (Fatehkia, Kashyap, and Weber 2018), which could be in part due to cultural norms against posting profile pictures.

These known biases may convince scholars that they should not trust data collection to a site over which they have relatively little control. However, the question is not whether Facebook is a flawless method of subject recruitment, but rather *how it compares to other data-gathering methods in the region*. Comparing online surveys to the mythical simple random sample in which the researcher draws balls from Polya's urn, as many learn in graduate school, does not adequately capture a situation in which the balls are sentient human beings. For example, the Arab Barometer, which has the widest scope and coverage of any survey research in the region, includes weights in all its surveys because even its direct household interviewing methods do not return samples that match population totals, requiring the use of weights to account for demographic imbalance.

Adjusting for Bias in Online Surveys Using MRP

Nonetheless, it is not easy to know how to best adjust an online survey to account for known biases. Online survey research has become increasingly popular in opinion polling research because it is ideally suited for a relatively novel adjustment technique known as multilevel regression with post-stratification (MRP) (Park, Gelman, and Bafumi 2004). The method became more prominent after Wang et al. (2014) successfully employed the technique to correct an online survey of Xbox users to more accurately predict the result of the 2012 U.S. election than conventional polling methods.

To employ MRP, the analyst needs access to a contingency table (usually derived from census data, although large surveys are a secondary option) that gives the population distribution of relevant demographic criteria, such as the distribution of citizens across urban/rural areas and by gender. To correct a survey for population imbalance, it is necessary to fit a model regressing the survey outcome on binary indicators for these same criteria within the survey, e.g., the gender and urban or rural location of the survey respondent. The model can then predict the average survey response by each combination of these criteria, and the model predictions can be re-weighted using the census-derived contingency table to match what the population looks like. Thornton and I employ this method to derive an estimate for the average distance that a Palestinian lives from an Israeli settlement–14 kilometers–despite the fact that our sample is heavily male-skewed (80% of the sample), by separately calculating the distance for all cells of a contingency table of the West Bank population by gender, district and age and summing to produce a weighted average.

Employing MRP effectively requires rethinking the survey design. First, it is important to know ahead of time what kind of census or large survey information is available to construct a contingency table, and questions in the survey must match the coding used in the census. Second, the survey collection design should focus on maximizing the size of the sample rather than balance per se. For example, in our Palestine survey, we targeted more ads at women than men to help address known imbalances, but we did not exclude male respondents as MRP works best with larger samples. An older male respondent can help correct for age imbalance, for instance, even if not gender imbalance. The important

sampling criterion is the total degree of diversity in responses rather than maintaining even quotas.

Finally, it is important to note that the population re-weighting procedure is only necessary if the aim is population inference, i.e., extrapolating the average survey response to the population. In many cases, scholars want to fit models that focus on the effect of covariates rather than estimating population quantities. In this case, simply including these demographic indicators as control variables is sufficient to account for how these design issues might affect a coefficient of interest in a regression model (Gelman 2007). By maximizing sample size–which is often much cheaper via Facebook ad targeting–and explicitly incorporating indicators that affect population biases, it can be straightforward to collect data that provide representative answers on important questions in MENA countries.

The Future of Online Surveys in the MENA Region

It does not seem that, following the end of the pandemic, online activity will decrease in the region. Rather, the increased reliance on online media for work, enjoyment and communication is likely to make social media an even more important domain for researchers to be active in. For example, the United Arab Emirates recently launched new visa programs to attract remote workers and help boost hotel and tourism revenues.[5]

In one hopeful example, the Arab Barometer's success in implementing its most recent round of surveys with online methods should help us understand how online methods compare to traditional household surveys. While face-to-face research can never be replaced, becoming more proficient at online research can make our research more inclusive of online communities and less vulnerable to regional instability and unforeseen events like pandemics.

---

[5] *The National*, "UAE to Issue Residency Visas for Remote Work," March 21, 2021. https://www.thenationalnews.com/uae/government/uae-to-issue-residency-visas-for-remote-work-1.1188299

References

Allam, Nermin, Gail Buttorff, and Marwa Shalaby. 2020. "Report: COVID-19 Pandemic Compounds Challenges Facing MENA Research." *MENA Politics Newsletter* 3 (2): 5–9.

Fatehkia, Masoomali, Ridhi Kashyap, and Ingmar Weber. 2018. "Using Facebook Ad Data to Track the Global Digital Gender Gap." *World Development* 107 (July): 189–209. https://doi.org/10.1016/j.worlddev.2018.03.007.

Gelman, Andrew. 2007. "Struggles with Survey Weighting and Regression Modeling." *Statistical Science* 22 (2): 1.

Grewal, Sharan, Tahir Kilavuz, and Robert Kubinec. *The Brookings Institution.* 2019. "Algeria's Uprising: A Survey of Protesters and the Military." https://www.brookings.edu/research/algerias-uprising-a-survey-of-protesters-and-the-military/.

Guiler, Kimberly G. 2020. "From Prison to Parliament: Victimhood, Identity, and Electoral Support." *Mediterranean Politics*, 1–30. https://doi.org/10.1080/13629395.2020.1721159.

Hoogeven, Johannes, Laura Rodriguez, and Atamanov Aziz. 2020. "Using Internet Surveys in the MENA Region During COVID-19: Will All Voices Be Heard?" https://blogs.worldbank.org/arabvoices/using-internet-surveys-mena-region-during-covid-19-will-all-voices-be-heard.

Kubinec, Robert, Haillie Na-Kyung Lee, and Andrey Tomashevskiy. 2020. "How to Get Away with Spreading COVID-19: Political Connections and Pandemic Response." *SocArXiv*, December. https://doi.org/10.31235/osf.io/68fpr.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls." *Political Analysis*, 375–85.

Thornton, Tessa, and Robert Kubinec. 2020. "Palestinian Attitudes Towards Settlers in the West Bank." *SocArXiv*, November. https://doi.org/10.31235/osf.io/nv35r.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2014. "Forecasting Elections with Non-Representative Polls." *International Journal of Forecasting*.

Wee, Jason, and Sophie Li. 2019. "Politics and Social Media in the Middle East and North Africa: Trends and Trust in Online Information," 15. https://www.arabbarometer.org/wp-content/uploads/AB_Media_Report_Final_Public-Opinion-2019-5.pdf.

Williamson, Scott, Kadir Yildirim, Sharan Grewal, and Mirjam Kuenkler. 2020. "Preaching Politics: How Politicization Undermines Religious Authority in the Middle East," 38.

Robert Kubinec, New York University Abu Dhabi, bobkubinec@gmail.com