

Comparative Politics Newsletter

The Organized Section in Comparative Politics of the American Political Science Association

Editors: Matt Golder (mgolder@psu.edu), Sona N. Golder (sgolder@psu.edu)

Editorial Assistant: Charles Crabtree

Editorial Board: Molly Ariotti, Xun Cao, Elizabeth C. Carlson, Amanda Fidalgo, Vineeta Yadav, Joseph Wright

Contents

Letter from the Editors: Sensitive Topics by Matt Golder and Sona N. Golder	1
I. Symposium: Studying Sensitive Political Phenomena	
Can Anonymity Promises Possibly be Credible in Police States? by Jesse Driscoll	4
Sensitive Data and the French State by David Laitin	7
Survey Methods for Sensitive Topics by Graeme Blair	12
Randomized Response: Foundations and New Developments by Daniel Gingerich	16
Capturing Ethnic Salience on Surveys in Developing Countries by Elizabeth C. Carlson	27
Lobbying, Corruption, and Non-Responses in Small Samples by Vineeta Yadav	32
II. Special Topic	
The Ethics of Field Experiments in Comparative Politics by Scott Desposato	37
III. Dataset	
Integrating Data on Ethnicity, Geography and Conflict: The Family of Ethnic Power Relations Datasets 2014 by Nils-Christian Bormann, Manuel Vogt, and Lars-Erik Cederman	41
IV. Other	
Contributors	46
Announcements and Additional Information	50

Letter from the Editors: Sensitive Topics

by Matt Golder & Sona N. Golder

The Pennsylvania State University

Welcome to the Spring 2015 issue of the Comparative Politics Newsletter. We are very happy to take over as editors, and would like to take this opportunity to thank the previous editors at the Hertie Institute in Berlin, Mark Hallerberg and Mark Kayser, for their excellent work over the last three years. We hope to maintain their high standards during our own editorial term.

One of the most important goals of the Comparative Politics Newsletter in our opinion is to engender a sense of community among comparative poli-

tics scholars around the world. To help with this, we have created a webpage for the Newsletter at <http://comparativenewsletter.com>. Please take a look at the webpage. You'll find information about our editorial team at Penn State, but, more importantly, you'll also find information about the contributors to our upcoming issues. As you'll see, our next issue will include a symposium on training graduate students in comparative politics. In addition, you'll find links to previous issues of the Newsletter. To make the website more useful, we also have a section on the Homepage that provides news and announcements that might be relevant to the members of the Comparative Politics Section. If you have anything that you would like to share with others, please contact us and we will be happy to post it on the website. You can contact us through the [Contact](#)

page or, alternatively, by using our Penn State email addresses (mgolder@psu.edu, sgolder@psu.edu). So that you do not have to keep checking the website to see if there have been any news updates, you can also follow us on Twitter [@cp_newsletter](https://twitter.com/cp_newsletter). We hope that by using the webpage and our Twitter account, we can pass along important news and announcements in a much more timely manner than has been possible in the past. If you have any suggestions for improving the utility of the webpage or Newsletter, we'd be very happy to hear them.

In line with previous editorial teams, the Newsletter will include symposia on various substantive and methodological issues, highlight new datasets of broad appeal, and generally inform the community about field-specific developments. With the exception of our inaugural issue, all future issues will also include a new section, tentatively called, *Letters to the Editors*. This section will include selected short comments (less than 300 words) from our readers in response to the material in the previous issue of the Newsletter. These comments might critique or build on the contributions from the previous issue, highlight some line of research or data source that was overlooked, or more generally provide additional information that might be of use to the Section membership. Our hope is that this new section can help engender and sustain more of a dialogue amongst comparative political scientists. We encourage you to submit your thoughts about the content of our current Newsletter using the [Contact Page](#) on the Newsletter website.

Our current issue includes a symposium looking at how researchers go about studying sensitive political phenomena, a special topic on the ethics of experiments in comparative politics, and an overview of the family of Ethnic Power Relations (EPR) datasets.

Our opening contribution to the symposium on researching sensitive political phenomena comes from [Jesse Driscoll](#), an Assistant Professor in the School of International Relations and Pacific Studies at the University of California, San Diego. Jesse spent over two years conducting ethnographic research in Tajikistan and Georgia for his forthcoming Cambridge University Press book, *Warlords and Coalition Politics in Post-Soviet States*. Many of the people Jesse interviewed were former combatants and militia members, some of whom had engaged in violent criminal activity both on behalf of, and against, the state. In his contribution, Jesse de-

scribes some of the safety, methodological, and ethical issues that he faced when conducting his research, particularly the tradeoff between promises of anonymity and emergent norms of data replicability.

Our second contribution comes from [David Laitin](#), a Professor in the Department of Political Science at Stanford University. David's essay looks at the relationship between sensitive data and the French state, a topic covered in his forthcoming Harvard University Press book, *Why Muslim Integration Fails in Christian-Heritage Societies* (co-authored with Claire L. Adida and Marie-Anne Valfort). Historically, the French state has prohibited the collection or processing of data that reveal the racial or ethnic origins (or political, philosophical or religious opinions, or union membership, or health, or sexual life) of people living in France. This prohibition is currently a matter of hot political debate in France due to social tensions in the immigrant *banlieues* and growing evidence of ethnic and religious discrimination in the public sphere. The French Prime Minister Manuel Valls recently went so far as to say that "territorial, social and ethnic apartheid" existed in France ([de La Baume, 2015](#)). The January 2015 terrorist attacks carried out by French-born radical Islamists on the offices of the satirical magazine *Charlie Hebdo*, which occurred after David had submitted his piece, have added further impetus to those calling on the French state to authorize the collection of ethnicity data. It's not clear how much will change, though. On February 5, the French president, François Hollande, stated that collecting ethnic data would be "*inutile*" (useless) and would not help solve social problems related to unemployment and educational underachievement ([Baumard, 2015](#)).

Our third contribution comes from [Graeme Blair](#), a Ph.D. Candidate at Princeton who will be an Assistant Professor at UCLA from 2016. When scholars use survey data, they implicitly assume that respondents have answered the survey questions truthfully. However, this assumption may be violated if respondents have incentives to conceal the truth, something that is likely to happen when the survey addresses sensitive topics. Graeme's essay provides an overview of various methods — survey administration protections, randomized response techniques, list experiments, and endorsement experiments — that can be used to elicit more truthful responses. He also discusses some of the critiques that can be made of each of these techniques.

[Daniel Gingerich](#), an Associate Professor in the Department of Politics at the University of Virginia, provides our fourth contribution. Daniel offers an in-depth look at how randomized response techniques can be used to study sensitive behavior and attitudes. Among other things, he describes several of the most commonly used randomized response models, outlines key recent developments in the use of these models, and discusses the findings of validation studies that have examined the performance of these models for a variety of sensitive topics both in political science and more generally. He concludes with some practical advice for implementing randomized response surveys in the field.

Our last two contributions to the symposium come from two of our colleagues who work on sensitive topics, [Elizabeth C. Carlson](#) and [Vineeta Yadav](#). Elizabeth is an Assistant Professor in the Department of Political Science and the Program on African Studies at Penn State. She uses survey and experimental methods to study political behavior and citizen preferences in Africa's new democracies. Some of her research attempts to identify the conditions under which Ugandan voters prefer a co-ethnic candidate for office, especially when that candidate is dubiously qualified. In her contribution, she reports the results of an experiment in which she examines reported levels of co-ethnic voting when (i) the ethnicity of the enumerator and respondent are matched, and (ii) when questions about co-ethnic voting are asked through a secret ballot embedded within a larger face-to-face interview. She finds that matching the ethnicity of respondents and enumerators exacerbates bias in reported co-ethnic voting, but that providing an opportunity for respondents to self-administer sensitive questions can lead to a more accurate reporting of co-ethnic voting.

Vineeta Yadav is an Assistant Professor in the Department of Political Science at Penn State. Her research focuses on the effects of institutions on economic development, with a particular emphasis on how institutions influence lobbying, corruption, and judicial empowerment. In her contribution, Vineeta points out that many of the survey methods developed for dealing with sensitive questions — randomized response techniques, list experiments, and endorsement experiments — require a large sample size for successful implementation. This can be problematic if one's population of interest is relatively small. Vineeta goes on to discuss this issue in the context of a business elite survey that she con-

ducted on corruption and lobbying in Brazil and India. She suggests that it is important to take into account the selection process that causes only some respondents to provide answers to sensitive survey questions.

Our inaugural issue of the Comparative Politics Newsletter also includes a special topics contribution by [Scott Desposato](#) on ethics and experiments in comparative politics. Scott is an Associate Professor in the Department of Political Science at UC San Diego. In 2013, he organized and hosted the NSF-funded [Ethics in Comparative Politics Experiments Conference](#). The goal of the conference was to identify critical issues with experiments in comparative politics, explore opinion in the field, and propose practical strategies for moving forward. In his contribution, Scott summarizes the central topics that came up during this conference, focusing on the importance of taking context into account, the role of local review, issues to do with deception and consent, and the impact that experiments can have in real political situations. A more detailed account of the proceedings of the conference will be published later this year in the [Routledge Series in Experimental Political Science](#).

Finally, [Nils-Christian Bormann](#), [Manuel Vogt](#), and [Lars-Erik Cederman](#) describe the family of Ethnic Power Relations (EPR) datasets. Nils-Christian and Manuel are both post-doctoral researchers in the International Conflict Research group at ETH Zurich; Lars-Erik is Professor of International Conflict Research in the Center for Comparative and International Studies at ETH Zurich. The EPR datasets integrate information on ethnicity, geography, and conflict for over 800 unique ethnic groups in 165 countries from 1946 to 2013. The individual components of the EPR dataset family can be downloaded from the GROW^{up} web portal at <http://growup.ethz.ch>.

We hope that you enjoy our first issue of the Comparative Politics Newsletter. If you have ideas for possible symposia or special topics, or would like to publicize a dataset of broad appeal, please contact us. Again, you can contact us through the Contact page of our webpage at <http://comparativenewsletter.com/contact> or simply use our Penn State email addresses (mgolder@psu.edu, sgolder@psu.edu). We look forward to working with you over the next four years.

Matt and Sona

References

- Baumard, Maryline. 2015. "François Hollande estimate inutile les statistiques." *Le Monde*.
- de La Baume, Maïa. 2015. "French Premier Says 'Apartheid' Is Leaving Minorities on the Fringe." *The New York Times*.

I. Symposium: Studying Sensitive Political Phenomena

Can Anonymity Promises Possibly be Credible in Police States? Interview Strategies in Post-War Authoritarian Regimes

by Jesse Driscoll

University of California, San Diego

"I get it. I know what you want. I understand. You want to make up a list, on your computer, of all our bad men. 'Terrorists.' You want to cross names off the list when they were killed or jailed. To see that we Tajiks can take care of our own. But we can. We did. You'll see."

Yuri, Dushanbe, 2007

In the course of researching certain topics — particularly topics related to civil war violence, criminality, and terrorism — the best data tends to come from people who are implicated in violence and criminality. Researchers are required to construct credible protections for terrible people to speak candidly about terrible things that they have done. I put a great deal of energy into worrying about the practicalities of how to protect the identities of these people while conducting research for my forthcoming book, *Warlords and Coalition Politics in Post-Soviet States*. The book attempts to explain how stable authoritarian regimes emerged after state failure in Central Asia and the Caucasus. This short essay, which draws heavily on material from my book, is meant to illuminate the practical trade-offs between anonymity promises — which are necessary for the researcher to get decent data in the first place, and to credibly present herself as an academic and not as a spy — and emergent norms of data replicability.

I traveled to Central Asia for the first time in 2005 and proceeded to spend nearly 25 uninterrupted months in Tajikistan and Georgia between 2006 and 2008. I employed ethnographic methods — including living in a remote village on the Tajik-Afghan border — to gain rapport with interview subjects. I gradually forged relationships with academics, journalists, expatriate professionals, current and former government employees, and eventually many former militia members in both countries. I conducted over 300 interviews, 173 of which were with former combatants.

At first, I naively tried to avoid implicating my respondents in descriptions of abjectly criminal activities. It quickly became obvious that this was impossible. I settled for a policy of total anonymity for all respondents. I recorded only the subject's first name, or a pseudonym, and interview date to preserve anonymity. In a few cases, the subject insisted that I record his full name. For my own safety, and that of my respondents, I never complied with these requests.

Many of my respondents used to be employed in the police or the army. At the time of the interview, many were marginally employed. Many of these men had paid severe costs in order to "capture the state" and had succeeded in transforming themselves from paramilitaries into uniformed members of state security services; later they had themselves been purged. Many held bleak and ironic attitudes, and could angrily recall the jarring transition from being "the law" to being "on the wrong side of the law" in a few arbitrary weeks. In early interviews, a strategy to collect individual-level characteristics to predict which militia members joined the state, which factions ended up with which jobs, and which factions were "weeded out," I had respondents sketch pictures of the security structures at different periods of time. Many early respondents gamely played along. Though the exercises quickly devolved into incoherence — pages filled with circles, lines, and scribbled names that I realized would never be coded or systematized — the process was intoxicating. Meandering conversation threads gradually gave way to more structured discussions about who was getting what, how and when side-switching between militia commanders was considered, and the like. For a time, I was sure I was getting the real story. For a time, this arcana made me feel cool, like I was a character on *The Wire*.

But it did not stay cool for very long. I was living in Kyrgyzstan in 2005 when the Andizjan events unfolded across the border in Uzbekistan. There was an attempted prison break that was put down with great violence by Uzbekistan's government. Many people died. For a few very uncomfortable weeks, it was not clear whether the refugees who had fled to Kyrgyzstan would spread a contagion of ethnic fear across the volatile Ferghana Valley or if the war in Afghanistan was creeping north.¹ I was living in Kyrgyzstan's capital city of Bishkek at the time, many hundreds of kilometers away from the violence. But I recall distinctly the feeling of vulnerability when the internet stopped working correctly. I decided that if I was going to continue operating in an unfriendly authoritarian environment, I needed to adapt. I quickly educated myself about how internet servers work. I stopped assuming my email communications were private. For important topics I began to rely upon pen and paper (see Figure 1). But these realizations gradually transformed my relationship with "the field." Especially in Tajikistan, it meant proceeding very slowly over months that turned into years. After a few frightening encounters, I was warned by a trusted advisor that my research design was making me look quite a bit like a spy, and that if I ran afoul of the wrong character in the wrong security bureaucracy, it could easily result in my permanent disappearance. The risks to my local contact network were, if anything, even more serious. I was asking the questions that locals had learned not to ask.²

Certain subject matter was carefully filtered by the respondents themselves. It was very rare for anyone to discuss money in specific denominations. I learned a lot about the banalities of racketeering, price-fixing, money laundering, and other mainstays of the post-Soviet underworld, but specific information about the dispensation of family assets was carefully guarded. Details were not volunteered and I never asked. Resentment towards

Figure 1: Why I Did Not Write Any Names Down On My Fieldnotes, Or Take Notes On My Computer

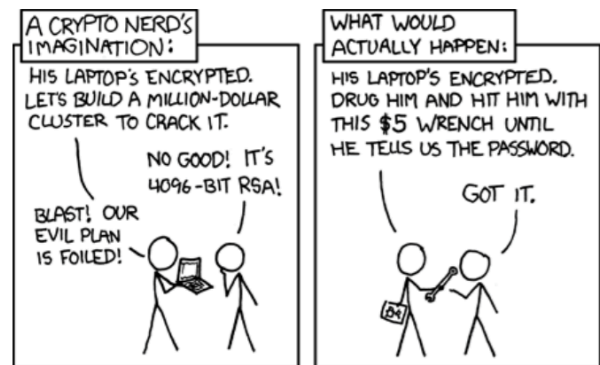


Image source: <http://xkcd.com/538/>.

the current generation of "big men" was tempered by the realities of the embedded favor economy. There are strong norms against taking the side of outsiders against the nation, tribe, or family. I never stopped being an outsider. My field interviews improved substantially when I accepted the limits of my role. I acquired a reputation for being primarily interested in political stories and having no particular local agenda. I also took conscious steps to control the subject matter, intentionally shying away from asking respondents about their own violent activities. The main way I did this was to make it clear that I was studying institutions — militia structures — and not people. I overcame my skepticism of questions like "in your group did you observe ..." when I realized that respondents could describe their own actions in the third person, with plausible deniability and emotional distance from actions that were regretted in retrospect.³ But interviews often sprawled, and what I ultimately ended up collecting were life histories. With time, as I learned to ask the right questions, and behind the veil of anonymity that my methods provided, I received rich anecdotes.

¹ Akiner (2005) went to Andizjan two weeks after the uprising to conduct interviews and concludes that the demonstrations were a "carefully prepared" attack on the Uzbek government organized by armed militants (part of the "Akromiya movement") who were multi-national in composition (30-31, 27-29). Her version of events is contradicted by the OSCE (2005), Ilkhamov (2006), Kendzior (2006), Bukharbaeva (2005), and Daly (2005). Although relations between the West and the government of Uzbekistan deteriorated significantly as a result of the Andizjan events, I am certain that most of the readers of this essay have never even heard of Andizjan; central Asia is very remote from American national interests.

² A Tajik doctoral student and personal friend was recently placed under house arrest in Tajikistan pending a trial for charges of treason and espionage associated with qualitative observational research (conducting interviews in Badakhshon) on behalf of a non-Tajik Principal Investigator. He was allowed to leave the country and currently resides in Canada, but he may never be allowed to return. Journalists and writers are routinely intimidated in Central Asia. Independent of each other, both of my regular translators in Tajikistan (both aspiring journalists) requested that I omit their names from my book and all future publications.

³ I acknowledge that there is some risk that the leading structure of the interview questions biased respondents towards remembering the worst in their counterparts, which may mean that the book's argument over-emphasizes the role of criminal sociopathy in small unit organizations. On the other hand, I am absolutely confident that the war brought a lot of sociopaths out of the woodwork in both Georgia and Tajikistan.

Not everyone trusted me at first. Some respondents left the interview not trusting me at all (and who could blame them?!). It is reasonable to second-guess the motives of networked strangers who arrive from far away states to ask questions about a war. Having extended conversations about post-Soviet security structures occasionally led respondents to ask pointed questions about my actual goals, leading to uncomfortable and conspiracy-laden conversations about whether old Russian phone taps had been replaced with American post-9/11 security assistance. I took notes but did not record interviews (Derluguian, 2005), and I am sure that if I had started taping voices I would have received less access and very different kinds of data. Though there is technically an amnesty law on the books, many respondents became visibly uncomfortable when the conversation turned to certain topics, especially their opinion of Tajikistan's long-sitting president Emomalii Rakhmon. But strange as it sounds, many of the respondents who assumed I was a spy, or affiliated with the U.S. military, were often *more willing* to engage with me. I can only speculate why this was the case. Perhaps they saw me as being part of the same fraternity of warrior-defenders, or they believed that we shared the experience of making bad youthful decisions and getting in over our heads.

After a few frightening encounters, I was warned by a trusted advisor that my research design was making me look quite a bit like a spy, and that if I ran afoul of the wrong character in the wrong security bureaucracy, it could easily result in my permanent disappearance.

It is also the case that some of my respondents were sociopathic liars, and that some respondents misrepresented basic facts about events, their beliefs, and their roles. Still, I listened hard, asked follow-up questions when I could, and usually stayed until the person wished to stop talking. But sometimes I got scared. The interview subjects always noticed. The interview ended shortly afterwards. That I conducted so many interviews is evidence that, from my perspective, the participants in these wars were not all thugs, gangsters, and sociopaths.

I could not have created the necessary intimacy if they were. Obviously there were strategic reasons to present distorted and self-serving versions of certain facts, but most respondents seem to remember doing what they thought was appropriate at the time — and some of it was horrible. After a few uncomfortable conversations, I tried to enforce a policy of stopping interviews when subjects divulged personal memories of violence or war crimes. But this was not always possible. Many interviews had a confessional aspect to them. Some men tried to shock me with graphic descriptions of tactics — inflicted or received — in order to test, embarrass, or entice me. As an author, I have learned that there is no catharsis to be gained by seeing these descriptions in print. Lurid descriptions of mutilations, written from the safety of the ivory tower, cheapen the horror of the experiences for participants. The deepest scars — post-traumatic stress and survivor guilt — are invisible. Ethnographers working in conflict zones eventually learn what police detectives and competent military professionals have understood for time immemorial: there are serious mental and emotional costs associated with treating sociopaths as subjects.⁴

Authoritarianism is more a set of practices than a kind of state, and sometimes fears that originate in authoritarian environments follow researchers home to their “open access” societies. I spent a lot of time worrying about what I would do if my field notes were subpoenaed by the FBI to assist the Russian or Tajik government in dismembering the networks that I accessed for my research. It is clear that even liberal states can, and if they deem it necessary will, deploy the law to place academics in horrible double-binds.⁵ After returning to the United States, and participating in conversations with a few West Point professors who teach classes that specialized in network analysis, I realized that to limit my own liability it was necessary to physically destroy all of my field notes that could be used to reconstruct contemporary networks (including all those pages of circles, sketches, and names described above). Some unlucky PI will likely be forced to choose between “protecting his sources” (e.g., releasing data that was acquired under explicit promises of confidentiality) and spending time behind bars, and I did not want it to be me. I suspect that in the future every researcher — qualitative

⁴I occasionally wonder how my life would be different if I had really read and absorbed E. Valentine Daniel (1996) before going into the field.

⁵Human subjects protections promised by Boston College were more or less moot once Northern Ireland police decided that they needed access to the names in the researcher's archive to prosecute the killing of Jean McConville. For more details, see <https://bostoncollegesubpoena.wordpress.com/>.

or quantitative — who collects data related to contemporary terrorism, criminality, the drug war, or civil war violence in general is going to have to self-define their role with a great deal more self-awareness.

People do not, as a rule, like to be studied. And as Jarvis Cocker observed, “everybody hates a tourist.” People do, as a rule, like the idea that their stories will be recorded for posterity. Ethnographers — by their invasive presence — force these decision heuristics into conflict. As the line between researcher and subject blurs, what occasionally emerges is a crucible for creative, cooperative theory-building. People can tell when their words are being received with empathy (i.e., when they are being treated as *subjects*) and when their words are being clinically recorded for some other purpose (i.e., when they are being treated as *objects*). I found that once a subject decides that the researcher is actually listening — and taking anonymity promises seriously — the researcher gets much better data.

References

- Akiner, Shirin. 2005. *Violence in Andijan, 13 May 2005: An independent assessment*. Technical report.
- Bukharbaeva, Galima. 2005. *Testimony: Statement of Galima Bukharbaeva, Correspondent—Institute for War and Peace Reporting*. Technical report Helsinki Commission Hearing.
- Daly, John C. K. 2005. “The Andijan disturbances and their implications.” *Central Asia-Caucasus Analyst* 29.
- Daniel, E. Valentine. 1996. *Charred Lullabies: Chapters in an Anthropography of Violence*. Princeton, NJ: Princeton University Press.
- Derluguian, Georgi M. 2005. *Bourdieu’s Secret Admirer in the Caucasus: A World-System Biography*. Chicago, IL: University of Chicago Press.
- Ilkhamov, Alisher. 2006. “The phenomenology of “Akromiya”: Separating facts from fiction.” *China and Eurasia Forum Quarterly* 4(2006):39–48.
- Kendzior, Sarah. 2006. “Inventing Akromiya: The role of Uzbek propagandists in the Andijan massacre.” *Demokratizatsiya: The Journal of Post-Soviet Democratization* 14(4):545–562.
- OSCE. 2005. *Preliminary findings on the events in Andijan, Uzbekistan, 13 May 2005*. Technical report OSCE Office for Democratic Institutions and Human Rights.

Sensitive Data and the French State

by David Laitin
Stanford University

Much attention has been directed at the obligations of scholars in protecting individuals from assaults on their privacy. Indeed, the entire IRB apparatus has been constructed on that concern. Less attention has focused on states and their obligations to protect their citizens from the data collection they sponsor and the use of those data for social analysis. In France, the tables are somewhat turned. There has been little attention to IRB approval of individual research projects, but the rules for the scientific use of state-collected data are a matter of high public concern. This is especially the case concerning data about the ethnic composition of France’s population.

This essay will first review France’s approach toward official ethnic data. It will then discuss the political pressures for change. Finally, it will describe the outlines of the new regime. This emerging regime seeks to balance an ideology of republicanism and a concern for the privacy of its citizens. In so doing, it recognizes the need to understand ethnic processes so that public officials can better address fundamental social, political, and economic problems.

*I. The French Republican Tradition*¹

French republican ideology emphasizes that the state has no interest in knowing the ethnic past of any of its citizens, as from the state’s point of view, all citizens are equally French. The nation, according to the French sociologist Dominique Schnapper is a “community of citizens” (Schnapper, 1998, 16–17). It is not, she has insisted, an ethnic group.² In her republican framework, ethnic attachments are irrelevant for national membership. Furthermore, state recognition of ethnic groups undermines national solidarity. As a member of the Constitutional Council (the highest constitutional authority in France) that had authority over issues of national accounting, her views had considerable sway.

Schnapper’s views on citizenship, ethnicity, and the nation, which she partially revised after retiring from the

¹This section draws from the Appendix of Adida, Laitin and Valfort (2015). Thanks to Patrick Simon for his comments on an earlier draft of this section.

²She moderates her position in Schnapper (2007).

Constitutional Council and returning to research, are deeply shared across the political spectrum in France. They are associated with the radical republican ideals that reached ascendance in France in the early twentieth century when the principal social divide was transformed from class conflict to religious conflict. The early 20th century divide between the seculars and the church reflected battles going back to the Reformation, as France moved back-and-forth between full support of the Catholic Church and then issuing edicts proclaiming toleration for all cults.³ The republican focus on a secularism, called *laïcité*, entailed a reining in of the gains made by the Church in getting full control over public education during the period of the Second Empire (1852-70). As the Second Empire collapsed amid the Paris Commune of 1871, and in a fit of republican frenzy signaling their hatred of the Church, radicals executed the Archbishop of Paris.

The Republicans won the first elections of the 3rd Republic in 1879 and began a “*grande laïcisation de l’État*”. Laws were passed repealing the Sunday closings of businesses, ending Church control over hospitals, and secularizing cemeteries. A law formally separating church and state was passed in 1905 granting freedom of conscience to all, with restrictions that could only be justified by need for public order. While only four “cults” (Catholic, Lutheran, Calvinist and Jewish) received official recognition, all others were free to practice. For the next century, Republicans made further assaults on Church subventions, though by the 5th republic (while the religious cleavage still revealed typical voting patterns) there were no more outstanding substantive issues on the republican agenda. But suddenly the issue of *laïcité* took on new meaning in 1989 with the “*foulard*,” or “headscarf,” incident in which confessionals and lays united as “seculars” against the public display of Muslim identity (Bowen, 2006).

One implication of the republican consensus is that the collection of ethnic data by the state has long been an anathema. This has severely restricted careful social analyses of immigration patterns. For example, in an important study seeking to measure intergenerational inequalities linked to immigration led by the state-

affiliated research group INSEE (*Institut National de la Statistique et des Études Économiques*), the researchers could not distinguish between descendants of French nationals born abroad (the “*pièdes noirs*”) and descendants of immigrants from the same country. The authors confessed that “the fact that they cannot be distinguished in our data leads to an underestimation of the negative effects of inadequate social capital among immigrants” (Meurs, Pailhé and Simon, 2006, 676).

The issue of ethnic data and its collection has continued to be debated among state statisticians and social scientists. The flavor of this debate has been nicely captured in a special issue of *French Politics, Culture and Society*. Its history is convoluted.⁴ In the grand republican era, official statistics recognized only those who were “French”, those who were “French by acquisition” (though with no legal status differentiating them from the first category), and those who were “foreigners.” However, in the Vichy era (for Jews) and today (for Arabs), republican principles were compromised as the state differentiated them from Catholic-heritage French.

Researchers in INED, as well as INSEE, were split between those who felt that republican ideals could best be met by addressing discrimination through the analysis of ethnic data and those who felt republican ideals would best be met with a state that did not recognize its citizens through an ethnic lens.

Despite these deviations, considered a stain on republican principles, French law continues to be resolutely republican. A 1978 law prohibited the collection or processing of data that reveal the racial or ethnic origins (or political, philosophical or religious opinions, or union membership, or health, or sexual life) of persons. The French Data Protection Authority (CNIL) was then established to assure more assiduous implementation of the law. In 2005 HALDE (High Authority for Antidiscrimination and Equality) was created and funded by the state, but it was unable to infer general trends of discrimination since it could not collect systematic data on

³See Bertossi (2012) for an analysis of the changing foci of republican thinking in France vis-à-vis the immigrant “other” over the past century. The following account relies on Zuber (2008).

⁴For a comprehensive discussion of that history, as the state statistical bureaus navigated between a desire for objectivity of categories and one that could accurately depict the legal and historical aspects of the French population, see Simon (1999).

⁵HALDE is now institutionally embedded in “Défenseur des Droits,” an independent constitutionally-mandated authority (since 2008) charged with ensuring the protection of citizens’ rights and liberties, and promoting equality (see <http://defenseurdesdroits.fr>). It appears it has

the employment of different ethnic or religious groups.⁵

II. Political Pressures for Change

With the obvious facts of discrimination (supplemented by a number of sociological investigations that skirted legal constraints) in the public realm, INED (*Institut National d'Etudes Démographiques*) researchers collected more socio-cultural information on the French population. Their survey in 1992, called “Mobilité Géographique et Insertion Sociale,” asked respondents about native language, something that served as a proxy for what they called “ethnic belonging.” This created a storm of protest within INED. Opponents to these surveys argued that the survey instrument made arbitrary judgments. Kurds, for example, were distinguished from Turks, but not Catalans from Spaniards (Blum and Guérin-Pace, 2008). Researchers in INED, as well as INSEE, were split between those who felt that republican ideals could best be met by addressing discrimination through the analysis of ethnic data and those who felt republican ideals would best be met with a state that did not recognize its citizens through an ethnic lens.⁶

The controversy has not abated. On the one hand, exceptions to the restrictions from 1978 were granted in the recognition of a need for a clearer understanding of the diverse nature of the population. A law passed in 2006 (“*loi pour l'égalité des chances*”) had provisions that opened the possibility for tests of discrimination that would allow researchers to indicate through subtle signals the ethnicity of job or housing applicants. On the other hand, there were public outcries against the relaxation of republican principles. An inflammatory petition published in the leftist newspaper *Libération* argued that ethnic data were not necessary in the battle against discrimination (Amadieu et al., 2007). Data, the article stipulated, could only provoke inter-ethnic confrontations (“*affrontements communautaires*”). Patrick Simon and co-authors from their perch as INED researchers responded in *Le Monde*, bringing the battle of the statisticians into the public realm. Simon then criticized the radical republicans for making a “choice of ignorance” by seeking “equality through invisibility” (Simon, 2008, 8). Amid this brouhaha, the provision in the 2007 law that allowed for a census of the ethnic/racial

backgrounds of the French population was subsequently declared unconstitutional by the Constitutional Council.

Having experienced the tensions in the immigrant *banlieues* in his role as Minister of Interior, Nicolas Sarkozy, on becoming French President, sought to give France the “statistical tools permitting it to measure its diversity.” He demanded that these tools be “objective and uncontested.” Yazid Sabeg — an Arab born in Algeria, educated by the Jesuits, with a Ph.D. from the Sorbonne, and subsequently a successful businessman — was appointed Commissioner of Diversity and Equality of Opportunity (Smith, 2005). Rather than drafting a new law, he appointed François Héran, at that time Director of INED and president of the European Association of Population Studies, to draft a report addressing the president’s charge (Eeckhout, 2009). Ultimately, the report sought a balance between “*tout ethnique*” and “*l’ethnicité zero*.” The balance would be in defining the circumstances and the guarantees that can make the collection of any such data useful and legitimate.⁷

The trend, then, became greater openness to ethnic data collection. In 2008-2009 a collaboration of INED and INSEE researchers constructed a super-sample of immigrants from diverse backgrounds in order to measure their trajectories of integration into French society (Beauchemin, Hamel and Simon, 2010). Second generation migrants were identified even though 95% of them were French citizens and, by republican standards, should not be distinguished from those French with deeper roots. Furthermore, the religion question was asked. While the data are heavily guarded and released only to members of the scientific community, this was a big step forward for data collection on ethnicities in France.

There is yet another force pushing France down this same path — European research institutions. The European Convention on Human Rights demands fulfillment of anti-discrimination regulations that will likely require France to produce data on its record in combating racial and religious discrimination. Also, EU scientific bureaus provide substantial research funds to academics who collect cross-national data, with ethnic issues get-

broader investigatory powers than did HALDE.

⁶Thomas (2012) describes how Chirac’s Nationality Commission (1988) and the subsequent *loi Méhaignerie*, which changed the nationality law, were able to create an uneasy compromise among the strands of republicanism that ran across France’s political system.

⁷Oddly, given the circumstances of the time with the Muslim question dominant in public debate, this report analyzed data collection for six types of ethnic discrimination but not that of religion. Religion, it is evident, was a cleavage too politically hot to handle.

ting substantial attention. Being part of Europe, in other words, has slowly pushed France toward a European norm of sociological investigations and surveys on ethnicity. Even Dominique Schnapper, as I pointed out earlier, has accommodated to these European pressures and now argues that it is “impossible, politically and morally, for researchers to renounce their role in the creation of the self-awareness of a democratic society by establishing knowledge that is as objective as possible” (Simon, 2008, 29).

III. A New Regime for Disseminating Secret Data

Access to sensitive data that is considered potentially threatening to individual privacy is now facilitated through a simple bureaucratic process by the [French Data Archives for Social Sciences \(Réseau Quetelet\)](#). The Quetelet network is under the responsibility of the Consultative Committee for Data in the Social Sciences (CCDSHS). It was created by decree in 2001 by the French ministries of the economy, employment, and national education and research. It reports directly to the minister in charge of research, and it has the responsibility to accede access to sensitive micro-data, to provide support for its utilization, and to document (through archiving and documentation) those data so that they are useful for research. Researchers register their scientific projects with the Réseau Quetelet and, if approved for meeting standards of scientific objectivity, can receive columns of data that are withheld from the public.

In a collaborative project with Jens Hainmueller in our new [Laboratory of Migration and Integration](#) at Stanford, we sought to re-analyze data from the *Trajectoires et Origines* dataset. However, the answers to the question of the professed religion of a survey respondent were not publicly available. Since our project was certified as scientifically objective by the Réseau Quetelet, we were quickly sent those data linked to an identification number for each respondent, thereby permitting easy merging.

However, with more sensitive data, there are imposed strict rules of confidentiality requiring accreditation from the Committee on Secret Statistics (Comité des Secrets). This committee, created in 1984, is an organ of the National Council of Statistical Information (CNIS); its mission is to give its opinion on whether these research projects have safeguards to ensure privacy, and whether they have sufficiently important ends

to justify any risks to individual privacy. In 2008, there was a modification of the law that allowed individual researchers to have access to a wide range of government statistics so long as their research purposes were truly scientific (or historical). The role of the committee is to assess whether the scientific purpose can be achieved in any way without using the sensitive data that was requested, to monitor the required promise that researchers will destroy the files containing the data shortly after the research project has been completed, and to assess whether the researcher has any personal stake in the findings. The committee meets every three months to review proposals, and researchers must appear before the committee to support their scientific claims.

Heavy fines are imposed for any misuse of the data provided to researchers. If approval is granted, the researcher is given a data box to store all of the results from the statistical models, and this can be done with remote access from selected locations within France. However, the final files with statistical tables must get reviewed by a representative of the Committee on Secret Statistics before downloading. But under no conditions would it be possible to download the raw data, as these data never leave the server and are encrypted. This limit may make it impossible to publish work in those journals that demand a replication dataset.

While getting the professed religion of survey respondents took only days to receive (though requiring a hard copy application sent to France), we needed to apply to the Comité des Secrets when we sought data on the precise department and year of arrival of immigrants in the French census. We proposed doing a regression discontinuity design on a program called the Contract of Welcoming and Integration (Contrat d’Accueil et d’Intégration), which was rolled out at different times across departments. We wanted to know if those immigrants who were subject to this Contrat integrated into France more successfully than those who were not subject to it (i.e. those who arrived in France months before the Contrat was imposed). Here, clearly, the disaggregation of the data could conceivably allow us to identify subjects as individuals rather than statistical observations, and thus extra provisions were necessary. After waiting several months to get onto the Comité’s calendar, our proposal was accepted, and we now have access to those data.

Now that we have been granted permission, what are the implications for publication in political science journals demanding access to the raw data as a condition for publication? There needs to be some accommodation, and a precedent for such an accommodation already exists. In a paper that is now in press at the *Journal of Political Economy* (JPE), my co-authors and I relied on data considered sensitive by INSEE and thus required an official authorization to get access to them (Algan, Hémet and Laitin, Forthcoming). Indeed, the nature of the data (information about individuals' national origins) and the very local geographic level at which the study is conducted (housing block) could, if used improperly, violate the privacy of French citizens. As a consequence, these data are not public and we got access to them as part of a convention between INSEE and Sciences Po, the institutional home of my co-authors. We were required to make use of these data within the confines of INSEE and we had to use a secured access point to analyze them. The *JPE* (after accepting the paper) questioned whether we could publish it without access to the raw data by other researchers that would permit replication. We stressed to the editor that these data are available to any researcher once INSEE agrees to the scientific merits of the project through a remote secured system (called CASD) that has been developed over the past five years in France. The *JPE* agreed to an exemption to its replication requirement after we agreed to provide all the necessary information explaining how to get access to these data through the CASD.

France has not lost touch with its republican traditions, and its focus on state responsibility towards its citizens rather than individual researcher responsibility to their subjects represents the French statist tradition. Yet French authorities are slowly coming to terms with the need to know more about their diverse populations. Their evolving policy around a high state council of secrets may sound highly bureaucratic to American ears. However, French concern that their citizens not be thought of as “Muslims” or “Africans” or “of immigrant background,” but rather as “French,” is not an ignoble concern. It is nonetheless an approach to the right to privacy that is increasingly cognizant of the need for the objective evaluation of its integration policies, much to the benefit of the social scientific community and public policy attempting to confront discriminatory processes.

References

- Adida, Claire L., David D. Laitin and Marie-Anne Valfort. 2015. *Why Muslim Integration Fails in Christian-Heritage Societies*. Cambridge, MA.: Harvard University Press.
- Algan, Yann, Camille Hémet and David D. Laitin. Forthcoming. “The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation.” *Journal of Political Economy*.
- Amadiou, Jean-François, Patrick Weil, Dominique Sopo, Samuel Thomas and Mouloud Aounit. 2007. “Engagement républicain contre les discriminations.” *Libération* 23.
- Beauchemin, Cris, Christelle Hamel and Patrick Simon. 2010. Trajectoires et origine. Documents de travail 16 INED and INSEE.
- Bertossi, Christophe. 2012. “French republicanism and the problem of normative density.” *Comparative European Politics* 10(3):248–265.
- Blum, Alain and France Guérin-Pace. 2008. “From measuring integration to fighting discrimination: The illusion of ‘ethnic statistics.’” *French Politics, Culture and Society* 26(1):45–61.
- Bowen, John. 2006. *Why the French Don't Like Headscarves: Islam, the State, and Public Space*. Princeton, NJ: Princeton University Press.
- Eeckhout, Laetitia Van. 2009. “Pour mesurer la diversité, Yazid Sabeg renonce à légiférer.” *Le Monde*.
- Meurs, Dominique, Ariane Pailhé and Patrick Simon. 2006. “The persistence of intergenerational inequalities linked to immigration: Labour market outcomes for immigrants and their descendants in France.” *Population* 61(5):645–682.
- Schnapper, Dominique. 1998. *Community of Citizens: On the Modern Idea of Nationality*. New Brunswick, NJ: Transaction Publishers.
- Schnapper, Dominique. 2007. *Qu'est-ce que l'intégration?* Paris, France: Gallimard.
- Simon, Patrick. 1999. “Nationality and origins in French statistics: Ambiguous categories.” *Population* 11:193–219.
- Simon, Patrick. 2008. “The choice of ignorance: The debate on ethnic and racial statistics in France.” *French Politics, Culture and Society* 26(1):7–31.
- Smith, Craig S. 2005. “What makes someone French?” *New York Times*.
- Thomas, Elaine. 2012. *Immigration, Islam, and the Politics of Belonging in France*. Philadelphia, PA: University of Pennsylvania Press.
- Zuber, Valentine. 2008. “La séparation des églises et de l'État en France et à Genève (1905-1907): Une solution pour deux histoires.” *French Politics, Culture and Society* 26(1):71–91.

Survey Methods for Sensitive Topics

by Graeme Blair

Columbia University

When comparativists rely on survey data, they implicitly invoke an important assumption: that respondents answered truthfully. This assumption may be violated when there are *incentives to conceal the truth*. Social pressures, fears of retaliation, and possible legal sanctions may drive respondents to answer questions in the least revealing way rather than entirely honestly. These incentives can operate when the topic is unmistakably sensitive — for example, in surveys about participation in protests or support for an unpopular policy — but they may also affect seemingly innocuous subjects like voter turnout. When there are incentives to conceal, our inferences about respondents (e.g. what proportion of them shared information with a militant) will be biased. Moreover, in some cases it will be difficult to determine the size or even the direction of the misreporting bias.

Incentives to conceal may even affect whether respondents agree to participate in a survey or to offer a response to a sensitive survey question. If the responses of participants — even truthful ones — differ from those of non-participants, then inferences about the population from the survey data (e.g. what proportion of Nigerians shared information with a militant) will suffer from nonresponse bias.

What can we do about these misreporting and non-response biases? In what follows, I review four survey techniques used by comparativists to address incentives to conceal truthful responses.¹ I first review survey administration practices designed to protect sensitive responses. For contexts in which these are insufficient, I review three experimental methods that can be used in addition that avoid soliciting exact answers to sensitive questions altogether. The experimental methods enable comparativists to ask survey questions that could not otherwise be asked due to ethical concerns and the risk of bias. However, these methods require additional assumptions that are often not testable, necessitating careful design and pilot testing. I conclude with a discussion of common critiques of the experimental techniques.

¹I will discuss how they can enable researchers to solicit truthful answers to binary “yes” or “no” questions, though most of these techniques can be extended to other kinds of outcomes such as numerical responses.

I. Survey Administration Protections

The first and most common approach is to build trust with respondents by implementing and communicating measures to protect sensitive answers through changes in how the responses are collected and stored. If respondents find these confidentiality measures convincing, they may be more likely to respond, and to respond truthfully. Specific approaches used by researchers include:

- Separating sensitive items from names, contact details, and questions that identify individuals such as exact age or family size (permanently, or via codes that can only be accessed by the researcher);
- Locking up paper surveys or encrypting electronic surveys;
- Ensuring that interviews take place in private locations without bystanders;
- Using interviewers who share the age/social group/gender/etc. of the respondent to address fears of outsiders asking sensitive questions;
- Self-administration on paper; via a recording (Chauchard, 2013); by touch-tone telephone; or on a smartphone, tablet, or computer.

In each case, the sensitive survey questions themselves are not modified. As a result, the great advantage to these measures is the simplicity of interpretation and analysis. The “yes” and “no” answers to a sensitive question can be directly analyzed with standard techniques such as means and logit or probit regressions. No additional assumptions beyond truthful responses are required.

II. Randomized Response Technique

The randomized response technique protects respondents by introducing random noise into their responses, so the responses could either reflect a truthful answer to the sensitive question of interest or an irrelevant response. To implement this technique, several design variants exist (Blair, Imai and Zhou, 2015). For example, in the forced response design, a randomization device such as a coin or a die directs each respondent to either automatically answer “yes” or “no,” or to answer the question of interest truthfully. An example on estimating support for coalition forces in Afghanistan illustrates the technique (Blair, Imai and Zhou, 2015):

For this question, I want you to answer yes or no. But I want you to consider the number of your dice throw. If 1 shows on the dice, tell me no. If 6 shows, tell me yes. But if another number, like 2 or 3 or 4 or 5 shows, tell me your own opinion about the question that I will ask you after you throw the dice. [TURN AWAY FROM THE RESPONDENT] Now you throw the dice so that I cannot see what comes out. Please do not forget the number that comes out. [WAIT TO TURN AROUND UNTIL RESPONDENT SAYS YES TO:] Have you thrown the dice? Have you picked it up? Now, during the height of the conflict in 2007 and 2008, did you know any militants, like a family member, a friend, or someone you talked to on a regular basis? Please, before you answer, take note of the number you rolled on the dice.

Individual responses are protected because a “yes” or “no” answer may be a truthful answer or it may indicate that the respondent rolled a 1 or a 6. To identify how many respondents said “yes” to the sensitive question, the “yes” and “no” responses are combined with properties of the randomizing device (i.e. a standard die has approximately a 1/6 probability of landing on each side).

Noise can also be introduced in other ways. In the mirrored question design, a coin or die rolled by the respondent in private assigns her to answer either the sensitive question or its inverse (“yes” and “no” flipped). In the unrelated question design, respondents answer the sensitive question or an unrelated innocuous question.

The chief advantage of the randomized response technique is that no individual response can be exactly identified by anyone. In addition, the level of protection — how much can be learned about sensitive individual responses — is directly controlled by the researcher, who chooses the randomizing device (coin, dice, spinner, etc.). The strong protection and control do come at a cost: compliance with the instructions is assumed, and this may be a strong assumption with low-education respondents or in contexts in which saying “yes” is itself sensitive. There are, however, designs and models to adjust for non-compliance.

For design and analysis advice, see [Gingerich \(2010\)](#)

and [Blair, Imai and Zhou \(2015\)](#).

III. List Experiments

The list experiment conceals individual responses to a sensitive survey item by aggregating those responses with responses to several other control questions. The respondent replies to a list experiment question with a count of the number of “yes” responses to a list of questions that includes the sensitive item. In this way, each individual’s “yes” or “no” response to the sensitive question is hidden within the “yes” and “no” responses to other questions.

A second experimental group is used to estimate the proportion of respondents who said “yes” to the sensitive item (the quantity of interest). In this group, an identical question is asked, except that the list excludes the sensitive item. The average response in this control group is subtracted from the average response to the original question to identify the proportion of respondents who said “yes” to the sensitive item. An example on estimating support for coalition forces in Afghanistan illustrates the technique ([Blair, Imai and Lyall, 2014](#)):

I’m going to read you a list with the names of different groups and individuals on it. After I read the entire list, I’d like you to tell me how many of these groups and individuals you broadly support, meaning that you generally agree with the goals and policies of the group or individual. Please don’t tell me which ones you generally agree with; only tell me how many groups or individuals you broadly support.

Sensitive Item Condition

Karzai Government

Foreign Forces

National Solidarity Program

Local Farmers

Control Item-Only Condition

Karzai Government

National Solidarity Program

Local Farmers

Now, please tell me how many of these groups or individuals do you broadly support?

Several list experiment applications have revealed increased respondent willingness to answer sensitive questions. However, a key disadvantage of the list experiment is that some individual responses are not protected. A respondent in the treatment group who answers that she supports all four groups is identified as a supporter of foreign forces in the example and a respondent who says no groups is identified as a non-supporter of foreign forces. Neither attitude is protected. As a result, some respondents may answer dishonestly to avoid having their views identified. There are methods to detect and adjust for this behavior (Blair and Imai, 2012), but the lack of complete protection may make all respondents more cautious.

For applications in comparative politics, see Corstange (2009), Gonzalez-Ocantos et al. (2012), and Meng, Pan and Yang (Forthcoming). For design and analysis advice, see Corstange (2009), Imai (2011), Blair and Imai (2012), Kramon and Weghorst (2012), Glynn (2013), Aronow et al. (2013), and Imai, Park and Greene (Forthcoming).

IV. Endorsement Experiments

The endorsement experiment is useful for measuring attitudes toward a political actor such as an elected official (see Rosenfeld, Imai and Shapiro (Forthcoming) for an inversion of the design to measure attitudes toward a policy). The method protects individual attitudes toward the sensitive actor by combining them with attitudes towards one or more public policies. In the endorsement condition, randomly assigned respondents are asked a question about a policy and told that the sensitive political actor endorses the policy. An individual's affect toward the actor is protected because a positive response could reflect affinity for either the policy or the actor.

To separate affect toward the actor from policy preferences, randomly assigned respondents are asked an identical question, but without the endorsement. This identifies policy preferences alone. Affect toward the actor is identified by subtracting average policy preferences from the average response to the endorsement question. This is the “endorsement effect.” An example from Lyall, Blair and Imai (2013) that solicits attitudes toward the Taliban illustrates:

Policy-Only Condition
A recent proposal ...

Endorsement Condition

A recent proposal by the Taliban ...

In both conditions underlined above:

...calls for the sweeping reform of the Afghan prison system, including the construction of new prisons in every district to help alleviate overcrowding in existing facilities. Though expensive, new programs for inmates would also be offered, and new judges and prosecutors would be trained. How do you feel about this proposal? Do you strongly agree, somewhat agree, are you indifferent, do you disagree, or do you strongly disagree with this policy?

The endorsement technique is perhaps the most protective of the methods because the direct sensitive question is not asked. Endorsement questions are also easy for enumerators and respondents to understand. There are two main downsides. First, the level of protection for respondents depends on the policies. In the example above, if most respondents disagree with the prison policy, an answer of “strongly agree” would indicate Taliban support. Second, the indirect nature of the question means that the magnitudes of endorsement effects are not directly interpretable without a behavioral assumption.

Examples of applications in comparative politics include Blair et al. (2013) and Lyall, Blair and Imai (2013). For design and analysis advice, see Bullock, Imai and Shapiro (2011) and Blair, Imai and Lyall (2014).

By mitigating the biases from incentives to conceal truthful responses, these four methods enable comparativists to ask questions that could typically only be asked before by building trust over long periods of time with small pools of respondents.

V. Analysis of the Experimental Methods

Regression and other standard analyses can easily be conducted for the randomized response, list, and endorsement techniques using free software in the R statistical environment with the “*rr*,” “*list*,” and “*endorse*”

packages (Blair, Imai and Zhou, 2015; Blair, Zhou and Imai, 2015b; Blair and Imai, 2010; Shiraito and Imai, 2013). The *rr* and *list* packages also enable researchers to use randomized response and list experiment questions as predictors in a regression. Analysis is no more complicated than running a regression in STATA or R. For example, the list experiment regression command is:

```
ictreg( y.variable ~ x.variable, treat =  
"treatment.variable", data = my.data )
```

VI. Critiques of the Experimental Methods

I now discuss five critiques that have been leveled at the experimental methods described above.

1. *There is still no incentive for respondents to answer truthfully.* This is an empirical question, and one for which there is not yet conclusive evidence. Nevertheless, in the small number of validation studies that have taken place, respondents have been more forthcoming with sensitive information and the estimates of known population parameters were closer to the truth using some of these methods compared to questions without protections (Rosenfeld, Imai and Shapiro, Forthcoming).

2. *The list experiment and randomized response technique are confusing to enumerators and respondents.* This is undoubtedly a problem. Carefully developing and pretesting instructions for respondents, directly training all enumerators, and identifying points of confusion by asking respondents a practice question can help.

3. *Respondents see through the designs and do not comply with them to avoid any risk.* An example illustrates this issue: in a list experiment in Afghanistan, zero of 2,754 respondents reported that they supported none or all of the groups mentioned — and these are the two responses that are not protected by the design (Blair, Imai and Lyall, 2014). This problem may be avoided by piloting different designs to find one that provides sufficient protection to encourage participation and compliance.

4. *Low power.* Each method requires a larger sample size than do direct questions. To improve power, researchers can use the “double list experiment” (Glynn, 2013), ask each control question separately in the list experiment control group (Corstange, 2009), ask multiple policy questions for the endorsement experiment, or combine responses from multiple measurements (see Aronow

et al. (2013) on direct questions and list experiments, and Blair, Imai and Lyall (2014) on list and endorsement experiments).

5. *Difficult to design.* Intensive fieldwork is needed to identify appropriate question designs and instructions. Multiple pilot tests are also often needed to design the control items for list experiments and the policies for the endorsement experiment.

VII. Discussion

By mitigating the biases from incentives to conceal truthful responses, these four methods enable comparativists to ask questions that could typically only be asked before by building trust over long periods of time with small pools of respondents. The choice of which method to use will depend on the context, but the choice can be informed by the theoretical differences discussed here and by careful pilot testing. Moreover, the methods are not mutually exclusive. Survey administration protections can be combined with the experimental methods, and multiple experimental methods can be used for measurement (Blair, Imai and Lyall, 2014). Future research should validate these methods in contexts studied by comparativists in order to increase confidence in them and provide more concrete advice on the choice of techniques.

References

- Aronow, Peter M., Alexander Coppock, Forrest W. Crawford and Donald P. Green. 2013. “Combining list experiment and direct question estimates of sensitive behavior prevalence.” *arXiv preprint arXiv:1312.1268*.
- Blair, Graeme, C. Christine Fair, Neil Malhotra and Jacob N. Shapiro. 2013. “Poverty and support for militant politics: Evidence from Pakistan.” *American Journal of Political Science* 57(1):30–48.
- Blair, Graeme and Kosuke Imai. 2010. “list: Statistical methods for the item count technique and list experiment.” Available at *The Comprehensive R Archive Network (CRAN)*.
- Blair, Graeme and Kosuke Imai. 2012. “Statistical analysis of list experiments.” *Political Analysis* 20(1):47–77.
- Blair, Graeme, Kosuke Imai and Jason Lyall. 2014. “Comparing and combining list and endorsement experiments: Evidence from Afghanistan.” *American Journal of Political Science* 58(4):1043–1063.
- Blair, Graeme, Yang-Yang Zhou and Kosuke Imai. 2015a. *Design and analysis of randomized response technique*. Technical report.

- Blair, Graeme, Yang-Yang Zhou and Kosuke Imai. 2015b. “rr: Statistical Methods for the Randomized Response Technique.” Available at *The Comprehensive R Archive Network (CRAN)*.
- Bullock, Will, Kosuke Imai and Jacob N. Shapiro. 2011. “Statistical analysis of endorsement experiments: Measuring support for militant groups in Pakistan.” *Political Analysis* 19(4):363–384.
- Chauchard, Simon. 2013. “Using MP3 players in surveys: The impact of a low-tech self-administration mode on reporting of sensitive attitudes.” *Public Opinion Quarterly* 77(S1):220–231.
- Corstange, Daniel. 2009. “Sensitive questions, truthful answers? Modeling the list experiment with LISTIT.” *Political Analysis* 17(1):45–63.
- Gingerich, Daniel W. 2010. “Understanding off-the-books politics: Conducting inference on the determinants of sensitive behavior with randomized response surveys.” *Political Analysis* 18(3):349–380.
- Glynn, Adam N. 2013. “What can we learn with statistical truth serum? Design and analysis of the list experiment.” *Public Opinion Quarterly* 77(S1):159–172.
- Gonzalez-Ocantos, Ezequiel, Chad Kiewiet De Jonge, Carlos Meléndez, Javier Osorio and David W Nickerson. 2012. “Vote buying and social desirability bias: Experimental evidence from Nicaragua.” *American Journal of Political Science* 56(1):202–217.
- Imai, Kosuke. 2011. “Multivariate regression analysis for the item count technique.” *Journal of the American Statistical Association* 106(494):407–416.
- Imai, Kosuke, Bethany Park and Kenneth F. Greene. Forthcoming. “Using the predicted responses from list experiments as explanatory variables in regression models.” *Political Analysis*.
- Kramon, Eric and Keith R. Weghorst. 2012. “Measuring sensitive attitudes in developing countries: Lessons from implementing the list experiment.” *Newsletter of the APSA Experimental Section* 3(2):14–22.
- Lyall, Jason, Graeme Blair and Kosuke Imai. 2013. “Explaining support for combatants during wartime: A survey experiment in Afghanistan.” *American Political Science Review* 107(04):679–705.
- Meng, Tianguang, Jennifer Pan and Ping Yang. Forthcoming. “Conditional receptivity to citizen participation evidence from a survey experiment in China.” *Comparative Political Studies*.
- Rosenfeld, Bryn, Kosuke Imai and Jacob Shapiro. Forthcoming. “An empirical validation study of popular survey methodologies for sensitive questions.” *American Journal of Political Science*.
- Shiraito, Yuki and Kosuke Imai. 2013. “Endorse: R package for analyzing endorsement experiments.” Available at *The Comprehensive R Archive Network (CRAN)*.

Randomized Response: Foundations and New Developments

by Daniel Gingerich

University of Virginia

I. Overview

Randomized response refers to a family of survey techniques designed to alleviate bias due to social desirability and legal concerns in studies of sensitive behaviors and attitudes. The distinguishing feature of randomized response surveys is the introduction of a randomizing device into the survey response process in order to guarantee the confidentiality of responses to potentially jeopardizing survey items. In a randomized response survey, a respondent’s observed response about a sensitive trait is jointly a function of the sensitive trait in question as well as the private (i.e. known only to the respondent) realization of the randomizing device.

The dependence of the observed response on the privately viewed realization of the randomizing device makes it impossible to deduce the respondent’s trait status based on the information she divulges to the enumerator. As a consequence, concerns about social approbation or potential legal sanctioning are hopefully alleviated, thereby liberating the respondent to respond freely about the sensitive trait. Indeed, each respondent’s trait status remains unknown to all actors involved in the collection and analysis of the survey, so there is no condition under which a court order or other form of government action could be used to compel researchers to reveal the sensitive behaviors of their respondents. While the realization of the randomizing device viewed by any given respondent is unknown to the analyst, the distribution of realizations generated by the device is known prior to conducting the survey. This *a priori* knowledge makes it possible for the analyst to estimate the proportion of individuals who bear the sensitive trait and develop a better understanding of the factors that contribute to its incidence.

Randomized response came into being nearly fifty years ago with the publication of a seminal paper written by the statistician Stanley Warner (Warner, 1965). In recent years, applications of the technique have included studies of topics as varied as abortion (Lara et al., 2006),

social security fraud (Lensvelt-Mulders et al., 2006), corruption within public bureaucracies (Gingerich, 2010, 2013), the prevalence of xenophobia and anti-Semitism (Krumpal, 2012), cheating by undergraduates (Scheers and Dayton, 1987; Kerkvliet, 1994a; Fox and Meijer, 2008), drug use (Kerkvliet, 1994b; Dietz et al., 2013), demand for pornography and prostitution (De Jong, Pieters and Fox, 2010), sexual attitudes (De Jong, Pieters and Stremersch, 2012), and the impact of anonymity on altruistic behavior in a laboratory setting (List et al., 2004; Franzen and Pointner, 2012). Given the growing interest within comparative politics in sensitive forms of political behavior such as corruption, clientelism, vote brokerage, violent anti-government protest, and support for extremist groups, the technique is one which may be of utility to much future work in the subfield. Moreover, randomized response data lends itself fairly straightforwardly to multivariate statistical analysis, thereby allowing researchers to better comprehend the factors that drive such sensitive forms of political behavior.

II. Examples of Randomized Response Models

As mentioned above, randomized response models refer to a family of survey questioning strategies as opposed to any specific technique. In order to give a sense of how randomized response surveys work, I concentrate here on describing several variants of randomized response surveys appropriate for studying a sensitive trait that is binary in form.¹ The key assumption that allows the analyst to identify the proportion of individuals bearing the sensitive trait based on the observed responses generated by these techniques is called *honesty given protection* (Gingerich, 2010). This assumption requires that, given the privacy guarantee provided by randomized response, respondents answer truthfully as prompted by the randomized response survey design. Throughout the discussion that follows, I will assume that the honesty given protection assumption holds.

Notation. Consider a setting in which each respondent i in a randomly selected sample of size n is queried indirectly about her status on an unobservable sensitive trait, $\theta_i \in \{0$ (“absent”), 1 (“present”)}. Our initial interest resides in estimating $\pi = \mathbb{E}[\theta_i] = \mathbb{P}(\theta_i = 1)$, the proportion of individuals who bear the trait of interest; later though we will also consider means of estimating the impact of determinants of the sensitive trait. The observed randomized response of a respondent is de-

noted by the binary outcome variable $Y_i \in \{0, 1\}$, with $n_1 = \sum_i^n Y_i$.

Warner Model. In the original Warner model, a randomizing device prompts respondents to respond “true” or “false” to one of two statements. The two statements are identical, save for the fact that one is in the positive form and the other is in the negative form. The particular statement to which the respondent is directed to respond depends on the private realization of the randomizing device, thereby ensuring confidentiality. Figure 1 on the next page gives an example of a question asked using the Warner variant of the randomized response technique. The sensitive trait of interest is whether or not the respondent would be willing to pay a bribe to avoid a traffic ticket. The respondent is given a spinner divided into two regions A and B (with the probability that the spinner lands in region A equal to p , where $p \neq \frac{1}{2}$). Prior to answering the question, she is directed to spin the spinner, observing the section where the spinner lands in private. If the spinner lands in section A, she is directed to respond true or false with respect to statement A (the statement in the positive form). If the spinner lands in section B, she is directed to respond true or false with respect to statement B (the statement in the negative form).

The confidentiality of respondents’ responses is provided by the fact that an observed response of “true” or “false” could refer to either the statement in its positive form or negative form, and, *ipso facto*, could in each case be either an admission of possession of the trait or denial of possession of the trait.

Let $Y_i = 1$ ($Y_i = 0$) denote the event that respondent i ’s observed response is “true” (“false”). Given the structure of the Warner response process and the honesty given protection assumption, the probability of each type of observed response is:

$$\mathbb{P}(Y_i=1) = p\pi + (1-p)(1-\pi) \quad (1)$$

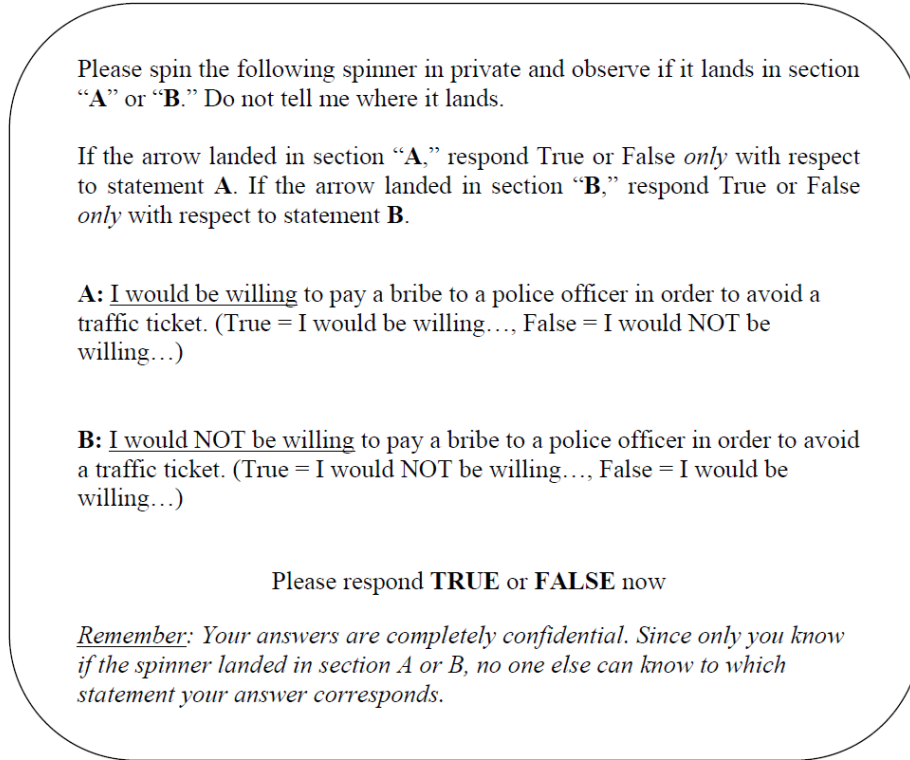
$$\mathbb{P}(Y_i=0) = p(1-\pi) + (1-p)\pi.$$

Letting $I(x)$ be an indicator function equal to 1 if x is true and 0 otherwise, the likelihood of the observed data in the sample is written,

$$\prod_i^n \mathbb{P}(Y_i=1)^{I(Y_i=1)} \mathbb{P}(Y_i=0)^{I(Y_i=0)}, \quad (2)$$

¹More comprehensive descriptions of the varieties of randomized response surveys, including techniques appropriate for multivalued or continuous sensitive traits, can be found in Fox and Tracy (1986), Chaudhuri and Mukerjee (1988), Gingerich (2010), and Blair, Imai and Zhou (2015).

Figure 1: An Example Warner Randomized Response Item



with corresponding log-likelihood equal to,

$$n_1 \ln \mathbb{P}(Y_i=1) + (n - n_1) \ln \mathbb{P}(Y_i=0). \quad (3)$$

Maximizing Eq. (3) with respect to the prevalence rate, π , gives the maximum likelihood estimate of the proportion of individuals who bear the sensitive trait:

$$\hat{\pi}_W = \frac{\frac{n_1}{n} + p - 1}{2p - 1}. \quad (4)$$

The above shows that the prevalence of the sensitive trait can be deduced from the observed proportion of affirmative responses to the randomized response question in conjunction with the known properties of the randomizing device (i.e. the probability that the spinner lands in region A).

Forced Response Model. A commonly used alternative to the Warner model is the so-called forced response model (Boruch, 1971). In the forced response model, respondents are given a randomizing device that generates one of three different outcomes, 1,2,3, say, having corresponding probabilities p_1 , p_2 , and $1 - p_1 - p_2$. The respondent begins by drawing a realization from the device. If the first outcome obtains, the respondent is directed to answer “true” or “false” to a statement indicating possession of the sensitive trait. If the second

outcome obtains, the respondent is directed to simply answer “true.” If the third outcome obtains, the respondent is directed to simply answer “false.”

In this model, the confidentiality of responses stems from the fact that some proportion of the observed “true” and “false” responses are based solely on the realization of the randomizing device and not on the characteristics or experiences of the respondents. As such, respondents prompted to respond in reference to the sensitive trait will know that their observed “true” or “false” response will not reveal their trait status.

As above, let $Y_i = 1$ ($Y_i = 0$) denote the event that respondent i ’s observed response is “true” (“false”). The structure of the forced response data generating process implies that the probability of each type of observed response is as follows:

$$\begin{aligned} \mathbb{P}(Y_i=1) &= p_1\pi + p_2 \\ \mathbb{P}(Y_i=0) &= p_1(1 - \pi) + 1 - p_1 - p_2. \end{aligned} \quad (5)$$

The log-likelihood of the sample for this model is as in Eq. (3), with the probabilities for the observed responses

replaced by those presented above. Maximizing the log-likelihood, the maximum likelihood estimate of the proportion of individuals who bear the sensitive trait for this model is:

$$\hat{\pi}_{FR} = \frac{\frac{n_1}{n} - p_2}{p_1}. \quad (6)$$

As before, the estimate of the prevalence rate is a simple function of the proportion of affirmative responses and the known properties of the randomizing device.

Crosswise Model. The Crosswise Model refers to a sensitive questioning strategy that is mathematically identical to the Warner version of randomized response described above, but one that utilizes an innocuous statement indicating membership in a non-sensitive group in place of a randomizing device (Yu, Tian and Tang, 2008; Tan, Tian and Tang, 2009). Figure 2 on the next page gives an example of a question asked using the Crosswise Model.

Respondents are presented with two statements and asked how many are true. The first statement is the innocuous statement, which indicates that the respondent's mother was born in October, November, or December. The second statement is the statement of interest, which, staying with our running example, denotes that the respondent would be willing to bribe a police officer in order to avoid a traffic ticket.

Protection is provided by constraining the manner in which respondents are permitted to respond. Respondents can give an answer of "A", which indicates that either both statements are true OR neither statement is true, or they can give an answer of "B", which indicates that one of the two statements is true. As a consequence of merging the responses about bribery with those of one's mother's birthday, it is impossible to discern whether or not a respondent is willing to bribe based on the observed response of either "A" or "B".

In order to use this questioning strategy, one has to choose an innocuous statement with special characteristics. First, the proportion of individuals for whom the statement is true must be known in advance by the analyst. Second, the innocuous statement must be structured so that the proportion of respondents for whom it is true is not equal to $\frac{1}{2}$. Third, the membership in the group indicated by the innocuous statement must

be private information (i.e. it must be known to respondents but not to survey administrators and it must be known by respondents to be unknown to administrators). Finally, membership in the group indicated by the benign statement must be uncorrelated with the sensitive behavior of interest. As in the example provided in Figure 2, innocuous statements based on the birthdays of friends or relatives are likely to satisfy all these criteria.²

Although the Crosswise Model has only been in existence for approximately six years, it has already been applied widely to topics such as cheating and plagiarism (Jann, Jerke and Krumpal, 2011; Höglinger, Jann and Diekmann, 2014), petty bribery and drug use (Gingerich et al., 2014), and tax evasion (Krumpal, 2012; Kundt, 2014). These studies have reported significant benefits to using the Crosswise Model over direct questioning and even over other variants of the randomized response technique.

III. The Protection-Variance Trade-off

Relative to direct questioning, the use of randomized response techniques necessarily entails a substantial reduction in the statistical precision of estimates of sensitive items. In this way, randomized response presents the applied researcher with a fundamental trade-off: greater protection of respondents — and, one hopes, lower bias due to legal or social desirability concerns — versus a loss of statistical precision due to the manner in which the technique intentionally introduces noise into the response process.

The highly transparent structure of randomized response designs makes it straightforward to assess the cost component of this trade-off (i.e. the increase in variance due to the use of a particular randomized response design). For the Warner Model, for instance, the variance of the estimate of the prevalence rate is:

$$\text{var}(\hat{\pi}_W) = \frac{\pi(1-\pi)}{n} + \frac{1}{n} \left[\frac{1}{16(p-.5)^2} - \frac{1}{4} \right], \quad (7)$$

where $\frac{\pi(1-\pi)}{n}$ denotes the sampling variance and $\frac{1}{n} \left[\frac{1}{16(p-.5)^2} - \frac{1}{4} \right]$ denotes variation due to the randomizing device. The second component of Eq. (7) captures

²To see that this model is mathematically identical to Warner randomized response, let $Y_i = 1$ ($Y_i = 0$) denote the event that respondent i 's observed response is "A" ("B") and let p denote the proportion of individuals in the population of interest whose mother was born in October, November, or December. Then the probabilities of the two observed responses are exactly as in Eq. (1).

Figure 2: An Example Crosswise Item

How many of the following statements are true?
- My mother was born in OCTOBER, NOVEMBER, OR DECEMBER
- In order to avoid paying a traffic ticket, I would be willing to pay a bribe to a police officer
<u>please indicate your answer below</u>
A. <u>both</u> statements are true OR <u>neither</u> statement is true
B. <u>one</u> of the two statements is true

Remember. Your mother's birthdate is unknown to anyone involved in the collection, administration, or analysis of this survey. As such, your confidentiality is guaranteed.

how the characteristics of the randomizing device affect the precision of the estimate of the prevalence rate. The variance of the estimate increases as the probabilities of being prompted to respond to the statement in the positive form and the statement in the negative form converge to one another (i.e. as p approaches $1/2$), and it decreases as the probabilities of being prompted to respond to the statement in the positive form and the statement in the negative form diverge from one another (i.e. as p approaches 0 or 1). Therein lies the trade-off between protection and variance. As p approaches $1/2$, the level of information contained in any observed response about whether or not the respondent has the sensitive trait gets smaller. This provides the respondent with greater protection. However, and for the exact same reason, the variance of the point estimate explodes as this happens. Given these considerations, a natural question for the applied researcher is how to structure the randomizing device so that the survey provides both a reasonable level of respondent protection and statistical precision. In the Warner model, for instance, how should one set p ? If one were willing to assume that the honesty given protection would hold at *any* level of protection then this problem has a simple solution: minimize variance by setting the protection level as low as possible (e.g. by setting p very close to 0 or 1 in the Warner Model). Of course, it is unreasonable to believe that the honesty given protection assumption would hold were one to do so.

One potentially useful way to think about the appropriate level of respondent protection is in terms of the minimally sufficient condition for plausible deniability. In other words, the applied researcher could attempt to gauge how low the level of protection could be set before respondents begin to feel that they could no longer plausibly deny possession of the trait based on their observed responses. Identifying this threshold is something that could possibly be explored within the context of focus group sessions. More formally, researchers might perform experiments examining how randomized response prevalence estimates change as a function of the level of protection, then use these findings, taken in conjunction with the relevant variance formulas, for optimally selecting the protection level afforded to respondents. As it stands now, however, the level of protection chosen in applied work seems to be chosen largely on the basis of intuition and sample size considerations. As such, I believe this is an area of sensitive survey design where further methodological development would be very useful.

IV. Extensions

I now consider several extensions to standard randomized response models. These include the incorporation of responses generated by direct questioning, explanatory models capturing the influence of predictors

of the sensitive trait, and item response theory models for modeling multiple randomized response questions.

Incorporating Direct Responses

The working theory that leads one to use a sensitive survey technique like randomized response is typically not that everyone lies under direct questioning, but rather that those who have the sensitive trait of interest generally do, or they do so at such a rate as to invalidate analyses based on direct questions about the trait. However, even if we believe that evasiveness under direct questioning would be extensive, in any given application it is probably unreasonable to assume that *all* individuals bearing the sensitive trait would lie or refuse to respond if queried directly. As such, sensitive survey techniques — used in isolation — throw away some potentially very useful pieces of information, namely, the affirmative responses of those respondents who would have been willing to tell us under direct questioning that they bear the sensitive trait.

Here I provide a brief overview of a framework developed in [Gingerich et al. \(2014\)](#) for jointly incorporating randomized response and direct responses in surveys of sensitive behavior. That paper shows that there is a sense in which the joint approach offers the best of both worlds: the resulting estimator of the prevalence rate (or the estimator of the influence of a determinant of the sensitive trait) inherits all of the bias-reducing advantages of randomized response, while at the same time enjoying significant improvements in statistical precision (relative to randomized response alone) due to the incorporation of direct responses. Moreover, the use of the joint response model permits one to estimate a series of diagnostic parameters that reveal the extent to which a sensitive survey technique is truly necessary to study the phenomenon of interest in a given population.

The joint response approach is easy to describe. Each respondent in the sample is first queried about her status on the sensitive trait using a randomized response technique. Then, at a later stage in the survey, each respondent is asked to respond directly to a question about her status, with the explicit option of “choose not to respond directly” offered to her. Thus, the observed outcomes for each respondent consist of a vector $Y_i = (y_i^D, y_i^A)$, where $y_i^D = \{0$ (“absent”), 1 (“present”), \emptyset (“unwilling to respond directly”)} is the observed response when i is queried about the sensi-

tive trait directly and $y_i^A \in \{0, 1\}$ is the observed response when i is queried about the sensitive trait using the randomized response (or crosswise) technique designed to guarantee anonymity. The observed response set is thus an array with six distinct elements, $\mathcal{Y} = \{(0, 0), (0, 1), (1, 0), (1, 1), (\emptyset, 0), (\emptyset, 1)\}$, with $k \in \mathcal{Y}$ representing an arbitrary element in this set. Henceforth, I will use the simplification $Y_i \in \mathcal{Y} = \{1, 2, \dots, 6\}$, where each natural number 1, ..., 6 represents one of the six distinct response combinations.

In addition to honesty given protection, the joint response approach makes an assumption called one-sided lying. This assumption is based on the assumed direction of social desirability bias in sensitive surveys. It holds that individuals who do not bear the sensitive trait never falsely claim that they do. Rather, the set of potential liars is limited to those respondents who do bear the sensitive trait. Let λ_θ^T , λ_θ^L , and $1 - \lambda_\theta^T - \lambda_\theta^L$ denote the probability that, when asked directly, a respondent whose status is θ tells the truth about her willingness to bribe, lies about her willingness, or refuses to answer the question, respectively. One-sided lying specifically implies that $\lambda_0^L = 0$.

Randomized response presents the applied researcher with a fundamental trade-off: greater protection of respondents — and, one hopes, lower bias due to legal or social desirability concerns — versus a loss of statistical precision due to the manner in which the technique intentionally introduces noise into the response process.

Now suppose that the analyst uses either the Warner randomized response model or the Crosswise Model as the sensitive question technique. Then given the assumptions of honesty given protection and one-sided lying, the probability of each combination of responses in the observed response set is as in Table 1 on the next page. The parameter vector to be estimated is $\xi = (\pi, \lambda_1^T, \lambda_1^L, \lambda_0^T)$, where the final three elements in this vector can be thought of as diagnostic parameters. These indicate the need (or lack thereof) to use a sensitive questioning technique to study the trait of interest in the target population and are substantively important in their own right. In particular, λ_1^T expresses the probability

Table 1: Outcome Probabilities for the Joint Response Model

Y_i	Outcome	Probability
1	$(y_i^D = 0, y_i^A = 0)$	$p\lambda_0^T(1 - \pi) + (1 - p)\lambda_1^L\pi$
2	$(y_i^D = 0, y_i^A = 1)$	$(1 - p)\lambda_0^T(1 - \pi) + p\lambda_1^L\pi$
3	$(y_i^D = 1, y_i^A = 0)$	$(1 - p)\lambda_1^T\pi$
4	$(y_i^D = 1, y_i^A = 1)$	$p\lambda_1^T\pi$
5	$(y_i^D = \emptyset, y_i^A = 0)$	$p(1 - \lambda_0^T)(1 - \pi) + (1 - p)(1 - \lambda_1^T - \lambda_1^L)\pi$
6	$(y_i^D = \emptyset, y_i^A = 1)$	$(1 - p)(1 - \lambda_0^T)(1 - \pi) + p(1 - \lambda_1^T - \lambda_1^L)\pi$

that an individual bearing the sensitive trait would be willing to respond honestly under direct questioning.

The log-likelihood for this model is equal to:

$$\ln L(\xi|Y) = \sum_{k=1}^6 n_k \ln \mathbb{P}_Y(k), \quad (8)$$

where $\mathbb{P}_Y(k)$ is the probability of observing $Y_i = k$ and $n_k = \sum_{i=1}^n I(Y_i = k)$ is the number of respondents exhibiting response category k . [Gingerich et al. \(2014\)](#) describes how to use the E-M algorithm to estimate the parameters of this model.

Explanatory Randomized Response

For many social scientists, the real attraction of sensitive survey techniques like randomized response rests with the fact that they can be easily utilized to study the impact of determinants of sensitive behavior. Not surprisingly, the theory and practice of randomized response regression models has a fairly long history ([Maddala, 1986](#); [Scheers and Dayton, 1988](#); [Van Der Heijden and van Gils, 1996](#); [Van Der Heijden, Bouts and Hox, 2000](#); [Lara et al., 2006](#); [Lensvelt-Mulders et al., 2006](#); [Gingerich, 2013](#); [Blair, Imai and Zhou, 2015](#)).

From a likelihood-perspective, setting up an explanatory model for the sensitive trait simply entails replacing the unconditional probability π with an appropriately parameterized conditional probability,

$$\pi_i = f(\mathbf{X}_i; \beta), \quad (9)$$

where \mathbf{X}_i is a vector of background characteristics and a constant, β is a parameter vector, and f is a function from the real line into the interval $[0, 1]$ such as the inverse-logit or probit function. Given the model for the sensitive trait, the analyst inserts π_i in place of π in the

expressions for the probabilities of each type of observed response.

If the outcome consists of randomized response data only, then the log-likelihood for the sample becomes:

$$\ln L(\beta|Y, \mathbf{X}) = \sum_{i=1}^n [I(Y_i = 1) \ln \mathbb{P}(Y_i = 1|\mathbf{X}_i) + I(Y_i = 0) \ln \mathbb{P}(Y_i = 0|\mathbf{X}_i)], \quad (10)$$

where $\mathbb{P}(Y_i = 1|\mathbf{X}_i)$ and $\mathbb{P}(Y_i = 0|\mathbf{X}_i)$ are defined based on Eq. (1) or Eq. (5), depending on which randomized response method is utilized.

Similarly, if the outcome data consist of both randomized responses and direct responses, then the log-likelihood is equal to:

$$\ln L(\xi|Y, \mathbf{X}) = \sum_{i=1}^n \sum_{k=1}^6 I(Y_i = k) \ln \mathbb{P}_Y(k|\mathbf{X}_i), \quad (11)$$

where the $\mathbb{P}_Y(k|\mathbf{X}_i)$ are defined as in Table 1 (for Warner randomized response or the Crosswise Method) and $\xi = (\beta, \lambda_1^T, \lambda_1^L, \lambda_0^T)$. The reader is referred to [Blair, Imai and Zhou \(2015\)](#) and [Gingerich et al. \(2014\)](#) for a discussion of estimation and inference with statistical models of this form.

Non-parametric or semi-parametric program evaluation techniques, such as matching, stratification, and propensity score weighting, can also be used to study the impact of drivers of sensitive behavior using randomized response data ([Gingerich, 2010](#)). Suppose that the goal of the analyst is to estimate the impact of a directly observed binary treatment variable, $T_i \in \{0, 1\}$, and that the outcome of interest is measured using the randomized response technique. Assuming that treatment assignment is ignorable given the covariates \mathbf{X} , the average treatment effect (i.e. the average unit-level impact of

being assigned $T = 1$ instead of $T = 0$) can be identified from the data as:

$$\tau = \mathbb{E}_X \{ \mathbb{E}[Y|T = 1, \mathbf{X} = x] - \mathbb{E}[Y|T = 0, \mathbf{X} = x] \} / c, \quad (12)$$

where $c = 2p - 1$ if the Warner randomized response or Crosswise Models are used, and $c = p_1$ if the forced response model is used. (The notation \mathbb{E}_X denotes that the outer expectation is taken with respect to the covariates). The numerator of the expression above is the average treatment effect calculated using the observed randomized responses as the outcome variable. Thus, the analyst can estimate the average treatment effect for the sensitive trait by first estimating the average treatment effect for the observed reports — using matching, stratification, or propensity score weighting — then dividing this quantity by the constant c .

Item Response Theory Modeling of Multiple Randomized Responses

In applied studies, researchers may be interested in measuring the latent sensitive attitudes of individuals or in developing an understanding of the influence of factors that contribute to such attitudes. A growing body of work on the use of item response theory models with randomized response data provides a framework for accomplishing these goals (Böckenholt and Van der Heijden, 2007; Böckenholt, Barlas and Van Der Heijden, 2009; Fox, 2005, 2010; Fox and Wyrick, 2008).

The setting of interest for the use of these techniques are instances in which the researcher would like to capture an attitude that is sensitive and inherently continuous. An example might be support for the use of torture in order to extract information from suspected terrorists. Rather than ask the respondent directly about this latent trait, since support for torture is an inherently ambiguous concept, the analyst asks the respondent a series of questions about support for particular practices that many would classify as torture: threatening a detainee's family members, sleep deprivation, rectal "feeding," waterboarding, and so on. Since support for each of these practices is also potentially sensitive, these questions are all asked in randomized response format.

Given this structure of questioning, the observed responses consist of binary responses to J different randomized response items, with the response to the j th item denoted by Y_{ij} . These items are tied structurally to θ_i , which denotes the value of the continuous and sen-

sitive latent trait of interest for respondent i . The goal of the analyst is either to estimate θ_i for all i or to estimate a model of determinants of θ_i . It is assumed that variation in the latent trait is responsible for variation in a series of J sensitive and unobservable items (e.g. variation in support for torture is responsible for variation in support for the specific practices described above). Let $\tilde{\theta}_{ij} \in \{0, 1\}$ denote the unobserved value on the j th sensitive item for respondent i (e.g. whether or not i supports waterboarding). The commonly used normal ogive item response model assumes that the responses on the items are linked to the latent trait through the following relationship:

$$\tilde{\pi}_{ij} = \mathbb{P}(\tilde{\theta}_{ij} = 1 | \theta_i) = \Phi(a_j \theta_i + b_j), \quad (13)$$

where Φ is the standard normal cdf and a_j and b_j are item parameters measuring the discriminating power and difficulty of item j , respectively.

The relationship between the underlying outcomes on the unobservable sensitive items and the observed randomized responses are then expressed via a modification of the equations in (1) or (5), depending on which particular randomized response model is utilized. If the Warner or Crosswise Model is used, for instance, the probabilities of the observed responses are written:

$$\mathbb{P}(Y_{ij}=1) = p\tilde{\pi}_{ij} + (1-p)(1 - \tilde{\pi}_{ij}) \quad (14)$$

$$\mathbb{P}(Y_{ij}=0) = p(1 - \tilde{\pi}_{ij}) + (1-p)\tilde{\pi}_{ij}.$$

In order to identify the parameters of the baseline randomized response IRT model, $\xi = (\theta, \mathbf{a}, \mathbf{b})$, it is typically assumed that θ_i is distributed normally with mean zero and unit variance. Alternatively, if the analyst wishes to estimate an explanatory model for the latent sensitive trait, then she may further specify the equation:

$$\theta_i = \mathbf{X}_i^\top \beta + \varepsilon_i, \quad \varepsilon_i \sim N(0, 1),$$

in which case the parameters to be estimated are $\xi = (\beta, \mathbf{a}, \mathbf{b})$. See Fox (2010) for a detailed discussion of a Bayesian approach to estimation and inference with models of this kind.

V. Evidence from Validation Studies

The performance of the randomized response technique in uncovering sensitive behaviors has been examined in a number of validation studies in which the true prevalence of the sensitive behavior in the sample is known by researchers in advance. These studies compare

randomized response to other questioning methods in terms of the proximity of the prevalence rates generated by each questioning method to the true rate. The overall tenor of the results produced by these studies, with a couple of exceptions, is generally favorable to the use of the method.³

Most common are validation studies comparing randomized response to direct questioning. In an early study of this kind, [Locander, Sudman and Bradburn \(1976\)](#) compared the performance of randomized response to three forms of direct questioning for estimating the prevalence of voter registration, library card ownership, bankruptcy declaration, voting in primary elections, and drunken driving arrests. Using a sample of 680 adults in Chicago, the authors found that randomized response produced estimates closest to the true values for the most sensitive items under consideration — bankruptcy and drunk driving — but the small number of observations per outcome prohibited them from reaching a definitive conclusion about the performance of the technique. More definitive findings were produced by [Lamb Jr. and Stem Jr \(1978\)](#), who compared the performance of direct questioning versus randomized response in estimating the incidence of failing course grades in a convenience sample of 312 upper-division undergraduates. They found that whereas both direct questioning and randomized response only slightly underestimated the proportion of students who failed courses, direct questioning severely underestimated the average number of failed courses and randomized response did a much better job in estimating this quantity.

Several validation studies have been based on public arrest records or court proceedings. In this vein, [Tracy and Fox \(1981\)](#) compared the performance of direct questioning versus randomized response in estimating the frequency of arrests in a sample of 530 arrestees in Philadelphia. The authors found that randomized response did better in estimating the overall mean number of arrests than direct questioning. In a more recent study, [Van Der Heijden, Bouts and Hox \(2000\)](#) compared the performance of computer-assisted self-interviews (CASI), face-to-face direct questioning, and two varieties of randomized response in estimating the prevalence of welfare and unemployment benefit fraud in a sample of 534 Dutch adults previously found guilty

of fraud. The authors found that the two randomized response procedures produced prevalence rates much closer to the true value than CASI or face-to-face direct questioning. In a similar investigation, [Wolter and Preisendörfer \(2013\)](#) performed a comparison of the level of misreporting of past convictions of criminal offenses across randomized response and face-to-face direct questioning in a sample of 552 German adults who had all been previously convicted of minor criminal offenses. These authors found that randomized response produced a prevalence estimate closer to the true value than direct questioning, but the difference between the two estimates was negligible.

Two validation studies have examined the incidence of smoking using biomarkers. In the first, [Akers et al. \(1983\)](#) examined the prevalence of teenage smoking using levels of salivary thiocyanate recorded in saliva as an indicator of past smoking. They found that the randomized response estimates of smoking prevalence were nearly identical to the direct questioning estimates and that the pattern of direct responses very closely tracked levels of salivary thiocyanate, implying that the use of randomized response was unnecessary to study the incidence of smoking in this population. A very different finding was encountered by [Fox, Avetisyan and Palen \(2013\)](#) when they compared the performance of randomized response versus direct questioning in detecting smoking behavior among 305 outpatients of a pulmonary department in a hospital in the Netherlands. In this study, the true smoking status of each patient was assessed by recording the CO level in exhaled air using a portable CO monitor. The authors found that being a smoker (based on the CO measure) much more powerfully predicted reported smoking when randomized response was utilized than when direct questioning was employed, thereby establishing the value of randomized response questioning for this particular population.

The most systematic and rigorous validation study involving randomized response is provided in a recent paper by [Rosenfeld, Imai and Shapiro \(Forthcoming\)](#). This work examines the performance of four survey methods for asking about sensitive topics: direct questioning, the list experiment, endorsement experiments, and the randomized response technique. In order to compare the questioning methods, the study used a survey of 2,655 voters who participated in the 2011 Missis-

³All of these studies focus on variants of randomized response techniques developed prior to the Crosswise Model. To the best of my knowledge, the Crosswise Model has not yet been evaluated in a formal validation study.

ssippi General Election (which featured a controversial anti-abortion referendum). Performance was gauged by comparing the proportion of “no” votes obtained using a given questioning method to the true proportion of “no” votes as recorded in the election (at both the county and state-level). Given Mississippi’s conservative bent, a “no” vote was the potentially stigmatizing behavior. The study found that randomized response performed better than all the alternatives, producing estimates of the proportion of “no” votes nearly identical to the true proportion (and with minimal variance relative to the other indirect questioning techniques).

VI. Some Practical Advice

I conclude by offering some practical advice to researchers who are contemplating the use of randomized response survey techniques in their applied work. The general point that needs to be made up front is that successfully implementing a randomized response survey is very hard work. Without careful attention to how these techniques are explained to and understood by respondents, randomized response surveys can easily fail to meet their objectives. Indeed, there are a couple of papers that have reported implementation failures in applications of randomized response, especially with the forced response variant of the technique (Edgell, Himmelfarb and Duchan, 1982; Azfar and Murrell, 2009; Holbrook and Krosnick, 2010). With these issues in mind, I have several concrete suggestions for maximizing the chances of success for a randomized response survey in the field.

First, the administration of randomized response surveys should ideally be conducted in a face-to-face format by enumerators who have undergone extensive training in the logic and implementation of the particular technique they will employ. Before responding to the randomized response questions, every respondent should receive an explanation of why the technique is being used and how it works. Thereafter, and prior to the presentation of the randomized response questions of interest, respondents should be presented with several examples of how hypothetical individuals, varying in trait status and their observed realizations of the randomizing device (or the benign group indicator if the Crosswise Model is used), would respond to a particular question according to the randomized response proto-

col. All this takes time, and survey researchers should be prepared to shorten other components of their survey instruments in order to give these items the attention they require. Indeed, researchers would do well to directly train the enumerators in administering the randomized response component of the survey, and require any survey firm they employ to utilize only the enumerators they have trained to deliver the survey in the field. In my mind, the importance of enumerator-respondent interaction in explaining the randomized response technique generally counsels against its use in telephone, mail, or internet surveys.⁴

As this overview makes clear, there are a variety of different forms of randomized response models, and there can be multiple ways in which to present questions that are based upon a given variant of the technique. In order to get a sense of which modality of the technique works best with a given population, researchers would also do well to make use of carefully selected focus groups prior to fielding a survey. Focus group sessions can be invaluable in helping the survey researcher to ascertain how framing randomized response questions differently can affect comfort with the technique, and they can also be very helpful in assessing the success of alternative scripts that introduce and explain the technique.

Once a variant of randomized response technique and its mode of presentation has been selected, researchers should field test the survey instrument in a small scale pilot study. In the pilot study, enumerators should be accompanied by silent observers from the research team. These individuals view the entire process of survey administration with an eye to detecting any challenges in the delivery of the randomized response component of the survey or any overt signs of respondent discomfort with the technique.

Finally, there are quality checks that can and should be incorporated into the survey instrument itself. These include follow-up questions subsequent to the execution of the randomized response component of the survey that assess respondents’ trust in, and understanding of, the technique. By allowing researchers to identify subgroups within the sample for whom the randomized response protocol was unclear or viewed with suspicion, the inclusion of such questions permits robustness exer-

⁴The one exception to this might be the Crosswise Model, which appears to have had success thus far in on-line surveys (Höglinger, Jann and Diekmann, 2014; Jann, Jerke and Krumpal, 2011; Kundt, 2014).

cises that verify whether or not the central findings of a study depend on the potentially problematic responses of such individuals.

References

- Akers, Ronald L, James Massey, William Clarke and Ronald M Lauer. 1983. "Are self-reports of adolescent deviance valid? Biochemical measures, randomized response, and the bogus pipeline in smoking behavior." *Social Forces* 62(1):234–251.
- Azfar, Omar and Peter Murrell. 2009. "Identifying reticent respondents: assessing the quality of survey data on corruption and values." *Economic Development and Cultural Change* 57(2):387–411.
- Blair, Graeme, Kosuke Imai and Yang-Yang Zhou. 2015. *Design and analysis of randomized response technique*. Technical report.
- Böckenholt, Ulf and Peter GM Van der Heijden. 2007. "Item randomized-response models for measuring noncompliance: Risk-return perceptions, social influences, and self-protective responses." *Psychometrika* 72(2):245–262.
- Böckenholt, Ulf, Sema Barlas and Peter G.M. Van Der Heijden. 2009. "Do randomized-response designs eliminate response biases? An empirical study of non-compliance behavior." *Journal of Applied Econometrics* 24(3):377–392.
- Boruch, Robert F. 1971. "Assuring confidentiality of responses in social research: A note on strategies." *The American Sociologist* 6(4):308–311.
- Chaudhuri, Arijit and Rahul Mukerjee. 1988. *Randomized response: Theory and techniques*. New York, NY: Marcel Dekker.
- De Jong, Martijn G., Rik Pieters and Jean-Paul Fox. 2010. "Reducing social desirability bias through item randomized response: An application to measure underreported desires." *Journal of Marketing Research* 47(1):14–27.
- De Jong, Martijn G., Rik Pieters and Stefan Stremersch. 2012. "Analysis of sensitive questions across cultures: An application of multigroup item randomized response theory to sexual attitudes and behavior." *Journal of Personality and Social Psychology* 103(3):543.
- Dietz, Pavel, Heiko Striegel, Andreas G. Franke, Klaus Lieb, Perikles Simon and Rolf Ulrich. 2013. "Randomized response estimates for the 12-month prevalence of cognitive-enhancing drug use in university students." *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 33(1):44–50.
- Edgell, Stephen E, Samuel Himmelfarb and Karen L. Duchan. 1982. "Validity of forced responses in a randomized response model." *Sociological Methods and Research* 11(1):89–100.
- Fox, James Alan and Paul E Tracy. 1986. *Randomized Response: A Method for Sensitive Surveys*. London, UK: Sage Publications.
- Fox, Jean-Paul. 2005. "Randomized item response theory models." *Journal of Educational and Behavioral statistics* 30(2):189–212.
- Fox, Jean-Paul. 2010. *Randomized Item Response Models*. In *Bayesian Item Response Modeling*. New York, NY: Springer.
- Fox, Jean-Paul and Cheryl Wyrick. 2008. "A mixed effects randomized item response model." *Journal of Educational and Behavioral Statistics* 33(4):389–415.
- Fox, Jean-Paul, M. Avetisyan and Van Der J. Palen. 2013. "Mixture randomized item-response modeling: A smoking behavior validation study." *Statistics in Medicine* 32(27):4821–4837.
- Fox, Jean-Paul and Rob R Meijer. 2008. "Using item response theory to obtain individual information from randomized response data: An application using cheating data." *Applied Psychological Measurement* 32(8):595–610.
- Franzen, Axel and Sonja Pointner. 2012. "Anonymity in the dictator game revisited." *Journal of Economic Behavior and Organization* 81(1):74–81.
- Gingerich, Daniel W. 2010. "Understanding off-the-books politics: Conducting inference on the determinants of sensitive behavior with randomized response surveys." *Political Analysis* 18(3):349–380.
- Gingerich, Daniel W. 2013. *Political Institutions and Party-Directed Corruption in South America: Stealing for the Team*. Cambridge, UK: Cambridge University Press.
- Gingerich, Daniel W., Virginia Oliveros, Ana Corbacho and Mauricio Ruiz-Vega. 2014. "When to protect? Using the crosswise model to integrate protected and direct responses in surveys of sensitive behavior." Working Paper.
- Höglinger, Marc, Ben Jann and Andreas Diekmann. 2014. "Sensitive questions in online surveys: An experimental evaluation of the randomized response technique and the crosswise model." *University of Bern Social Sciences Working Paper No 9*.
- Holbrook, Allyson L and Jon A Krosnick. 2010. "Social desirability bias in voter turnout reports tests using the item count technique." *Public Opinion Quarterly* 74(1):37–67.
- Jann, Ben, Julia Jerke and Ivar Krumpal. 2011. "Asking sensitive questions using the crosswise model: An experimental survey measuring plagiarism." *Public Opinion Quarterly* 76(1):32–49.
- Kerkvliet, Joe. 1994a. "Cheating by economics students: A comparison of survey results." *The Journal of Economic Education* 25(2):121–133.
- Kerkvliet, Joe. 1994b. "Estimating a logit model with randomized data: The case of cocaine use." *Australian Journal of Statistics* 36(1):9–20.

- Krumpal, Ivar. 2012. "Estimating the prevalence of xenophobia and anti-semitism in Germany: A comparison of randomized response and direct questioning." *Social Science Research* 41(6):1387–1403.
- Kundt, Thorben Christian. 2014. "Applying 'Benford's Law' to the Crosswise Model: Findings from an online survey on tax evasion." Available at SSRN 2487069 .
- Lamb Jr., Charles W. and Donald E. Stem Jr. 1978. "An empirical validation of the randomized response technique." *Journal of Marketing Research* 15(4):616–621.
- Lara, Diana, Sandra G. García, Charlotte Ellertson, Carol Camlin and Javier Suárez. 2006. "The measure of induced abortion levels in Mexico using random response technique." *Sociological Methods and Research* 35(2):279–301.
- Lensvelt-Mulders, Gerty J.L.M., Peter G.M. Van Der Heijden, Olav Laudy and Ger Van Gils. 2006. "A validation of a computer-assisted randomized response survey to estimate the prevalence of fraud in social security." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169(2):305–318.
- List, John A., Robert P. Berrens, Alok K. Bohara and Joe Kerkvliet. 2004. "Examining the role of social isolation on stated preferences." *American Economic Review* 94(3):741–752.
- Locander, William, Seymour Sudman and Norman Bradburn. 1976. "An investigation of interview method, threat and response distortion." *Journal of the American Statistical Association* 71(354):269–275.
- Maddala, Gangadharrao S. 1986. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press.
- Rosenfeld, Bryn, Kosuke Imai and Jacob Shapiro. Forthcoming. "An empirical validation study of popular survey methodologies for sensitive questions." *American Journal of Political Science* .
- Scheers, N.J. and C. Mitchell Dayton. 1987. "Improved estimation of academic cheating behavior using the randomized response technique." *Research in Higher Education* 26(1):61–69.
- Scheers, N.J. and C. Mitchell Dayton. 1988. "Covariate randomized response models." *Journal of the American Statistical Association* 83(404):969–974.
- Tan, Ming T., Guo-Liang Tian and Man-Lai Tang. 2009. "Sample surveys with sensitive questions: A nonrandomized response approach." *The American Statistician* 63(1):9–16.
- Tracy, Paul E. and James Alan Fox. 1981. "The validity of randomized response for sensitive measurements." *American Sociological Review* 46(2):187–200.
- Van Der Heijden, Peter G.M. and Ger van Gils. 1996. Some logistic regression models for randomized response data. In *Statistical modeling. Proceedings of the 11th international workshop on statistical modeling. Orvieto, Italy*. pp. 15–19.
- Van Der Heijden, Peter G.M., Jan Bouts and Joop Hox. 2000. "A comparison of randomized response, computer-assisted self-interview, and face-to-face direct questioning eliciting sensitive information in the context of welfare and unemployment benefit." *Sociological Methods and Research* 28(4):505–537.
- Warner, Stanley L. 1965. "Randomized response: A survey technique for eliminating evasive answer bias." *Journal of the American Statistical Association* 60(309):63–69.
- Wolter, Felix and Peter Preisendörfer. 2013. "Asking sensitive questions: An evaluation of the randomized response technique versus direct questioning using individual validation data." *Sociological Methods and Research* 42(3):321–353.
- Yu, Jun-Wu, Guo-Liang Tian and Man-Lai Tang. 2008. "Two new models for survey sampling with sensitive characteristic: Design and analysis." *Metrika* 67(3):251–263.

Capturing Ethnic Salience on Surveys in Developing Countries

by Elizabeth C. Carlson

The Pennsylvania State University

The literature on public opinion in the U.S. demonstrates that social norms against expressing racist attitudes cause misestimates on a variety of important measures, including racial animosity (Wittenbrink, Judd and Park, 1997); support for race-related policies such as school integration (Berinsky, 1999); and intended vote for non-white political candidates (Hopkins, 2009; Terkildsen, 1993). Similar taboos against expressing negative attitudes toward certain ethnic, racial, or religious groups also exist elsewhere and presumably bias survey results in these locations as well. Unfortunately, our ability to identify and combat such biases is often limited in developing country contexts. One well-tested means for decreasing social desirability bias is to allow respondents to report their answers in private — by self-administering the survey on paper or a computer — rather than reporting their opinions out loud to another person (Krysan, 1998; Krysan and Couper, 2003). However, this type of privacy is hard to offer in contexts where many respondents are functionally illiterate and thus cannot self-administer a written survey. To be

accessible to the full population, surveys in the developing world are usually conducted orally: an enumerator reads the questions to the respondent and records the responses on her behalf. The choice faced by those conducting surveys on sensitive issues in the developing world is thus between gathering genuine opinions from the subset of the population that can read and write, and selecting a representative sample of the population, while acknowledging that survey mode used to gather the data from such a sample increases the chances that the results will include bias.

Of course, neither of these options is particularly attractive. The question that motivates my own recent work — the conditions under which Ugandan voters prefer a coethnic candidate for office, especially when that candidate is dubiously qualified — cannot be answered under either approach. Since ethnic salience varies across class and education (Eiffert, Miguel and Posner, 2010; Bossuroy, 2011), I cannot reduce my sample to only the literate. But identifying the conditions under which ethnic voting occurs requires, at minimum, an accurate measure of when voters do and do not vote along ethnic lines. Social censure of ‘tribalism’ is a serious concern if we gather vote preference through oral interviews. Though some Africanists argue that African voters are open about their ethnic loyalties (Eiffert, Miguel and Posner, 2010), most data suggests that voters strongly prefer not to reveal such preferences to an enumerator. Though voting returns in many African elections make it clear that voters support candidates of their own ethnicity, most survey respondents who are asked directly report that they do not consider ethnicity when voting, or even that such preferences are morally wrong (Lindberg and Weghorst, 2010; Posner, 2005). We can therefore expect that, without intervention to reduce bias, survey respondents in African countries will systematically under-report voting for candidates of their own ethnicity on public opinion surveys, affecting our understanding of when such voting is likely.

In my study, I used two simple interventions intended to increase the likelihood that respondents would accurately report their true preferences, while maintaining an oral framework accessible to all members of the population (Carlson, Forthcoming, 2014). In the first intervention, I adapted the delivery of the survey to allow private reporting by illiterate respondents. Chauchard (2013) uses such a technique in India. He recorded questions on MP3 players and allowed respondents to self-

report their answer from a menu of options represented with familiar symbols. In Chauchard (2013)’s example, however, the entire survey was self-administered. This limits the content of the survey to only certain types of closed-ended questions. My intervention was much smaller: I embedded a single question answered with a secret ballot into a larger face-to-face interview. Such a method also happens to require only pen and paper, which may be an advantage for very low-resource projects. A second strategy, used with mixed results in the U.S. (Hatchett and Shuman, 1975; Campbell, 1981; Finkel, Guterbock and Borg, 1991) and proposed by Africanists such as Adida et al. (2014), is to match the ethnicity of enumerators and respondents. This intervention is based on the assumption that respondents will be more honest about ethnic preferences when they do not need to worry about offending a member of another group. In my results, I find that allowing limited self-administration effectively reduces bias, while ethnic matching exacerbates the problem.

In my study, I asked 800 Ugandans, drawn from 32 villages, to report their preference between two hypothetical candidates whose ethnicity and qualification for office — such as education and record of performance in prior office — were randomly selected from a list of possible traits. A sample vignette is shown in Figure 1 below. In the vignettes, ethnicity was signaled by the candidate’s city of origin. Performance was indicated by what the candidate accomplished in prior office. In the sample vignette, the first candidate paved a road. The second candidate did not have any accomplishments to note, which in the Ugandan context signals incompetence or corruption. The dependent variable of interest is the percent of respondents who selected a candidate of their own ethnicity. If the interventions work to reduce socially desirability bias, the percent reporting support for a coethnic, regardless of his performance, should increase.

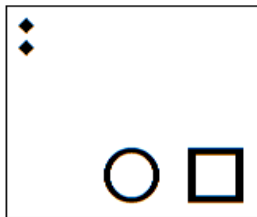
Figure 1: Sample Candidate Pairing

Candidate One. The candidate is from Hoima. He has a Master’s Degree in African Development. Before running for president, he was an LCI [village chief]. While he was LCI, he paved the community’s major road. If elected president he promises to create new jobs.

Candidate Two. The candidate is from Soroti. He has a university degree in Tourism. Before running for president, he was a Member of Parliament. If elected president he promises to improve government health care. (*Note: This candidate is cued as having no record of provision, simply by leaving out mention of such provision.*)

Participants in the experiment were recruited into a 45-minute long survey about local welfare and public goods quality that contained both closed- and open-ended questions and was conducted orally one-on-one with an enumerator. Respondents were then assigned to small groups. Members of each group listened to their assigned vignettes and reported their preference via a secret pen and paper ballot. To enable illiterate respondents to mark their ballot without assistance, the candidates were indicated as a “circle” and a “square” (or “box” as appropriate in the local language). Figure 2 below shows the ballot.

Figure 2: Ballot



There were two key variations in the context of the voting experiment. First, respondents faced two different levels of social exposure: while most respondents cast only a secret ballot, the members of approximately 1/3 of the groups were also assigned to report their preference aloud to everyone in the group. Second, the groups varied in their ethnic composition: conditional on the ethnic diversity of the local community from which the respondents were drawn, whether a respondent voted only with members of her own ethnic group or in a mixed group was random. The first variation tests whether granting respondents privacy will reduce bias; the second tests whether asking respondents to report ethnic preferences to members of only their own ethnicity does the same.

The first conclusion of the experiment is that privacy works to reduce social desirability bias. Those who

cast only secret ballots were significantly more likely to report a preference for a coethnic candidate than those who reported their preferences publicly. Among voters who reported their votes aloud, coethnics had no advantage: 50% of coethnics won their contests, as did 50% of non-coethnic candidates. This result strongly suggests that Ugandan voters do not see any advantage to coethnic candidates. Yet, when voters were allowed to report their vote entirely in secret, it is clear that voters *do* actually hold such a preference. In the secret-ballot condition, coethnic candidates were 6% more likely to win than were non-coethnic candidates. The vote shares of coethnic and non-coethnics in the secret ballot condition are significantly different from their vote shares in the public voting treatment, as well as significantly different from one another. These results indicate that allowing limited self-administration on sensitive questions, rather than requiring oral responses, can reduce bias to the point that it actually changes our conclusions.

These results indicate that allowing limited self-administration on sensitive questions, rather than requiring oral responses, can reduce bias to the point that it actually changes our conclusions.

Reducing observation also decreases socially undesirable responses outside of the experimental Ugandan context. Since almost all existing data on public opinion in Africa was gathered with in-person interviews, I cannot test the effects of allowing full privacy on respondents' reported vote choice. However, I can leverage variation in the number of people in front of whom respondents reported their opinions. Despite enumerators' efforts to prevent it, about a quarter of respondents on the massive, multi-country Afrobarometer public opinion survey nevertheless complete their interview in front of family or friends. If this happens, it is noted at the end of the survey. I pool the results from Round Four (2011-2012) of the Afrobarometer from every country where the major presidential candidates were of different ethnicities, for a total sample of over 25,000 respondents across 18 countries. Controlling for country fixed effects and individual-level characteristics that might make a respondent more likely to be observed, such as age or gender, I find that as the number of observers increases, the likelihood that a respondent will report a preference for a candidate of their own ethnicity decreases. Those whose interviews were observed by family

or neighbors are 10% less likely to report support for a coethnic presidential candidate than those whose surveys were conducted only by an enumerator, at a high level of statistical significance. These results suggest that reducing the number of observers to zero — by allowing complete self-administration, such as with a secret ballot — might increase reported support for coethnic candidates even more.

Though my results indicate that allowing respondents privacy to report their vote preferences will reduce social desirability bias, they also indicate that getting accurate results does not require adapting an entire survey to be self-administered. It is reasonable to be concerned that embedding a secret ballot question in an otherwise oral survey might prime the very issues we are hoping to de-emphasize. There is some evidence in my data that this sort of priming is occurring. In the experiment, approximately 40% of the candidates in the vignettes were assigned to be of the respondent's ethnicity, which is far higher than chance would predict in a country as ethnically diverse as Uganda. By the final round of the experiment, when voters had already heard three vignettes and cast a ballot three times, many respondents seem to have realized they were being asked to reveal ethnic preferences and adjusted their responses accordingly. The votes cast in the third round of the experiment strongly de-emphasized ethnicity regardless of whether votes were cast publicly or privately.

The question is whether the priming that occurred in the private ballot condition produces different answers than the priming that would have occurred without the use of a private ballot: does it, for example, emphasize that such attitudes should be sensitive or secret and exacerbate bias on later questions? I can provide an initial test of this using responses from the post-survey. After participating in the voting experiment, respondents completed a post-survey one-on-one with an enumerator. The questions were potentially sensitive, including two questions explicitly about ethnic salience. On one of these questions — whether the respondent prioritized their ethnic or national identity — the responses of those who voted publicly and privately were the same (67% of respondents in both groups say they prioritize their national identity.) On a second question — whether the respondent would ever marry someone of another group — the groups are significantly different, but in a way that suggests that allowing private voting dampened bias even on later questions asked aloud.

Those who were allowed to vote privately in the voting experiment were 3% more likely to admit in the post-survey that they would never marry someone of another group. The difference is significant at standard levels and robust to a variety of respondent-level controls. Altering survey protocol to allow self-administration of a single question, therefore, does not seem to worsen bias on later questions, and may even reduce it.

The second conclusion of the study is that while privacy works, ethnic matching does not. If it did, we would expect that those respondents assigned to vote in groups composed of only their own coethnics would be more likely to report support for coethnic candidates than those assigned to groups of mixed ethnicity. Instead, those who voted in fully homogenous groups were significantly *less* likely to report a preference for a coethnic candidate. Those who voted in a group comprised only of their coethnics selected coethnics 50% of the time, whereas those who voted in ethnically heterogeneous groups selected coethnics 57% of the time. Though the ethnic composition of the group is correlated with the ethnic diversity in the local community (a diverse group cannot be constructed out of a fully homogenous community), the estimated treatment effect is robust to controls for local ethnic diversity. This result can again be replicated outside the experimental context. Using the same pooled Afrobarometer data I described above, with country-level fixed effects, and a variety of controls, I find that respondents are 7% less likely to report support for the candidate of their own ethnicity when they are interviewed by a coethnic enumerator.

One explanation for this finding is that interactions with non-coethnics may prime ethnic conflict or competition for resources in a way that intra-ethnic interaction does not, triggering a defensive reaction against the idea of a non-coethnic in power and generating a preference for coethnic politicians; this explanation has nothing to do with social desirability bias. However, we also know that some Africans are more generous when playing experimental games with coethnic strangers than they are when playing with non-coethnics (Habyarimana et al., 2007). This appears to be because ethnic networks are so dense that no exchange between coethnics can ever be truly anonymous: exposure to coethnics is more consequential than exposure to non-coethnics. If this is the case, matching respondents with coethnic enumerators may actually produce *stronger* bias in the direction of so-

Table 1: Ethnic Diversity and Social Exposure on Reported Preferences for Coethnic Candidates in Uganda

	Private voting	Public voting	All votes
Ethnically diverse voting group	0.61 (0.03)	0.58 (0.04)	0.60 (0.02)
Ethnically homogenous voting group	0.53 (0.02)	0.44 (0.04)	0.50 (0.02)
All groups	0.55 (0.02)	0.50 (0.03)	0.54 (0.01)

cial norms. The results of my experiment, summarized in Table 1 on the next page, suggest that both mechanisms are in operation: those who were assigned to vote in an ethnically diverse group are more likely to vote for a coethnic whether they vote privately or publicly. However, exposure to social pressure most strongly affects the reported preferences of those in ethnically homogenous groups, indicating that the desire to conform to social norms is stronger among coethnics. This may also explain why observation by family and neighbors, who will almost always be coethnics, has such a strong effect in reducing reported ethnic voting in the survey results presented above. Together, my results provide some guidance for those gathering data on ethnic politics in places where ethnicity is a sensitive subject and surveys must be conducted orally. Though it is very low-cost, and may be effective in some contexts, a strategy of ethnic matching between respondents and enumerators is risky: my results indicate that it exacerbates bias in Uganda and in a pooled African sample. It may also do so in India, the Middle East, and other areas where intra-ethnic social networks are strong and intra-group social sanctions are particularly consequential. The better option is to provide an opportunity to allow self-administration of sensitive questions so that respondents do not need to report their preferences aloud: my results are just one case in which this strategy has been shown to be effective. What my results contribute is a finding that gathering self-administered data on some of a survey's questions does not require writing an entire survey to be self-administered. Adaptation on a few of a survey's most sensitive questions can improve data quality with minimal investment and effect on the rest of the survey's data.

References

Adida, Claire, Karen Ferree, Daniel Posner and Amanda Robinson. 2014. "Who's asking? Interviewer coethnicity effects in African survey data." Unpublished Manuscript.

- Berinsky, Adam. 1999. "The two faces of public opinion." *American Journal of Political Science* 43(4):1209–1230.
- Bossuroy, Thomas. 2011. "Individual determinants of ethnic identification." Dauphine University Paris IRD Working Paper.
- Campbell, Bruce. 1981. "Race-of-interviewer effects among southern adolescents." *Public Opinion Quarterly* 45:231–244.
- Carlson, Elizabeth C. 2014. "Social desirability bias and reported ethnic voting on African surveys." Afrobarometer Working Paper No. 144.
- Carlson, Elizabeth C. Forthcoming. "Ethnic voting and accountability: A choice experiment in Uganda." *World Politics*.
- Chauchard, Simon. 2013. "Using MP3 players in surveys: The impact of a low-tech self-administration mode on reporting of sensitive attitudes." *Public Opinion Quarterly* 77(S1):220–231.
- Eiffert, Benn, Edward Miguel and Daniel Posner. 2010. "Political competition and ethnic identification in Africa." *American Journal of Political Science* 54(2):494–510.
- Finkel, Steven, Thomas Guterbock and Marian Borg. 1991. "Race-of-interviewer effects in a pre-election poll: Virginia 1989." *Public Opinion Quarterly* 55(3):313–330.
- Habyarimana, James, Macartan Humphreys, Daniel Posner and Jeremy Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review* 101(4):709–725.
- Hatchett, Shirley and Howard Shuman. 1975. "White respondents and race of interviewer effects." *Public Opinion Quarterly* 39(4):523–527.
- Hopkins, Daniel J. 2009. "No more Wilder effect, never a Whitman effect: When and why polls mislead about black and female candidates." *Journal of Politics* 71(3):769–781.
- Krysan, Maria. 1998. "Privacy and the expression of white racial attitudes." *Public Opinion Quarterly* 62:506–544.
- Krysan, Maria and Mick Couper. 2003. "Race in live and virtual interviews: Racial deference, social desirability and activation effects in attitude surveys." *Social Psychology Quarterly* 66(4):364–383.
- Lindberg, Staffan and Kieth Weghorst. 2010. "Are swing voters instruments of democracy or farmers of clientelism? Evidence from Ghana." Working Paper.

- Posner, Daniel. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge: Cambridge University Press.
- Terkildsen, Nayda. 1993. "When white voters evaluate black candidates: The processing implications of candidate skin color, prejudice, and self-monitoring." *American Journal of Political Science* 37(4):1032–1053.
- Wittenbrink, Bernd, Charles M. Judd and Bernadette Park. 1997. "Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures." *Journal of Personality and Social Psychology* 72:262–74.

Lobbying, Corruption, and Non-Responses in Small Samples

by Vineeta Yadav

The Pennsylvania State University

The illegal nature of corruption has posed formidable obstacles for researchers studying this complex phenomenon. In the absence of hard objective data, scholars have most often relied on surveys of citizens, firms, and experts to gather perceptions of corruption (Olken and Pande, 2011; Treisman, 2007). However, survey questions on corruption often suffer from high non-response, and possibly false response, rates because they are considered intrusive, because admitting to participating in corrupt practices is considered socially undesirable, and because respondents may fear retaliation from officials for admitting participation in, or knowledge of, corruption (see, for example, reports documenting the implementation of the European Bank for Reconstruction and Development's [ERBD] 2002 Business Environment and Enterprise Performance Survey [BEEPS]).

As Blair, Imai and Zhou (2015), Krumpal (2013), and Tan, Tian and Tang (2009) discuss, the survey methodological toolkit now includes several techniques — randomized responses, list experiments, endorsement experiments — that can help scholars address the problems, such as non-responses and false responses, that are associated with sensitive survey questions. These various techniques focus on reassuring respondents that their responses will be confidential by using methods that share one key characteristic — they sacrifice the ability of the analyst to identify individual responses.

For many research questions, including those related to corruption, sacrificing this ability does not typically impose a high substantive analytical cost. Several recent corruption studies have therefore used these techniques to address some of the challenges posed by sensitivity (Winters and Weitz-Shapiro, 2013; Gingerich, 2010).

However, there are important research agendas where the relationship between the respondent's own behavior and his perception of a collective or aggregate outcome *is* the central research question. This is most notably the case for the burgeoning literature on lobbying and corruption. The central focus in this literature is on studying how the lobbying practices of an individual special interest group — especially its decision to use money rather than other resources — influence its perception of corruption in that country (Yadav, 2011; Bennedsen, Feldmann and Lassen, 2009; Campos and Giovannoni, 2007). The loss of the ability to identify the corruption level reported by a specific respondent in this case deprives the researcher of the ability to study this question.

The second challenge this specific research agenda poses for analysts is that the sample sizes required for successfully implementing the most commonly used techniques for addressing sensitivity are larger than the entire population sizes of the key elite populations that are at the center of the lobbying-corruption nexus. All of these techniques generally require sample sizes of 1,000 or more to be effective. For example, Gingerich (2010, 364) finds that randomized response techniques underperform compared to direct responses for samples smaller than 1,000 and recommends using samples of 3,000 or more. Corstange (2009, 55) finds that list experiments are not suitable for sample sizes smaller than 1,000 and recommends sample sizes of 2,000 or more. This constraint makes these techniques feasible and appropriate for studying perceptions of corruption among large populations such as voters, firms, and bureaucrats, but not for small elite populations. Unfortunately, several populations that are highly relevant for the study of corruption, such as those of party leaders and business lobbies, are too small for these techniques to be used appropriately.

Considerable empirical evidence shows that the actors most likely to be involved in, and knowledgeable about, corruption in parliament and the executive are not citizens or firms, who rarely interact directly with

these bodies, but elite groups in society, such as business interest groups and very large firms, who interact with them frequently (OECD, 2012, 2014; Kalniņš, 2011; Thomas and Hrebendar, 2008; EBRD, 2002). While citizens and firms do interact with the bureaucracy and the legal system, they do so only intermittently (Ayyagari, Beck and Demirguc-Kunt, 2007; International, 2007). In contrast, business lobbies routinely interact with the bureaucracy and the legal system in order to influence policy legislation, implementation, and legal adjudication (OECD, 2012, 2014; Kalniņš, 2011; Thomas and Hrebendar, 2008; EBRD, 2002). Importantly, studies show that in most countries around the world (the U.S. being a notable exception), lobbying is done primarily by formal business associations and chambers rather than by firms themselves or hired third party lobbyists (OECD, 2012, 2014; Kalniņš, 2011; Thomas and Hrebendar, 2008). Indeed, firms that are members in business associations and chambers often identify such intercession with elected officials, the bureaucracy, or the judiciary as one of the most valuable benefits of association membership (OECD, 2012, 2014; Kalniņš, 2011; Thomas and Hrebendar, 2008; EBRD, 2002).

The sample sizes required for successfully implementing the most commonly used techniques for addressing sensitivity are larger than the entire population sizes of the key elite populations that are at the center of the lobbying-corruption nexus.

The problem for corruption scholars is created by the fact that while citizens and firms number in the millions and thousands, the total number of associations, chambers, and other bodies involved directly in lobbying in most countries (with the exception of the U.S. and a handful of wealthy countries) numbers in the hundreds and may cross a thousand for bigger, more sophisticated economies (OECD, 2012, 2014; Kalniņš, 2011; Thomas and Hrebendar, 2008). Very large firms can and often do lobby individually in their own interests. Their numbers, however, are not high enough to boost the population sizes of active business lobbies over the thresholds required for using the survey techniques mentioned previously (Ayyagari, Beck and Demirguc-Kunt, 2007). The small numbers of organized business associations and

very large firms therefore render the use of the standard repertoire of techniques for addressing sensitive topics ineffective and ill-advised in most developing countries.

Finally, an additional concern that few extant studies of corruption address is that the sensitivities of respondents, whether masses or elites, may vary for different types of corrupt behavior. The common use of the term “corruption” encompasses a wide range of phenomena, including corruption in parliament, the executive office, political parties, individual politicians, bureaucrats, judges, police, etc. Different corruption sub-types may be subject to varying sensibilities of social desirability. We would, therefore, expect the sensitivity of respondents and consequently their willingness to respond, and to respond honestly, to corruption related questions to potentially vary across questions asking for their opinions on these different corruption sub-types. With a few notable exceptions, however, the vast majority of surveys on corruption use questions which do not differentiate between these different corruption sub-types.¹ As a result, we have little leverage in understanding which of these sub-types is driving non-responses to a question on “corruption” and on which sub-type of corruption the respondent is in fact expressing his opinion.

The measurement problem this creates is substantial in the cross-country context. Since countries vary in the factors that are believed to influence different sub-types of corruption, non-responses to “corruption” questions may be driven by sensitivity to different sub-types of corruption across (and, of course, within) countries, and responses may be dominated by perceptions of different sub-types of corruption as well. The rate of non-response, as well as the origins and magnitude of the selection bias that non-responses introduce, therefore varies in an opaque way across countries. As a result, whether scholars use survey questions on corruption from individual surveys or indices of corruption that combine information from different surveys, it remains unclear exactly what concept is being examined.

I now briefly illustrate how identifying and collecting data on less sensitive factors that influence non-responses to corruption can allow us to address some of these problems. Specifically, I use such data to study what influences respondents to participate in, or avoid answering, corruption questions and to correct for the

¹The BEEPS and the UN’s International Crime Victimization Surveys do ask questions on distinct corruption behaviors; however, most cross-country surveys do not.

selection bias that participation decisions introduce in comparing corruption levels and behaviors across countries. By asking narrow questions that query respondents on specific sub-types of corruption, we can study whether patterns of non-responses on sensitive corruption questions are similar or different across sub-types of corruption both within and across countries. I use evidence from a 2005–2006 survey of business lobbies in Brazil and India to illustrate these points.

Although Brazil and India both have a reputation for corruption, they have different institutional designs. Existing research tells us that these institutional differences should influence not just the level of aggregate corruption in each country but also the prominence of various corruption sub-types (Graf Lambsdorff, 2005). This suggests that the sensitivity of respondents to questions querying them on different sub-types of corruption may vary across sub-types of corruption within each country and across the two countries, leading to correspondingly different patterns of non-responses and false responses. When I conducted my survey in 2005–2006, the aggregate number of business associations, regional and sectoral chambers, and other bodies involved in policy lobbying at any level was 1,492 in Brazil and 880 in India.² These total *population* sizes are too small to effectively implement techniques such as randomized responses, list experiments, or list endorsements to deal with corruption related sensitivity. Therefore, this research agenda required a different approach.

Given the prominence of business lobbying in Brazil and India (dos Santos and de Costa, 2014; Schneider, 2013; Yadav, 2008), one possibility was to use information on the lobbying tactics of businesses — information, media, public protests, money — to gain leverage on non-responses. This is because the choice of lobbying tactics conceivably influences how willing business lobbies are to respond to questions on different types of corruption. In both countries, lobbying was unregulated but not illegal, corporate financial contributions to parties and politicians were legal but capped, and reporting and disclosure laws for financial donors and recipients were poorly enforced (Yadav, 2011). At first blush, these practices do not make collecting information on lobbying

seem like a promising avenue. What makes this a viable strategy, however, is that, even though it was unregulated, lobbying did not suffer from a social desirability bias among the key actors (the business establishment, government officials, and elected officials) involved in the policymaking process.

In both countries, these policy players recognized business lobbying as useful for acquiring policy information and expertise, and saw it as an endeavor that was legitimate and essential for good policymaking and representation (dos Santos and de Costa, 2014; Schneider, 2013; Yadav, 2011). This view was driven significantly by the fact that lobbying is practiced openly and actively in democracies such as the U.S., U.K., and Germany, countries that are held up as aspirational economic models. In this context, the fact that the study was being sponsored by a U.S. university certainly reduced the discomfort associated with reporting lobbying behaviors as well.

These trends on lobbying legitimacy suggest that questions to business interest groups about their lobbying tactics are unlikely to provoke high non-response rates. If this is the case, then they may offer leverage on understanding who responds to corruption related questions in each country and the relationships across non-responses on different sub-types of corruption, they may account for the bias created by respondents selecting out of responding to each corruption sub-type, and, importantly, they may allow us to study how an individual respondent's lobbying tactics influences her perceptions of corruption in each country. In turn, comparing these results across the two countries may elucidate some of the potential problems that cross-country corruption measures might be hiding.

The surveys were given to high level officials from organized and formal business interest groups from 2005 to 2006.³ The realized sample sizes were 158 in Brazil and 179 in India. Questions asking groups *who* they lobbied (executive office, political parties, individual legislators, regional and sectoral caucuses) saw non-response rate of 3.2% and 7.2% in Brazil and India. Questions querying *how* they lobbied these targets (using technical information, financial contributions, media campaigns, public

²There were, and still are, no existing databases that list all of the organizations involved in lobbying in either country. I put these sampling frames together by consulting and compiling listed organizations from various sources myself. This has been the technique that other researchers, including Frye (2002) and the World Bank's WBES and BEEPs survey supervisors, have adopted. See Yadav (2011) for more details on the survey and sampling designs.

³Please see Yadav (2011) for the exact questions that were used to operationalize these concepts. All were modeled after similar questions asked in the BEEPs 2002 surveys conducted by the World Bank and the EBRD.

protests) saw non-response rates of 4.4% and 7.2% in Brazil and India. In contrast, questions asking respondents how a certain form of corruption “had an impact on businesses in your sector” drew high, but varying, non-response rates for different corruption sub-types in both countries. Figure 1 below illustrates these trends for eight different sub-types of corruption. Notably, non-response rates vary substantially across questions on different types of corruption within each country, ranging from 19.6% to 44.3% in Brazil and 12.2% to 52.8% in India.

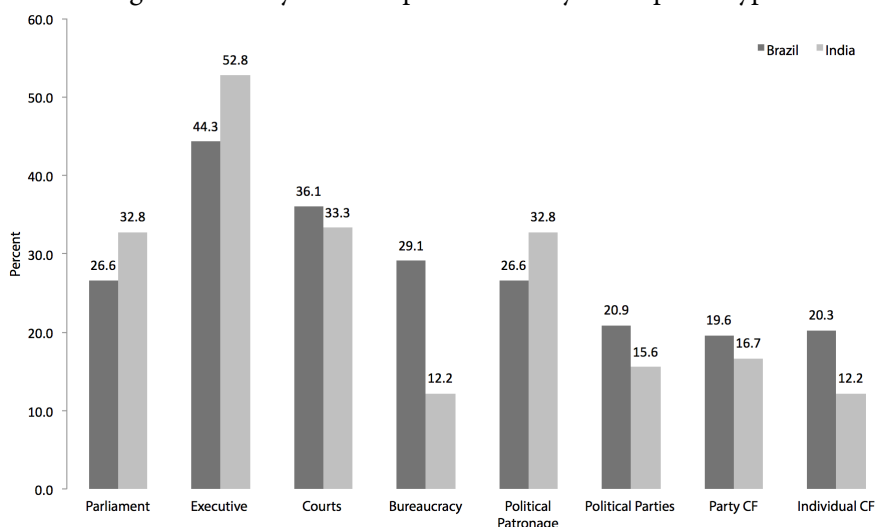
Given the low non-response rates for lobbying questions, it is possible to model a group’s decision to participate in answering questions about each of the eight corruption sub-types listed in Figure 1 as a function of whether that group used technical information, public protests, media campaigns, or money to lobby policy-makers on an issue of vital importance to their members while controlling for other factors that might potentially influence this choice as well. Corruption scales ranged from 1, indicating no impact on the respondent group’s members from that specific type of corruption behavior, to 4, indicating a decisive impact on its members. We can now directly address the question of how a lobby’s own tactics — especially whether or not it used money for lobbying — influence its perception of different types of corruption by estimating a series of ordered probit models explaining corruption levels with a selection stage modeling the choice to respond as a function of lobbying tactics on resources.

The outcome variable in each ordered probit model is one of the eight corruption sub-types listed in Figure 1: parliamentary corruption, executive corruption, party corruption, patronage, bureaucratic corruption, judicial corruption, corruption in party campaign finance, and corruption in individual campaign finance. The selection stage models the decision to participate in responding to the question on that specific corruption sub-type as a function of that group’s tactical choice to use money, information, or other tactics to lobby policy-makers and a few other theoretically relevant variables.

Applying this strategy to Brazil and India allows us to see how these results may be similar or different across these countries (i.e. whether non-responses are driven by similar factors and whether the use of money for lobbying has a similar effect on participation and reported corruption levels in the two countries). Below, I highlight a few key findings from such an analysis of parliamentary, executive, bureaucratic, and judicial corruption. The full results from this analysis are available at <http://sites.psu.edu/vineetayadav/>.

Participation Decisions: Results from the selection stage show that, with the exception of bureaucratic corruption in Brazil, the decision to respond to corruption questions and the decision to report a certain level of corruption were highly correlated in both countries. Resource lobbying tactics seemed to rarely affect the decision to respond to any corruption questions in Brazil. In-

Figure 1: Binary Non-Response Rates by Corruption Type



Note: Data are from a 2005-2006 survey of high level officials in business lobbies in Brazil (dark grey) and India (light grey). There were 158 respondents in Brazil and 179 respondents in India. CF denotes campaign finance.

stead, group-specific features, such as the fate of group members under economic reforms, as well as features of a country's overall business environment, such as transparency in policymaking and the legal status of political contributions, influence these decisions. In contrast to Brazil, Indian business groups lobbying with money were significantly less likely to respond to any type of corruption question, and those using information were significantly more likely to respond to questions on all corruption types except executive corruption. Thus, in India, differences in lobbying tactics led to non-respondent samples that differed systematically from respondents, and these samples differed across corruption types.

Lobbying Tactics and Reported Corruption Levels:

Results from the outcome stage of the ordered probit models show that accounting for the effects of resource lobbying tactics on the decision to respond to corruption questions influences the results we obtain regarding the effects of resource tactics on reported corruption levels. Resource tactics now only influence bureaucratic corruption levels in India and have no influence on any reported corruption levels in Brazil.

The target of the resource lobbying is a significant predictor of corruption levels in both countries. Groups lobbying legislators in Brazil report lower levels of all types of corruption. Groups in India lobbying parties and the executive report higher levels of parliamentary and bureaucratic corruption. In contrast to these results, models which do not account for the respondent's choice to answer find that groups lobbying with technical information report lower parliamentary and executive corruption levels in Brazil and those lobbying with money report higher parliamentary corruption in India. Collectively, these results suggest that not accounting for respondents' decisions to respond to corruption questions in cross-country analyses of corruption survey data may be introducing significant and variable levels of bias into our results regarding corruption.

In conclusion, the evidence presented here suggests that surveys can continue to serve as an effective way of collecting data to answer specific questions on the corruption agenda when researchers are constrained in the choice of techniques they can use due to small sample sizes. In order to do so, however, it is essential to account for respondents' selection into non-responses, especially in cross-country contexts. This is because countries may

vary considerably in the extent to which their respondents systematically refuse to answer questions and in the reasons why they refuse. The results also highlight how important it is to disaggregate the concept of corruption into its component sub-types and then deal with the sensitivity issues specific to each sub-type. While the discussion here focuses on lobbying and corruption, these issues are pertinent to many situations where small elite populations must be surveyed and where respondents are being queried on their opinions regarding conceptually fuzzy aggregate terms that are composed of distinct sensitive individual components such as the quality of governance or rule of law in non-democracies.

References

- Ayyagari, Meghana, Thorsten Beck and Asli Demircuc-Kunt. 2007. "Small and medium enterprises across the globe." *Small Business Economics* 29(4):415–434.
- Bennedsen, Morten, Sven E Feldmann and David Dreyer Lassen. 2009. *Strong firms lobby, weak firms bribe: A survey-based analysis of the demand for influence and corruption*. Technical report EPRU Working Paper Series, Department of Economics, University of Copenhagen.
- Blair, Graeme, Kosuke Imai and Yang-Yang Zhou. 2015. *Design and analysis of randomized response technique*. Technical report.
- Campos, Nauro F. and Francesco Giovannoni. 2007. "Lobbying, corruption and political influence." *Public Choice* 131(1-2):1–21.
- Corstange, Daniel. 2009. "Sensitive questions, truthful answers? Modeling the list experiment with LISTIT." *Political Analysis* 17(1):45–63.
- dos Santos, Luiz Alberto and Paulo Mauricio Teixeira de Costa. 2014. Brazil: Lobby Regulation, Transparency and Democratic Governance. In *Lobbyists, Governments and Public Trust, Volume 3*. OECD pp. 109–123.
- EBRD. 2002. *The Business Environment and Enterprise Performance Survey 2002: A brief report on observations, experiences and methodology from the survey*. Technical report EBRD.
- Frye, Timothy. 2002. "Capture or exchange? Business lobbying in Russia." *Europe-Asia Studies* 54(7):1017–1036.
- Gingerich, Daniel W. 2010. "Understanding off-the-books politics: Conducting inference on the determinants of sensitive behavior with randomized response surveys." *Political Analysis* 18(3):349–380.
- Graf Lambsdorff, Johann. 2005. *International Handbook of the Economics of Corruption*. Northampton, MA: Edward Elgar.
- International, Transparency. 2007. *Global corruption report*

- 2007: *Corruption and judicial systems*. Transparency International.
- Kalniņš, Valts. 2011. "Transparency in lobbying: Comparative review of existing and emerging regulatory regimes." *Centre for Public Policy, PROVIDUS*.
- Krumpal, Ivar. 2013. "Determinants of social desirability bias in sensitive surveys: A literature review." *Quality and Quantity* 47(4):2025–2047.
- OECD. 2012. *Private Interests, Public Conduct: The Essence of Lobbying*. In *Lobbyists, Governments and Public Trust, Volume 2*. OECD.
- OECD. 2014. *Transparency in Lobbying Activities*. In *Lobbyists, Governments and Public Trust, Volume 3*. OECD.
- Olken, Benjamin A and Rohini Pande. 2011. *Corruption in developing countries*. Technical report National Bureau of Economic Research.
- Schneider, Ben Ross. 2013. "Institutions for effective business-government collaboration: Micro mechanisms and macro politics in Latin America." *IDB Working Paper No. IDB-WP-418*.
- Tan, Ming T., Guo-Liang Tian and Man-Lai Tang. 2009. "Sample surveys with sensitive questions: A nonrandomized response approach." *The American Statistician* 63(1):9–16.
- Thomas, Clive S. and Ronald J. Hrebendar. 2008. "Understanding interest groups, lobbying and lobbyists in developing democracies." *Journal of Public Affairs* 8(1-2):1–14.
- Treisman, Daniel. 2007. "What have we learned about the causes of corruption from ten years of cross-national empirical research?" *Annual Review of Political Science* 10:211–244.
- Winters, M. and R. Weitz-Shapiro. 2013. "Third-person versus second-person vignettes in survey experimental research on sensitive topics." Unpublished Manuscript.
- Yadav, Vineeta. 2008. "Business lobbies and policymaking in developing countries: The contrasting cases of India and China." *Journal of Public Affairs* 8(1-2):67–82.
- Yadav, Vineeta. 2011. *Political Parties, Business Groups, and Corruption in Developing Countries*. Oxford, UK: Oxford University Press.

II. Special Topic

The Ethics of Field Experiments in Comparative Politics

by Scott Desposato

University of Zurich and University of California, San Diego

By now most political scientists are probably aware of the controversy surrounding a recent field experiment conducted in Montana by Stanford and Dartmouth faculty (Michelson, 2014; Murphy, 2014; Stanford, 2014; Otani, 2014; Willis, 2014; Dennison, 2014). The scholars sent mailers to a reported 100,000 Montana voters with content regarding candidates for elected judgeships. Many have expressed surprise and indignation over the study, for its potential impact on an election outcome, for features of the mailer's design, and for possible IRB irregularities. But for this audience, the design should be familiar because it is one that has been applied in dozens of comparative politics experiments all over the world. Indeed, some of those comparative politics studies were illegal, lacked local IRB approval, and never informed, consented, or debriefed subjects¹.

Experiments are not new to political science, but most of our history with them involves fully informed and consenting undergraduate students playing simple and safe games in campus laboratories in the United States. Over the last 15 years, however, the number of experiments conducted by political scientists has increased exponentially.² These experiments can involve hundreds of thousands of subjects, sometimes uninformed and unconsenting, in very different contexts, with treatments that may fundamentally affect subjects' lives.

These changes hold great promise for advancing our research — identifying causal mechanisms, testing theories, and generalizing models across contexts. But they are generating ethical challenges, especially for comparative politics. Many of these challenges are new and unique to social science, and they do not have pre-existing solutions in other fields.

¹Ironically, the recent Montana study differs from most similar experiments in that the mailer noted that it was part of a study. Had the researchers not disclosed their affiliations, we probably never would have heard of the study.

²See empirical analysis of the number and type of experiments in Morton and Williams (2010), Druckman et al. (2006), and Desposato (2015).

To start a dialogue on these new challenges, I held a conference on ethics in comparative politics experiments at UCSD ([conference website](#); NSF# 1251510). The goal was to identify critical issues, explore opinion in the field, and propose practical strategies for moving forward. The proceedings are to be published later this year in the *Routledge Series in Experimental Political Science*.

In this essay, I will briefly review some of the issues we examined. We can organize the most common issues scholars are encountering into several categories: context, local review, deception and consent, and impact. The last three usually are manifest together in field experiments.

I. Context

The first set of issues are a byproduct of the globalization of experimental political science. Experiments that have no imaginable risks in the United States can cause problems elsewhere in the world. For example, Kim Dionne, Augustine Harawa, and Hastings Honde's chapter shows how standard compensation plans can cause conflict between subjects and non-subjects in contexts of extreme poverty and inequality. Becky Morton and Jonathan Rodgers discuss questions of religion, including how simple economic games may violate religious norms in some countries. Eddie Malesky examines the Common Rule exemption for public officials and asks how we should define "officials" in other contexts. For example, are inherited village leadership positions really public officials with no protections (Dionne and Honde, 2015; Morton and Rogers, 2015; Malesky, 2015)?

Based on the discussion at the conference, many of these are "easy" issues in that they would be relatively straightforward to fix through low-cost design changes, and most agreed that we should in fact adapt our designs to local context. For example, to avoid religious problems with treatments that look like gambling (which is forbidden in Islam), Morton and Rodgers discuss recasting a game to be about choosing traffic routes. To address tensions over compensation, Dionne, Harawa, and Honde propose strategies for explaining sampling and payoffs to subjects. Other contextual questions are not resolved. For example, if subjects frequently encounter violence in their everyday life, can we expose subjects and enumerators to levels of risk that might be common in their country, but which would be unacceptable back

home (Driscoll, 2015)?

II. Local Review

Questions of local review are more complex. Many scholars from the United States and Europe that are conducting experiments overseas do not have any local approval of their projects. Yet some countries have strict rules regarding special research visas, government permission, and local ethical review of protocols. In many cases, these requirements seem designed to prevent any research from every happening; in other cases they are opportunities to extract rents from scholars. At the conference, we focused on the cases of Brazil, China, Ecuador, Malawi, Mexico, and Vietnam, and found widely varying human subjects' rules and opinions on how we should proceed.

Some felt that experiments without foreign permission would be acceptable. Others, including scholars from the developing world, were indignant that first-world researchers would fly in on tourist visas, conduct experiments without permission, and slip out of the country, essentially running "under the radar" experiments.

Some issues were not controversial. Most agreed that compliance would be appropriate if the procedures were straightforward and reasonable. Similarly, most agreed that some form of local review — even if just informal review by local scholars or collaborators — is always appropriate. One benefit to local review is feedback on whether your intervention is contextually appropriate or has any risks you did not imagine.

But when considering cases where formal procedures are onerous or the government agencies incompetent or corrupt, there was clear disagreement on how we should proceed. Some felt that experiments without foreign permission would be acceptable. Others, including scholars from the developing world, were indignant that first-world researchers would fly in on tourist visas, conduct experiments without permission, and slip out of the country, essentially running "under the radar" experiments. One particularly grey area involves authoritarian regimes — where research rules may be designed to protect the regime rather than the subjects, and where

compliance would either mean not doing the research or only doing research that is beneficial to an oppressive regime.

Debate about local review also illustrates some of the differences between political science and other fields. Medical experiments which are often much higher risk would never proceed overseas without proof of local review; most medical funding agencies would require proof of local approval of the experiment. Political science experiments are generally extremely low risk to subjects, but may be perceived as threatening to regimes and could be even harder to get approved than medical studies. For the time being, the National Science Foundation and many of our university IRB's are not requiring any local approval of political science experiments.

One possible strategy is for scholars to test the waters of local review and if it seems like a bottomless pit, try working with local academics to develop a reasonable and effective review process. An example of this is the case of Brazil, where the University of São Paulo has established an Ethics Committee that will review foreign scholars' projects.

III. Deception, Consent, and Impact

The other most common issues involve deception, consent, and impact. The first two are enduring questions across bioethics, where the limits of consent and deception are continually being explored and debated (see, for example, Koenig (2014)). In political science there are also long-standing debates about whether deception is ever appropriate, even in laboratory environments. But more recently, and of greatest concern for comparative politics, these three issues are manifest together in field experiments.

Field experiments are interventions conducted "in the real world", where some feature of the natural environment is manipulated by researchers. Often subjects have no idea that they are participating in a study. These designs are especially promising because they allow us to observe the effect of the treatment in a real application, not in an artificial laboratory setting where many features of the experiment may not generalize to the real

world. At their best, field experiments provide more generalizable knowledge about how our treatments work and could aid in policy implementation and theory building.

My impression is that there is a significant divide in the field on the ethics of many field experiments. Proponents note that the treatments usually have only a trivially small risk of harm to individual subjects and are typical of everyday experiences. For example, it is difficult to imagine how a political mailer from a researcher poses a serious risk, especially when citizens receive dozens or even hundreds during campaigns. Other experiments might involve fake job applications to businesses or constituency mail to elected officials. But what is one extra resume or letter to an office that might deal with hundreds daily? And when field experiments involve consequential goods, like health care or clean water, most experiments do not violate the standard of care principle: the control group's situation is never made worse, they just do not receive the treatment. Further, random assignment of public goods is probably more fair than letting a corrupt political process perform the assignment.

Others are concerned with several features of field experiments. One position holds informed consent as paramount; thus any manipulation of subjects without their consent is simply unacceptable. A second source of tension is that while individual risks may be low, aggregate costs and risks may be large. For example, suppose we deceive a subject into just fifteen minutes of looking at a fake resume, as part of a study on employment. If we have 5,000 subjects, that means we have used 31 weeks of deception-induced full time labor.³ A third is that even when risks to subjects are trivially small, people often resent being manipulated without their permission for scholars' own purposes. Angering and upsetting subjects can be considered a form of harm to individuals. Field experiments might also be considered ethically questionable if they harm the broader discipline, for example, if a study results in all experiments being banned by an unfriendly government.⁴

Field experiments in the study of illegal activities, including corruption and vote buying, are particularly controversial, as treatments may involve breaking local

³(5,000 subjects × 15 minutes per subject) / (60 minutes per hour × 8 hours per day × 5 work days per week) = 31 weeks

⁴At the same time, too much concern with not making people or governments upset might stifle scholarship and violate norms of academic freedom. But the point illustrates that the choices we make about individual projects can affect our colleagues' research as well.

law. Scholars have solicited illegal activity by subjects, and committed minor crimes as part of field experiments. One perspective is that these are critically important topics and deserve our attention, and that deceptive — and sometimes illegal — designs are the only way to study such topics. For others, these experiments are simply wrong.

Especially disturbing to some are interventions in real elections or other political processes, where a political scientist's research has the potential to change vote shares or even an election outcome — potentially affecting millions of subjects and bystanders without their consent. As Zimmerman (2015) has pointed out, elections are always a zero sum game, so any impact of a treatment can cause harm. In addition, some are concerned that intervention in real politics may compromise our objectivity and integrity as scientists. On the other hand, one could argue that many campaign field experiments are normatively good — for example, if a political mailer provides factual information to voters or increases turnout, can we say that the intervention is a normative good, regardless of its impact on election results?

IV. Discussion

These are just a few of the perspectives on field experiments. My own opinion is that we can dodge or at least minimize many of the controversial cases with modest design changes. As I have argued elsewhere, the most critical issue is that of informed consent. Most controversies fade away when those affected by the study knowingly agree to participate. Some field experiments have designs that are amenable to informed consent, and whenever possible, we should request it. For cases where consent is impractical or ruins a design, there are alternative forms of consent that one might employ (Humphreys, N.d.). One can also minimize problems by treading lightly. For example, one could conduct a power analysis and minimize sample size. Why treat 100,000 subjects when 5,000 would be sufficient for hypothesis testing? We might also adopt a norm that subjects in field experiments be compensated for their time — either directly, or through payment to some public goods provider. Finally, a norm of post-study disclosure to all affected parties provides stronger incentives for researchers to anticipate problems and build solutions into designs.

There are many other issues I do not have space to discuss here, some of which are covered in the book. I'll conclude with a few general recommendations to consider before conducting your next experiment.

1. *Risk is not just about the expected physical harm to an individual subject.* Thinking ethically about experiments means thinking about individual subjects, aggregate costs, bystanders, and enumerators. Further, even if the study is safe and no one gets hurt, there can be costs to the discipline from making subjects and governments angry at us. As we have learned recently, people can get very upset about experiments that seem safe and innocuous.

2. *Be honest about the incentives of the profession.* Acknowledge and accept that we all have a conflict of interest when it comes to ethics and research with human subjects. Protecting subjects and others affected by interventions almost always makes research harder. At the same time, the discipline rewards the studies that are bigger and more exotic. Our own enthusiasm for the research and the professional success it may bring can nudge us toward ignoring real issues. The stakes are highest for those in the most precarious career situations: junior faculty and graduate students. This means that senior faculty should provide leadership on how to ethically conduct experiments, through mentoring, education, and example.

3. *Do not use ethics to reject the experimental revolution.* Experiments hold great promise for the discipline and should be embraced as one of many tools for studying comparative politics. Most comparative experiments are safe, legal, and have effectively zero risk. The problems we've encountered are rare, and in many cases, easily avoidable.

4. *Own it.* Do not outsource ethical questions to a third party. I have heard multiple scholars, when asked about the ethics of a project, respond, "Well, it got through IRB!" IRBs are imperfect institutions largely created to protect university access to federal funds, and IRB approval is not perfect absolution. Responsibility for ethics rests with the principal investigator.

5. *Be part of the dialogue.* There is increasing interest in ethics across the discipline, and there will almost certainly be several roundtables on ethics at Midwest and APSA. These may lead to new APSA Ethics Guidelines

for research. This dialogue will affect all of us, and you should be a part of it.

References

- Dennison, Mike. 2014. "Montana's top political cop wants action, apology from Dartmouth and Stanford." *The Billings Gazette*.
- Desposato, Scott. 2015. Introduction. In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. Scott Desposato. Routledge Studies in Experimental Political Science London, UK: Routledge.
- Dionne, Kim Yi, Augustine Harawa and Hastings Honde. 2015. The Ethics of Exclusion when Experimenting in Impoverished Settings. In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. Scott Desposato. Routledge Studies in Experimental Political Science London, UK: Routledge.
- Driscoll, Jesse. 2015. Prison States and Games of Chicken. In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. Scott Desposato. Routledge Studies in Experimental Political Science London, UK: Routledge.
- Druckman, James N., Donald R. Green, James H. Kuklinski and Arthur Lupia. 2006. "The growth and development of experimental research in political science." *American Political Science Review* 100(4):627–635.
- Humphreys, Macartan. N.d. "Manipulation and consent." Unpublished Manuscript.
- Koenig, Barbara. 2014. "Have we asked too much of Consent?" *The Hastings Center Report* 44(4):33–34.
- Malesky, Edward. 2015. Who is a Public Official? In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. Scott Desposato. Routledge Studies in Experimental Political Science London, UK: Routledge.
- Michelson, Melissa R. 2014. "Messing with Montana: Get-out-the-vote experiment raises ethics questions." *Western Political Science Association*.
- Morton, Rebecca B. and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality From Nature to the Lab*. New York: Cambridge.
- Morton, Rebecca and Jonathan Rogers. 2015. Religion, Experiments, and Ethical Concerns. In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. Scott Desposato. Routledge Studies in Experimental Political Science London, UK: Routledge.
- Murphy, Katy. 2014. "Stanford, Dartmouth sidestep legal action against researchers' election mailers." *Mercury News*.
- Otani, Akane. 2014. "Stanford and Dartmouth's bungled voter experiment spans three states." *Businessweek*.
- Stanford. 2014. "Stanford responds to voter mailer."
- Willis, Derek. 2014. "Professors' research project stirs political outrage in Montana." *New York Times*.
- Zimmerman, Brigitte. 2015. Ethical Concerns Surrounding Research Interventions in Democratic Processes. In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. Scott Desposato. Routledge Studies in Experimental Political Science London, UK: Routledge.

III. Dataset

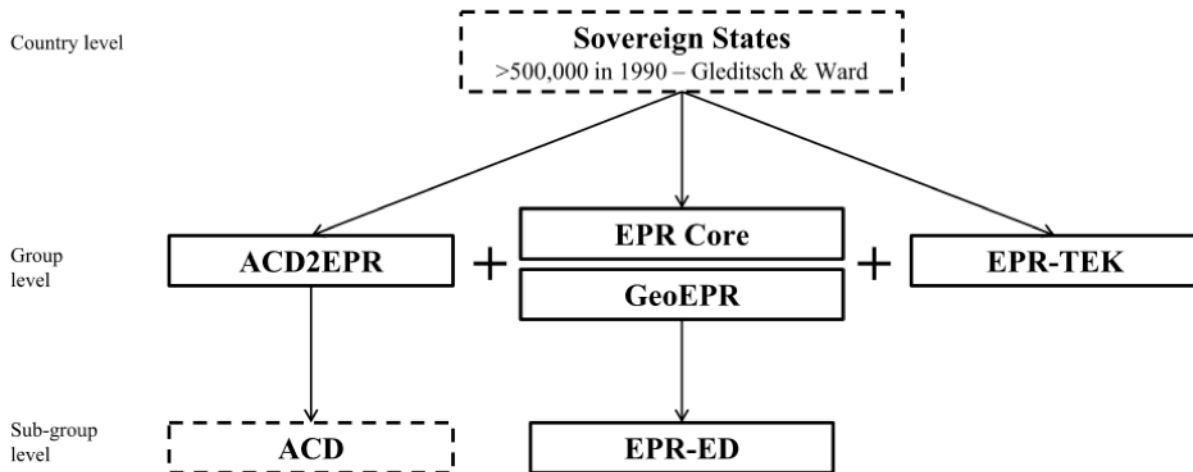
Integrating Data on Ethnicity, Geography, and Conflict: The Family of Ethnic Power Relations Datasets 2014

by Nils-Christian Bormann, Manuel Vogt,
& Lars-Erik Cederman

ETH Zürich

Since the early 1990s, empirical research on the consequences of ethnic diversity has been the subject of much academic debate among economists (Alesina, Baqir and Easterly, 1997), comparativists (Ordeshook and Shvetsova, 1994), and IR scholars (Denny and Walter, 2014). Moving from highly aggregated measures of ethnic diversity such as the ethno-linguistic fractionalization (ELF) index, political scientists have turned to more detailed group-level datasets that cover linguistic, religious, and racial ethnic groups (Gurr, 1995). This essay introduces the latest release of the *Ethnic Power Relations* (EPR) data, version 2014, which includes information on over 800 unique ethnic groups in 165 countries between 1946 and 2013. Consisting of five constituent datasets, the EPR Family 2014 provides time-varying information on ethnic groups' access to executive power, relative group size, their geographic settlement patterns and involvement in intrastate wars, transnational linkages, and sub-group dimensions such as language and religion. While all variables are coded on the group-level, we also provide country-level versions of the data that can be used to study outcomes such as democratization. In addition, the 2014 version features several novelties, such as a new variable of territorial autonomy, and for the first time, comprehensive documentation in the form of the *EPR Atlas*. In the following, we explain the

Figure 1: The structure of the EPR Dataset Family 2014



Note: Bold boxes designate EPR Family datasets; dashed boxes identify datasets to which EPR is linked.

outstanding features of the EPR Family 2014 by tracing its evolution from earlier versions and by comparing it to alternative data sources on ethnicity.

Providing a list of politically relevant ethnic groups, the *EPR Core* dataset is at the center of the EPR data universe (see Figure 1). EPR defines ethnicity as a subjectively experienced sense of commonality based on a belief in common ancestry and shared culture. An ethnic group is considered politically relevant if at least one political organization has claimed to represent its interests at the national level or if its members are subjected to state-led political discrimination (Cederman, Wimmer and Min, 2010).¹ By broadening the notion of “political relevance” to include majority and dominant groups, the EPR group list is more inclusive than the Minorities at Risk (MAR) dataset by Gurr (1995), and it accounts for changes in the set of politically relevant groups when new actors make political claims along new or altered ethnic identity lines.² For example, EPR considers Whites, Blacks, and Native Americans to be the only politically relevant groups in the United States until 1965. Afterwards, Hispanics, Asian Americans, and Arab Americans become politically relevant as several civil society organizations that make claims on behalf of these groups are founded in the wake of the

anti-discrimination legislation of the Johnson administration.

The central feature of all EPR versions is the ordinal scale of ethnic groups’ access to executive power that distinguishes three major categories: (1) one group controls the executive, (2) representatives of an ethnic group share government power, and (3) members of an ethnic group are excluded from the highest ruling body of a state (e.g., a democratic cabinet, a military junta, or a royal court). Each of these three main categories includes sub-classifications of an ethnic group’s power status (see Table 1 and textbox on the next page).

Access to sub-national governments such as federal states does not count towards this power-scale, which focuses solely on the access of group representatives to the central government. This classification of groups’ power status has undergone an important change in the 2014 version, as previous EPR versions code regional autonomy within the ordinal power scale for groups excluded from the central government, such as the Naga in India, but not for groups included in the government. In contrast, the EPR Core 2014 dataset removes the regional autonomy category from the relative power access and provides a new variable that assesses groups’ access to

¹ Based on collaboration with researchers at UCLA, the original version of EPR v. 1.0 covers the period 1946-2005. Collected by the International Conflict Research Group at ETH, the direct predecessor to the current release is the dataset EPR-ETH 2.0, which extends coverage until 2009. A separate version assembled by researchers from UCLA (labeled EPR version 3) also builds directly on the first release of EPR and extends coverage to 2010. While this version introduces information about the “cultural contents” of ethnic groups, it does not offer any spatial extension nor any compatibility with EPR Family 2014.

² The new AMAR dataset by Birnir et al. (2014) lists more than 1,200 “socially relevant” ethnic groups and provides various sub-groups. However, additional variables are only coded for a small random sample of the larger universe of groups.

Table 1: Power Status, 2013

Power status	N	Percent	EPR Power Categories
<i>One-group rule</i>			
Monopoly	24	3.4%	<i>Monopoly</i> : Only representatives from one ethnic group are included in the executive (e.g., Sunni Arabs in Saudi Arabia).
Dominant	42	6.0%	<i>Dominant</i> : The government consists predominantly of members from one ethnic group but some token members of other groups are included without any real influence (e.g., Sinhalese in Sri Lanka).
<i>Power-sharing</i>			
Senior partner	83	11.9%	<i>Senior Partner</i> : A group shares power but its representatives have more influence than its partners (e.g., Ovambu in Namibia).
Junior partner	164	23.5%	<i>Junior Partner</i> : A group shares power but its representatives have less power than its partners (e.g., Kavango in Namibia).
<i>Excluded</i>			
Powerless	312	44.7%	<i>Powerless</i> : A group lacks representation in the executive (e.g., Bosniaks in Croatia).
Discriminated	67	9.6%	<i>Discriminated</i> : The state actively and intentionally targets members of an ethnic group in the political realm (e.g., Sahrawis in Morocco).
Self-exclusion	6	0.9%	<i>Self-exclusion</i> : This category describes ethnic groups that control some territory of the state and have declared this region to be independent from the central government (e.g., Armenians in Nagorno-Karabakh).
Total	698	100%	

executive power on the regional level *independently* of their national power status. Representatives of the Punjabi-Sikhs were included in most of India's ruling coalitions, and they have politically dominated the Punjab since the redrawing of its boundaries in 1966. In earlier versions of EPR, the Punjabi would simply have been coded as being included in the central government, but the 2014 release now codes the Punjabis as both included in a power-sharing government and controlling their own state. This change enables researchers to investigate governmental and territorial power-sharing for a large number of ethnic groups in pre- and post-war settings (Cederman et al., 2015). The de facto coding of access to executive power on the national and sub-national level also sets the EPR data apart from data sources that provide information on the de jure provisions of peace agreements in post-conflict states (Hartzell and Hoddie, 2007).

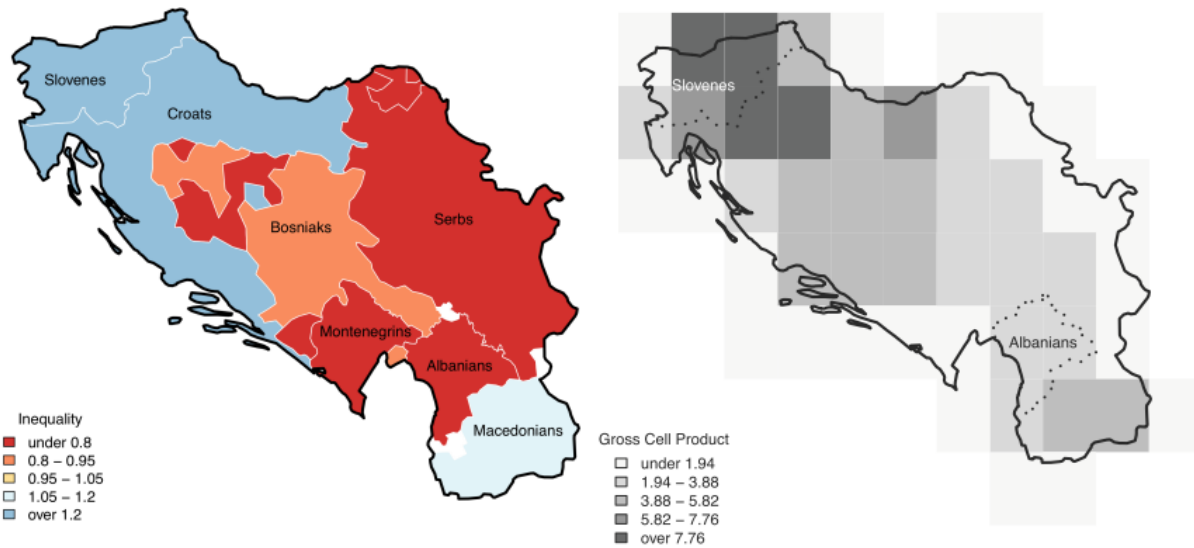
A unique feature of the EPR Family is the geo-coding of ethnic groups' settlement patterns in the GeoEPR dataset, which was introduced by Wucherpfennig et al. (2011). Providing shape files for the vast majority of ethnic groups that are territorially concentrated, GeoEPR provides digital time-variant maps of ethnic groups' home territories. Beyond extending the temporal scope to 2013, GeoEPR 2014 introduces the "statewide" category that encompasses groups with a presence in virtually every part of a country. These groups previously

fell into the "dispersed" category that now only comprises the subset of minority groups that live in different parts of the country. Beyond serving as useful graphical illustrations, the GeoEPR maps offer the basis for measures of an ethnic group's territorial concentration and open up new avenues for data collection on the ethnic group-level. Cederman, Weidmann and Gleditsch (2011), for example, use the GeoEPR maps to derive estimates of group inequality by overlaying it with geo-referenced income data (Nordhaus, 2006). Figure 2 on the next page shows an example of the GeoEPR maps. It shows the ethnic geography in Yugoslavia before its dissolution in 1990. GeoEPR also traces changes in groups' settlement patterns due to changes in country borders as in Yugoslavia or mass-scale migration within the same country as occurred after the Russian invasion of Georgia in 2008.

The conflict coding of the EPR family establishes a link to the UCDP/PRIO Armed Conflict (ACD) Database (Gleditsch et al., 2002; Themnér and Wallenstein, 2014).³ The ACD2EPR data classify rebel groups as ethnic if they recruit fighters from a particular ethnic group and make public claims on behalf of that group. Coding individual rebel organizations rather than entire civil wars, the ACD2EPR data offer a more nuanced classification of conflicts, which enables conflict researchers to compare multiple rebel groups that fight on behalf of the same ethnic group, such as the Karen National Union

³For UCDP, see the Uppsala Conflict Data Program (<http://www.pcr.uu.se/research/UCDP/>). PRIO stands for Peace Research Institute Oslo (<http://www.prio.org/>)

Figure 2: EPR groups in Yugoslavia in 1990 with inequality estimates (left) and G-ECON income data (right)



and the Karen National United Party in Burma, with rebel groups that unite several ethnic groups against the government, such as the African National Congress during South Africa’s Apartheid regime that united Asians, Blacks, and Coloreds (Wucherpfennig et al., 2012).

Going beyond the confines of the nation-state, the newly integrated Transborder Ethnic Kin (EPR-TEK) dataset identifies all EPR groups with settlements in at least two countries through nominal matching. Thus, groups in different countries are coded as transborder kin if they share the same ethnographic name, including synonyms. The Kurds who live in Turkey, Iran, Iraq, and Syria exemplify such a transnational ethnic group, and thereby represent about half of all groups in the EPR dataset that feature a cross-border link. In contrast to the MAR dataset, which only codes whether groups are linked to other groups in neighboring states, the combination of EPR-Core and EPR-TEK data also provides information on the political status of transborder ethnic kin groups. Researchers can therefore differentiate if groups in one country can obtain support from a transborder group that holds government power or from excluded groups, and therefore trace more specific mechanisms (Cederman et al., 2013).

The second novel addition to the EPR Dataset Family is the Ethnic Dimensions (EPR-ED) data, which

identify the linguistic, religious, and racial segments of all EPR groups.⁴ This is the first dataset that codes both multiple cleavage dimensions and several segments within a cleavage dimension for ethnic groups. Hence, this approach refrains from representing ethnic groups as always being exclusively linguistic, religious, or racial. In contrast, other researchers often link ethnic groups to one primary category such as “ethno-linguistic” or “ethno-religious” (Sorens, 2011), thus treating these groups as internally homogenous and one-dimensional. Drawing on the Ethnologue database on languages (Lewis, 2009), the Joshua Project’s coding of religious identities (2011), and many secondary sources, the EPR-ED data enable researchers to go beyond the ethnic group level by constructing estimates of within-group fractionalization and dyadic between-group differences along any combination of linguistic, religious, or racial cleavages.

The individual components of the EPR dataset family can be downloaded from the GROW^{UP} web portal at <http://growup.ethz.ch>. In addition to featuring a Public Front-End (PFE) for visual inspection of ethnic groups in time and space, the system provides a user-friendly Research Front-End (RFE), which allows users to assemble and download customized panel datasets composed of EPR-related variables. The data offered via the RFE are pre-aggregated to the level of group-years and

⁴The term “race” here refers to ethnic groups’ origins from particular world regions, such as Europe, Sub-Saharan Africa, Oceania, etc. These regional origins — expressed at the individual level through certain phenotypical markers (particularly skin color) — have become relevant as social categories in the context of the European colonization of the world and the related process of racial classification.

country-years. Formatted in order to facilitate statistical analysis, the selection of variables includes conflict onset and incidence dummies, as well as various peace-years variables and other temporally defined conflict indicators.⁵

In summary, the EPR Dataset Family 2014 provides scholars with a wide range of instruments to improve their understanding of politicized ethnicity in general, and ethnic conflict in particular. Thanks to its extensive scope, the integrated datasets facilitate the study of a variety of questions. However, much work still remains to be done. In this regard, we want to highlight three promising directions of future data collection. First, it is important to identify the actual political actors that are at the roots of the collective action undertaken by ethnic groups. For this purpose, we plan to launch a new data collection project, the EPR-Organizations Dataset (<http://www.r4d.epr.ethz.ch>), which will offer information on ethnically based political organizations in all countries of the world. The goal is to break up the monolithic concept of ethnic groups by taking into account the diverse agendas and claims of different organizational representatives, and to study ethnic mobilization processes. Second, broadening the spectrum of political violence covered by EPR beyond civil war, upcoming extensions of the ACD2EPR dataset will link EPR groups to all UCDP actors, including those involved in one-sided violence, non-state conflict, and inter-state war. Finally, going beyond the nominal coding of EPR-TEK, future versions of the dataset will adhere to a claim-based coding that allows for changes in transnational ethnic identifications.

References

- Alesina, Alberto, Reza Baqir and William Easterly. 1997. "Public goods and ethnic divisions." *Quarterly Journal of Economics* 114(4):1243–1284.
- Birnie, Jóhanna K., Jonathan Wilkenfeld, James D. Fearon, David D. Laitin, Ted Robert Gurr, Dawn Brancati, Stephen M. Saideman, Amy Pate and Agatha S. Hultquist. 2014. "Socially relevant ethnic groups, ethnic structure, and AMAR." *Journal of Peace Research* 52(1):110–115.
- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why do ethnic groups rebel? New data and analysis." *World Politics* 62(1):87–119.
- Cederman, Lars-Erik, Kristian Skrede Gleditsch, Idean Salehyan and Julian Wucherpfennig. 2013. "Transborder ethnic kin and civil war." *International Organization* 67(2):389–410.
- Cederman, Lars-Erik, Nils B. Weidmann and Kristian Skrede Gleditsch. 2011. "Horizontal inequalities and ethnonationalist civil war: A global comparison." *American Political Science Review* 105(3):478–495.
- Cederman, Lars-Erik, Simon Hug, Andreas Schädel and Julian Wucherpfennig. 2015. "Territorial autonomy in the shadow of conflict: Too little, too late?" *American Political Science Review*.
- Denny, Elaine K. and Barbara F. Walter. 2014. "Ethnicity and civil war." *Journal of Peace Research* 51(2):199–212.
- Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed conflict 1946–2001: A new dataset." *Journal of Peace Research* 39(5):615–637.
- Gurr, Ted Robert. 1995. *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. United States Institute of Peace.
- Hartzell, Caroline A. and Matthew Hoddie. 2007. *Crafting Peace: Power-Sharing Institutions and the Negotiated Settlement of Civil Wars*. University Park, PA: Penn State Press.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the world*. Vol. 9 SIL International Dallas, TX.
- Nordhaus, William D. 2006. "Geography and macroeconomics: New data and new findings." *Proceedings of the National Academy of Sciences of the United States of America* 103(10):3510–3517.
- Ordeshook, Peter C. and Olga V. Shvetsova. 1994. "Ethnic heterogeneity, district magnitude, and the number of parties." *American Journal of Political Science* 38(1):100–123.
- Sorens, Jason. 2011. "Mineral production, territory, and ethnic rebellion: The role of rebel constituencies." *Journal of Peace Research* 48(5):571–585.
- Themner, Lotta and Peter Wallensteen. 2014. "Armed conflicts, 1946–2013." *Journal of Peace Research* 51(4):541–554.
- Wucherpfennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman and Andreas Wimmer. 2011. "Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset." *Conflict Management and Peace Science* 28(5):423–437.
- Wucherpfennig, Julian, Nils W. Metternich, Lars-Erik Cederman and Kristian Skrede Gleditsch. 2012. "Ethnicity, the state, and the duration of civil war." *World Politics* 64(1):79–115.

⁵The EPR 2014 data are available in their raw format here: <http://www.icr.ethz.ch/data/index>. The new Research Front End that allows customized downloads will go online in March 2015.

Editors and Contributors

Matt Golder



Matt Golder is an Associate Professor in the [Department of Political Science](#) at [The Pennsylvania State University](#). He received his Ph.D. in 2004 from [New York University](#). His research looks at how political institutions affect democratic representation, with a particular focus on electoral rules and party systems. In addition to articles in leading journals, such as the *American Journal of Political Science*, the *British Journal of Political Science*, the *Journal of Politics*, and *Political Analysis*, he has also published a textbook on comparative politics, *Principles of Comparative Politics*. He is currently working on two research projects. The first looks at negative campaigning in a multi-party context, while the second examines the relationship between religious participation, economic development, and political attitudes. In addition to serving as chair of APSA's section on [Representation and Electoral Systems](#) (2011-2013), he is also a member of the executive board for the [Making Electoral Democracy Work](#) project led by André Blais at the University of Montreal and the advisory board for the [Electoral Integrity Project](#) led by Pippa Norris at Harvard University. More information can be found at his [website](#) and on his [Google scholar profile](#).

Sona N. Golder



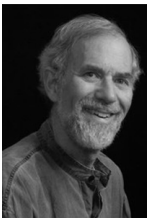
Sona Golder is an Associate Professor in the [Department of Political Science](#) at [The Pennsylvania State University](#). She received her Ph.D. in 2004 from [New York University](#). She studies political institutions, with a particular interest in coalition formation. Her early research and first book, *The Logic of Pre-Electoral Coalition Formation*, examines the determinants and consequences of electoral alliances in parliamentary democracies. She is currently working on a book project that draws together her more recent work on the post-election government formation process in Europe. Her research has been published in many of the discipline's leading journals, including the *American Journal of Political Science*, the *British Journal of Political Science*, the *Journal of Politics*, and *Political Analysis*. She has also published a textbook on comparative politics, *Principles of Comparative Politics*. She is also involved in the women in methods group — she was the organizer and host for the [4th Annual Visions in Methodology \(VIM\) Conference](#), she serves as a VIM mentor for female graduate students and junior faculty, and she is a member of the diversity committee for APSA's Political Methodology Section. More information can be found at her [website](#) and on her [Google scholar profile](#).

Jesse Driscoll



Jesse Driscoll is an Assistant Professor at the [School of International Relations and Pacific Studies](#) at the [University of California, San Diego](#). He received his Ph.D. in 2009 from [Stanford University](#) and was a Post-Doctoral Academy Fellow at the [Harvard Academy for International and Area Studies](#). He has conducted extensive fieldwork in the Former Soviet Union, and is currently overseeing a number of research projects looking at how individuals living in conditions of acute insecurity respond to survey questions about their behavior. His work has appeared in the *Journal of Conflict Resolution, Research and Politics*, and the *Journal of Survey Statistics and Methodology*. His book *Warlords and Coalition Politics in Post-Soviet States* is forthcoming from Cambridge University Press. His research has been funded by the National Science Foundation.

David Laitin



David Laitin is the James T. Watkins IV and Elise V. Watkins Professor in the [Department of Political Science](#) at [Stanford University](#). He conducts research on political culture, ethnic conflict, and civil war, and his field expertise spans Somalia, Nigeria, Catalonia, Estonia and France. His research has been published in all of the discipline's leading journals. He has also published books with Cambridge University Press and Oxford University Press. He has been a recipient of fellowships from the [Howard Foundation](#), the [Rockefeller Foundation](#), the [Guggenheim Foundation](#), and the [Russell Sage Foundation](#). He is an elected member of the [American Academy of Arts and Sciences](#) and the [National Academy of Sciences](#). More information can be found at his [website](#).

Graeme Blair



Graeme Blair is a pre-doctoral fellow at [Columbia University](#) through [Experiments in Governance and Politics](#) and the [Earth Institute](#). He will receive his Ph.D. in 2015 from [Princeton University](#) and will be an Assistant Professor in the [Department of Political Science](#) at [UCLA](#) from 2016. His research examines both why civilians support armed groups and why groups living near valuable assets like oil fields are often able to force the state to make policy concessions and fiscal transfers. In addition, he has developed new methods for sensitive survey questions. His work has appeared in the *American Political Science Review*, the *American Journal of Political Science*, and *Political Analysis*. He has also won the [2013 Pi Sigma Alpha Award](#) for best paper at Midwest Political Science Association. More information can be found at his [website](#) and on his [Google scholar profile](#).

Daniel Gingerich



Daniel Gingerich is an Associate Professor in the [Department of Politics](#) at the [University of Virginia](#). He also directs the [Quantitative Collaborative](#), a University-wide institute designed to disseminate new advances in the quantitative social sciences and foster greater interdisciplinary collaboration. He received his Ph.D. in 2007 from [Harvard University](#). His research focuses on understanding the causes and consequences of corruption and clientelism in Latin America and the Caribbean, as well as on developing new methodologies to study these phenomena. He has published articles in journals such as the *British Journal of Political Science*, *Political Analysis*, and the *Quarterly Journal of Political Science*. His book, *Political Institutions and Party-Directed Corruption in South America: Stealing for the Team*, was published by Cambridge University Press in 2013. His research has been funded by the [National Science Foundation](#). More information can be found at his [website](#).

Elizabeth C. Carlson



Elizabeth C. Carlson is an Assistant Professor in the [Department of Political Science](#) and the [Program on African Studies](#) at [The Pennsylvania State University](#). She received her Ph.D. in 2011 from the [University of California, Los Angeles](#) and has held pre- and post-doctoral fellowships at [Stanford](#) and [Yale](#). She uses survey and experimental methods to study political behavior and citizen preferences in Africa's new democracies, as well as how citizen behavior shapes government performance and accountability. Her work is forthcoming in *World Politics* and has been funded by the [National Science Foundation](#). She is a member of [Experiments in Governance and Politics](#). More information can be found at her [website](#) and on her [Google scholar profile](#).

Vineeta Yadav



Vineeta Yadav is an Assistant Professor in the [Department of Political Science](#) at [The Pennsylvania State University](#). She received her Ph.D. in 2007 from [Yale University](#). Her research focuses on the effects of institutions on economic development, with a particular emphasis on how institutions influence lobbying, corruption, and judicial empowerment. In addition to several articles, she has also published two books: *Political Parties, Business Groups, and Corruption in Developing Countries* and *Democracy, Electoral Systems and Judicial Empowerment in Developing Countries*. A third book, *Corruption in Dictatorships*, is forthcoming with Cambridge University Press. She is currently working on two new research projects. The first examines the origins of commercial courts and their effect on economic development and civic rights. The second looks at how the evolution of political learning and support influences institution building and democratic consolidation. More information can be found at her [website](#) and on her [Google scholar profile](#).

Scott Desposato



Scott Desposato is an Associate Professor in the [Department of Political Science](#) at the [University of California, San Diego](#). Before coming to San Diego, he was an Assistant Professor at the [University of Arizona](#), a [Harvard Academy Scholar](#), and a [Research Fellow](#) at [Princeton University](#). He received his Ph.D. in 2001 from the [University of California, Los Angeles](#). His general research interests include democratic institutions, campaigning, mass behavior, and political methodology. Specific projects have examined redistricting in the United States, electoral rules and federalism in Brazil, party-switching by politicians, and statistical methods for studying legislatures. His work has appeared in the discipline's leading journals such as the *American Journal of Political Science*, the *American Political Science Review*, and the *Journal of Politics*. In 2013, he organized the [NSF-funded Ethics in Comparative Politics Experiments Conference](#). More information can be found at his [website](#) and on his [Google scholar profile](#).

Nils-Christian Bormann



Nils-Christian Bormann is a post-doctoral researcher in the [International Conflict Research \(ICR\)](#) group at [ETH Zurich](#). Prior to taking his position with the ICR, he was a visiting researcher at the [University of Pittsburgh's Department of Political Science](#), where he was funded by a [DocMobility Grant](#) from the [Swiss National Science Foundation](#). He received his Ph.D. in 2014 from [ETH Zurich](#). His research interests include ethnic coalitions and power-sharing, civil wars, ethnic conflict, democratization, electoral rules, and spatial methods. His work has appeared in *Electoral Studies*. More information can be found at his [website](#).

Manuel Vogt



Manuel Vogt is a post-doctoral researcher in the [International Conflict Research \(ICR\)](#) group at [ETH Zurich](#). He received his Ph.D. in 2013 from [ETH Zurich](#). For the past five years, he has managed the [EPR-ETH dataset](#) and coordinated the two updates of the dataset. His research interests include ethnic conflict, (post-conflict) democratization, electoral institutions and party systems, and Latin American and African politics. His work has appeared in *Latin American Politics and Society* and in several edited volumes. More information can be found at his [website](#).

Lars-Erik Cederman



Lars-Erik Cederman is a Professor of International Conflict Research in the [Center for Comparative and International Studies](#) at [ETH Zurich](#). Before coming to Zurich, he was an Associate Professor at [Harvard University](#), an Assistant Professor at the [University of California, Los Angeles](#), and a University Lecturer and Fellow at [Somerville College, Oxford University](#). He received his Ph.D. in 1994 from the [University of Michigan](#). His research interests include computational modeling, international relations theory, nationalism, integration and disintegration processes, and historical sociology. He has published articles in journals such as the *American Political Science Review*, *International Organization*, and *Proceedings of the National Academy of Sciences*. More information can be found at his [website](#) and on his [Google scholar profile](#).

Announcements

- REPAL 2015 is a new conference on Latin American political economy. The conference is 7-8 July in Montevideo. For more information see <http://redeconomiapoliticaamlat.com>.
- Please see the call for papers announcement for a special issue dedicated to the topic of 'Political Activism in Europe: Numbers and Trends' in the online journal PACO - <http://siba-ese.unisalento.it/index.php/paco/announcement/view/36>.
- The IPSA Committee on Concepts and Methods seeks an editor, or team of editors, for a two-year term. This is an opportunity to make a significant contribution to one of the leading discussion forums for conceptual and methodological issues in Political Science. For more information please visit <http://comparativenewsletter.com/node/16>.

About the Section

The [Organized Section in Comparative Politics](#) is the largest organized section in the American Political Science Association (APSA) with over 1,300 members. The purpose of the Section is to promote the comparative, especially cross-national, study of politics and to integrate the work of comparativists, area studies specialists, and those interested in American politics. The Section organizes panels for APSA's annual meetings; awards annual prizes for best paper, best article, best book, and best dataset; and oversees and helps finance the publication of the Newsletter. For more information, please visit the Section's website.

About the Newsletter

The goal of the Comparative Politics Newsletter is to engender a sense of community among comparative politics scholars around the world. To this end, the Newsletter publishes symposia on various substantive and methodological issues, highlights new datasets of broad appeal, prints short comments from readers in response to materials in the previous issue, and generally informs the community about field-specific developments. Recent symposia have looked at the varieties of authoritarianism, the global economic crisis, and field experiments. Upcoming issues will focus on graduate training in comparative politics and the politics of space. It is published twice a year — once during the Spring and once during the Fall. The Newsletter is currently edited by [Matt Golder](#) and [Sona N. Golder](#) at The Pennsylvania State University.

How to Subscribe

Subscription to the APSA-CP Newsletter is a benefit to members of the Organized Section in Comparative Politics of the American Political Science Association. To join the section, check the appropriate box when joining APSA or renewing your Association membership. You may join the APSA online at <http://www.apsanet.org/content.asp?contentid=4>.

Copyright 2015 American Political Science Association.
Address: 1527 New Hampshire Ave, NW Washington, DC 20036-1206
Phone: (202) 483-2512 | Fax: (202) 483-2657 | Email: apsa@apsanet.org