ESSAYS IN THE STUDY OF LARGE-SCALE SURVEY DATA

by

Gregory Daniel Eady

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Political Science
University of Toronto

© Copyright 2018 by Gregory Daniel Eady
Abstract

Essays in the Study of Large-scale Survey Data

Gregory Daniel Eady
Doctor of Philosophy
Graduate Department of Political Science
University of Toronto
2018

For the study of politics, the recent growth in the size, affordability, and availability of data and computational resources has been staggering. Large and novel data sources are now beginning to figure prominently in the work of some of the most impressive current research. This dissertation builds on these developments by demonstrating how the growing availability of large-scale survey data can open up new opportunities to answer questions that have, until now, remained difficult to address. To this end, this dissertation (1) proposes new methods for the analysis of large-scale survey data, and (2) seeks to answer theoretically important questions that can benefit from these methods and/or data.

In the first paper, I seek to answer why some survey respondents answer truthfully to sensitive survey questions, while others do not. I propose a solution to this problem by developing a multivariate regression-based technique to model whether survey respondents provide one response to a sensitive item in a list experiment, but answer otherwise to a direct question. The method is applied to a large-scale list experiment to investigate misreporting about prejudice concerning women in politics. In the second paper, I examine how researchers can measure the ideology of ordinary citizens and political actors on the same scale. I propose a Bayesian Aldrich-McKelvey scaling method that both incorporates survey respondents’ ideological self-placement into the scaling model and accounts for the fact that typical ideological placement scales are ordinal. The model is applied to large-scale data from recent Cooperative Congressional Election Studies to investigate whether Bernie Sanders was ideologically out of step with the national electorate during recent Presidential elections. In the final paper, I examine the effects of major terrorist attacks on public attitudes toward refugees. Using data from a large-scale natural
experiment, the paper shows that the 2015 Islamic State terrorist attacks in Paris increased (1) anxiety over refugee resettlement; (2) perceptions of refugees as a security and cultural threat; and (3) opposition to resettlement. The findings are highly relevant to our understanding of public reactions to major terrorist attacks, and the responses of political entrepreneurs in their aftermath.
# Contents

List of Tables vii

List of Figures viii

1 Introduction 1

1.1 Introduction ................................................. 1
1.2 The Data and Computational Social Science Revolution ................. 2
1.3 Organization & Overview ...................................... 5

2 The Statistical Analysis of Misreporting 10

2.1 Introduction .................................................. 10
2.2 Proposed Methodology ...................................... 11
  2.2.1 The list experiment setup ............................... 12
  2.2.2 The direct question and misreporting .................. 13
  2.2.3 Respondent types ........................................ 14
  2.2.4 Maximum likelihood estimator .......................... 15
  2.2.5 Model optimization ..................................... 16
2.3 Monte Carlo Simulations ..................................... 18
  2.3.1 Simulation study 1 ....................................... 19
  2.3.2 Simulation study 2 ....................................... 20
  2.3.3 Simulation study 3 ....................................... 20
2.4 Empirical Application ....................................... 23
  2.4.1 Does the left misreport their prejudices more than those on the right? 23
  2.4.2 Data and survey design ................................. 24
  2.4.3 Results .................................................. 26
2.5 Conclusion .................................................. 29
3 How Should We Measure Citizen Ideology? 31
  3.1 Introduction ................................................................. 31
  3.2 AM Scaling and Differential Item Functioning ....................... 33
  3.3 Statistical model ............................................................ 35
  3.4 Empirical Comparison and Application ................................ 38
    3.4.1 Political ideology during the mid-term and Presidential elections .... 40
    3.4.2 Was Bernie Sanders ideologically out of step during the 2016 Presidential campaign? ........................................... 43
  3.5 Conclusion ................................................................. 44

4 International Terrorism & Public Attitudes toward Refugees 47
  4.1 Introduction ................................................................. 47
  4.2 The Syrian Refugee Crisis and the Paris Attacks .................... 49
  4.3 Terrorism and Attitudes toward Refugees ............................. 50
    4.3.1 Terrorism and emotion ............................................... 50
    4.3.2 Terrorism and threat perception ................................... 52
  4.4 Terrorism and Political Mobilization ................................. 53
  4.5 Effect Duration ............................................................ 54
  4.6 Hypotheses ................................................................. 55
  4.7 Research Design ......................................................... 55
    4.7.1 The Paris attacks as natural experiment ......................... 56
    4.7.2 Survey design ....................................................... 57
  4.8 Results ........................................................................... 58
    4.8.1 Emotions and attitudes toward refugees and resettlement .......... 58
    4.8.2 Evidence of causal mechanisms ..................................... 62
    4.8.3 Political mobilization ............................................... 66
  4.9 Discussion and Conclusion .............................................. 70

5 Conclusion 73
  5.1 Introduction ................................................................. 73
  5.2 Motivation and contributions ............................................ 73
    5.2.1 Bigger data, smaller groups ......................................... 74
    5.2.2 Bigger data, bigger models ......................................... 76
  5.3 Survey research and the growth of ‘Big Data’ ....................... 77

A The Statistical Analysis of Misreporting: Appendix 80
  A.1 Standard list experiment model ......................................... 80
  A.2 Proposed method for alternative case ................................ 81
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Respondent types defined by the response ((Z^*_i, D_i))</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Observed-data likelihood</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>Weights, (w_i^{(3)}), as calculated in the E-step</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Multivariate list experiment results</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>Predictive performance of AM Scaling Models</td>
<td>40</td>
</tr>
<tr>
<td>3.2</td>
<td>Proportion of voters and co-partisans who are more liberal than candidates and parties</td>
<td>46</td>
</tr>
<tr>
<td>4.1</td>
<td>Emotions and perceptions of threat OLS regression results</td>
<td>59</td>
</tr>
<tr>
<td>4.2</td>
<td>Support for refugee resettlement ordinal logistic regression results</td>
<td>63</td>
</tr>
<tr>
<td>4.3</td>
<td>Estimates of causal mechanisms</td>
<td>65</td>
</tr>
<tr>
<td>4.4</td>
<td>Willingness to contact MP regarding resettlement logistic regression results</td>
<td>69</td>
</tr>
<tr>
<td>A.1</td>
<td>Respondent types for standard list experiment</td>
<td>81</td>
</tr>
<tr>
<td>A.2</td>
<td>Respondent types defined by the response ((Z^<em>_i, D_i)) for the case in which responding (Z^</em>_i = 0) is to provide the sensitive response</td>
<td>82</td>
</tr>
<tr>
<td>A.3</td>
<td>Observed-data likelihood for the case in which (Z^* = 0) is the socially undesirable response</td>
<td>83</td>
</tr>
<tr>
<td>A.4</td>
<td>Weights (w_i^{(1)}) as calculated in the E-step for the case in which (Z^* = 0) is the socially undesirable response</td>
<td>83</td>
</tr>
<tr>
<td>A.5</td>
<td>Multivariate list experiment results (control-items sub-model)</td>
<td>87</td>
</tr>
<tr>
<td>C.1</td>
<td>Balance checks for data collected within 2 days of the Paris terrorist attacks</td>
<td>95</td>
</tr>
<tr>
<td>C.2</td>
<td>Regression results to examine sample composition before and after the Paris attacks</td>
<td>96</td>
</tr>
<tr>
<td>C.3</td>
<td>Support for refugee resettlement ordinal logistic regression results</td>
<td>97</td>
</tr>
<tr>
<td>C.4</td>
<td>Regression results comparing anxiety indexes including and excluding anger indicator</td>
<td>99</td>
</tr>
<tr>
<td>C.5</td>
<td>Second-stage causal mechanism regression models</td>
<td>104</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Simulation study 1 results ........................................ 19
2.2 Simulation study 2 results ........................................ 21
2.3 Simulation study 3 results ........................................ 22
2.4 Agreement with statement regarding gender equality in politics ........... 26
2.5 Predicted probabilities of support for and misreporting of a belief that “women are as competent as men in politics” ........................................... 29

3.1 AM scaling stylized example ........................................ 34
3.2 Comparison of AM Models including and excluding self-placement ........ 39
3.3 Comparison of linear and ordinal AM scaling ................................ 41
3.4 AM Scaled Senator Ideology (CCES) .................................. 42
3.5 Posterior of political ideology for voters and political actors ................... 45

4.1 Anxiety toward and sympathy for refugees across time ......................... 60
4.2 Security and cultural threat indexes across time .................................. 61
4.3 Support for refugee resettlement across time .................................... 62
4.4 Estimated effect of terrorist attacks on public support for refugee resettlement ... 64
4.5 Willingness to contact Member of Parliament across time ...................... 67
4.6 Willingness to contact an MP regarding resettlement before and after Paris attacks 68

A.1 Control-items sub-model for Simulation Study 2 ......................... 85
A.2 Misreport sub-model for Simulation Study 2 ................................ 86

B.1 Comparison of Hare et al. (2015) AM scaling model and model including self-placement ........................................ 90
B.2 Comparison of heterogeneous and homogeneous ordinal AM scaling models (political actor ideology $\xi_j$) ........................................ 92
B.3 Comparison of heterogeneous and homogeneous linear AM scaling models (respondent ideology $\theta_i$) ........................................ 93
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.1 Effect of attacks on anxiety index, including and excluding 'anger' indicator</td>
<td>98</td>
</tr>
<tr>
<td>C.2 Economic threat index across time</td>
<td>100</td>
</tr>
<tr>
<td>C.3 Retrospective economic evaluations across time</td>
<td>101</td>
</tr>
<tr>
<td>C.4 Sensitivity analysis for sequential ignorability assumption</td>
<td>102</td>
</tr>
<tr>
<td>C.5 Sensitivity analysis for homogeneous interaction assumption for cultural threat index</td>
<td>105</td>
</tr>
<tr>
<td>C.6 Mentions of Paris terrorist attacks in national media</td>
<td>106</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Introduction

In the quantitative study of politics, the growth in the size, affordability, and availability of data and computational resources over the past two decades has been staggering. Large and novel data sources now figure prominently in the work of some of the most impressive current research in the political and social sciences. For example, King, Pan and Roberts (2013, 2014) use millions of social media posts and a large-scale field experiment to show that the Chinese government uses censorship on social media to prevent collective action; Bonica (2013, 2014) introduces scaling methods for national campaign finance data to estimate the ideology of candidates and campaign donors on a common ideological scale; and Wang et al. (2015) demonstrate that sophisticated statistical modeling techniques can be applied to highly unrepresentative X-box data to forecast state-level electoral results with accuracy. These examples vary in substance, but they share a common foundation in their development of and reliance on computationally intensive methods and/or the use of large and novel data sets to answer theoretically and politically important research questions. The growth in the accessibility of these computational resources and in the production and availability of new sources of data now enable researchers to answer questions that would otherwise remain unanswered.

This dissertation follows in the footsteps of these studies by presenting three articles that aim to ask and answer research questions that rely extensively on the availability of large data sets and/or access to extensive computational resources. In the articles presented, I seek to ask questions that are motivated by a desire to provide methodological tools to researchers and to answer research questions that would otherwise be impossible to answer without the recent

---

1In two articles presented in this dissertation, the fitting of the proposed methods to simulated and real-world data number in the weeks. Much of this computational workload was conducted on Calcul Québec’s large-scale computing network. I thank Félix-Antoine Fortin for facilitating access to and providing support for Calcul Québec’s computing network.
growth in the availability of novel large-scale data sources.

In particular, this dissertation pursues a broader agenda to examine how the availability of large-scale survey data can be used to motivate methodological innovation and answer substantive questions of theoretical importance. In the first article, I ask who is more likely to misreport their responses to sensitive survey questions, and develop a methodological technique and accompanying open-source statistical software to answer this question. The method is applied to a large-scale data set regarding Canadians’ attitudes toward the capabilities of women in politics. In the second article, I develop a method to examine and correct for perceptual bias in survey respondents’ understanding of their own political ideology and that of political candidates, parties, and institutions. I apply the method to a large-scale dataset to estimate the ideology of the electorate and national and local political representatives on a common ideological scale. Finally, in the last article, I turn from methodological developments for large-scale data sources to the in-depth study of a novel large-scale time series cross-sectional survey that permits examination of the causal effects of a major terrorist attack on attitudes toward refugees and political mobilization. The third article, which presents results from a natural experiment, demonstrates how political science’s increasing focus on causal inference can benefit from the availability of large-scale time series cross-sectional survey data.

1.2 The Data and Computational Social Science Revolution

In quantitative political science and the social sciences more generally, the defining characteristic of the last two decades has been the growth both in researchers’ access to large data sources and the computational resources to analyze them. Although the social sciences have been slower to take advantage of these resources than have their cousins in the natural and physical sciences (Lazer et al., 2009), these changes are now remaking the landscape of social inquiry.

The most obvious manifestation of these changes is the growth in access to and the use of novel and large-scale data sources. These data are often described primarily by their size, as “Big Data” (e.g. Patty and Penn, 2015; Ashworth, Berry and Bueno de Mesquita, 2015; Nagler and Tucker, 2015; Grimmer, 2015; Lazer and Radford, 2017), a term that generally refers to data that are many orders of magnitude larger than those from traditional sources. Such data are often unstructured and take forms that are frequently foreign to researchers who are accustomed to working with well-arranged rectangular data sets. Indeed, previous applications of the term “Big Data” were specifically given to data whose size and shape were unmanageable given the currently available software.

But the power of the data that have emerged from the growth of information and communications technology is not typically their absolute size, but rather their novelty, and with more traditional sources such as surveys, their relative size. The primary benefit of these data is that
they allow researchers to tackle research questions that would be difficult to answer with data from traditional sources or with those collected through traditional means, but of small size. Emblematic of this shift in the availability and use of novel data sources is the growing amount of research whose data sources in the past would have been extremely time- and resource-intensive, or been wholly unavailable. Examples include the use of audio of oral arguments at the U.S. Supreme Court to measure emotions to infer judges attitudes toward cases (Dietrich, Enos and Sen, 2016); the use of video and audio from police body cameras to examine levels of respect shown toward blacks and whites during traffic stops (Voigt et al., 2017); the use of local rainfall data to examine the effect of protest attendance on the adoption of protest movements’ defining issue positions (Madestam et al., 2013); and the use of field experimental data collected through the apartment sharing application AirBnB to examine discrimination in the sharing economy (Edelman, Luca and Svirsky, 2017). These examples demonstrate that it is more often the ease of access to and cost of data that have defined recent changes in social science research, rather than the size of those data.

Of course, many of these new data sources are not only remarkable in their novelty, but also in their size. The most noteworthy case is the availability of large-scale social media data. These data have been put to good use through examination of the causes and consequences of network structure online (e.g. Barberá, Wang, Bonneau, Jost, Nagler, Tucker and González-Bailón, 2015; Coppock, Guess and Ternovski, 2016; Bond et al., 2017); descriptions of the shape and polarization of political conversation (e.g. Barberá, Jost, Nagler, Tucker and Bonneau, 2015; Boutyline and Willer, 2017); the prediction of political outcomes (e.g. Beauchamp, 2017; Zeitzoff, Forthcoming); and the development of ideological scaling methods for network data (e.g. Barberá, 2015; Bond and Messing, 2015).

However, despite the frequent emphasis on changes in the social science due to the availability of exceptionally large-scale data sets, some of the most important consequences of the data and computational revolution are on a smaller scale. The biggest of these in survey research is the increasing affordability of large public opinion surveys, brought about by the introduction of online survey panels (Clarke et al., 2008). Aside from some notable exceptions (e.g. Wang et al., 2015), the size of current public opinion surveys are not ‘Big Data’. The Canadian and American National Election studies have increased in size, for example, from roughly 1,000 to 2,000 respondents per study in the second half of the twentieth century to roughly 5,000 to 10,000 in the last decade. Other important examples are the Cooperative Congressional Election Study in the U.S., which in its 2016 iteration contains roughly 50,000 observations, and the Local Parliament Project in Canada, which during the 2015 federal election collected roughly 40,000 survey responses.

The increase in the size of these samples are thus modest — a single order of magnitude. But they have important consequences for the scope of developments in political methodology, the
analysis of sub-groups in researchers’ populations of interest, and the examination of changes in opinion across time. For example, the availability of larger-scale survey data and increases in computational power have led to rapid growth in research that aims to measure the ideology of citizens and legislators on the same scale (e.g. Bafumi and Herron, 2010; Tausanovitch and Warshaw, 2013; Lo, Proksch and Gschwend, 2014; Hare et al., 2015; Saiegh, 2015; Jessee, 2016; Ramey, 2016). The development and application of these methods are now helping to bridge the gap between the extensive literature on legislator ideology and the literature concerning the ideology of individual citizens. This research would be impossible without increases in sample size to permit public opinion to be estimated at smaller levels of geography and increases in computational power to fit models developed for these data.2

Increases in computational power and survey size have also encouraged researchers to ask questions and implement methods that would otherwise be left unaddressed. One area of research that has benefitted greatly from these developments is that which examines methods for estimating sub-national public opinion and uses the resulting estimates to better understand the causes of policy change (e.g. Lax and Phillips, 2009a, b; Warshaw and Rodden, 2012; Tausanovitch and Warshaw, 2013; Enns and Koch, 2013; Leemann and Wasserfallen, Forthcoming). For example, researchers have asked to what degree municipal governments are responsive to the preference of citizens at the municipal level (Tausanovitch and Warshaw, 2013); what the effects of public opinion at the state level are on the adoption of policies affecting gays and lesbians (Lax and Phillips, 2009a); and how we can measure state- and district-level public opinion and whether it is tied to referenda and voting behavior (Lax and Phillips, 2009b; Warshaw and Rodden, 2012).

Furthermore, methods to measure sensitive attitudes and behaviors, which are well-known to reduce bias at the cost of efficiency (i.e. sample size) (see Rosenfeld, Imai and Shapiro, 2016, 797-798), have recently grown in use. The development and application of methods for their analysis have benefitted substantially from the increases in computational power and survey size, and many methods have arisen to take advantage of this (e.g. Gingerich, 2010; Bullock, Imai and Shapiro, 2011; Imai, 2011; Blair and Imai, 2012; Blair, Imai and Zhou, 2015; Gingerich et al., 2016).

A final example is the growth in the use of time series cross-sectional survey data, which have grown substantially in size and frequency and have been put to academic and public use most visibly through electoral forecasting models (e.g. Lock and Gelman, 2010; Linzer, 2013; Wang et al., 2015). The frequency of polling in particular — a result of the ease and affordability of survey data collection through online surveys (Clarke et al., 2008) — has permitted researchers to more precisely demonstrate changes in public opinion across time. This has resulted in well-

2These models frequently rely heavily on the computational power put to use by the recent development of Bayesian inference engines such as Stan (Carpenter et al., 2017) and JAGS (Plummer, 2016).
publicized successes in election forecasting, most prominently during the U.S. 2012 Presidential election (for a review, see Pasek, 2015) and most recently during the 2017 U.K. general elections (Lauderdale and Rivers, 2017).3

This dissertation aims to build on these developments by asking and answering questions that depend on the substantial increases in survey sizes and computational power that have been made in recent years, and providing new methodological tools for researchers to do the same. In the following section, I summarize the goal and contribution of each article.

1.3 Organization & Overview

First article In the first paper, I ask who is more likely to misreport responses to sensitive survey questions. This is an important question because (1) survey responses to sensitive questions play an essential role in a wide variety of research programs and (2) it is well-known that survey respondents who are asked directly about socially sensitive attitudes and behaviors often provide responses that do not represent the truth. Respondents instead frequently provide survey responses to be consistent with social norms and expectations (Edwards, 1953, 1957; Warner, 1965; Sigall and Page, 1971; Bradburn et al., 1978; Himmelfarb and Lickteig, 1982; Fisher, 1993; Berinsky, 1999; Johnson and Van de Vijver, 2002).

To address this concern, researchers often turn to indirect measurement techniques. Indirect questioning techniques encourage truthful responses from survey respondents by providing anonymity inherently, as characteristic of survey question design: indirect questions are devised in ways to ensure that survey responses, in themselves, do not reveal whether specific respondents hold socially unacceptable beliefs. The investigation of sensitive survey questions through indirect questioning is now a rapidly growing research programs. This research has been greatly strengthened by the recent development of a research program to develop methodological tools to enable responses to indirect questions to be investigated with much greater sophistication than was previously possible (Gingerich, 2010; Bullock, Green and Ha, 2010; Imai, 2011; Blair and Imai, 2012; Blair, Imai and Lyall, 2014; Imai, Park and Greene, 2015; Blair, Imai and Zhou, 2015). These methodological advances have focused on the development of methods to permit researchers to model responses to sensitive statements, to test research design assumptions, and to increase statistical efficiency.

The act of misreporting on sensitive questions, however, has typically been treated as a nuisance or bias in need of elimination. The first article thus highlights the substantive and methodological importance of turning the problem of measurement error due to misreporting on its head by focusing on misreporting as a methodologically and substantively important object

---

3The YouGov statistical model and forecast were conducted by political scientists Benjamin Lauderdale, Doug Rivers, Delia Bailey, and Jack Blumenau using a Bayesian multilevel regression and post-stratification (MRP) forecasting model.
of interest in and of itself. As demonstrated through empirical application, the method proposed in this article has the potential to answer research questions that are of considerable political relevance and scientific interest. In particular, it demonstrates that those on the ideological left are more likely to misreport their attitudes regarding the capabilities of women in politics than are those on the ideological right. Theoretically, I link this to differences in the strength of social norms regarding the expression of prejudiced beliefs that vary across the ideological spectrum.

The method is thus motivated by a theoretical desire to understand a vexing problem in social science research, but one that has been made difficult by the need for a substantial amount of data. The method takes advantage of the fact that surveys that include indirect measurement techniques are now relatively large. Furthermore, as the article demonstrates through simulation, large survey samples are necessary to examine the form of misreporting demonstrated in the article. This is because (1) indirect measurement techniques in themselves minimize bias in survey responses due to misreporting with a decrease in efficiency, and (2) theoretical interest is in the sub-sample of respondents who both hold a sensitive belief and misreport it, reducing statistical power even further. The method thus provides an important and useful new tool for the analysis of a critical but under-analyzed dimension of research into sensitive beliefs, attitudes, and behaviors. I provide statistical software written in the open-source statistical package R to encourage its use.

Second article In the second article, I examine how the electorate understands their own political ideology and that of national and sub-national political actors. In particular, I address a common but rarely investigated class of perceptual problems in survey research: that for some survey questions, such as ideological placement scales, survey respondents frequently perceive the scale on which they give their responses differently between one another (Brady, 1985). What might be deemed “Conservative” by one person on a typical ideological placement scale, for example, might be deemed “Moderate” by another. As a result, while respondents on the ideological left might consider a political party such as the Conservative Party or Republican Party to be extremely conservative, someone on the right might consider these parties to be moderate. The consequence is that survey respondents’ perceptions of their own ideology and that of political actors are not, in their raw form, measured on a common scale. This fact prevents comparison of responses between respondents, and between respondents and their political representatives.

The goal of the second paper is therefore to propose a method to correct for this problem and to apply this method to understand the distribution of ideology among the electorate and national and local political representatives in recent election years. To do so, I build on the classic work by Aldrich and McKelvey (1977), which proposes a statistical method to re-scale responses to ideological placement scales on to a common scale. The growth of large-scale
national and multi-national surveys that ask surveys respondents to place national as well as local-level actors such as governors and congressional representatives has led to the increasing use of Aldrich-McKelvey (AM) scaling to examine differences between citizens and political representatives (e.g. Saiegh, 2009; Hollibaugh, Rothenberg and Rulison, 2013; Lo, Proksch and Gschwend, 2014; Bakker et al., 2014; Saiegh, 2015; Ramey, 2016; Hare et al., 2015). A recent Bayesian implementation of AM scaling, moreover, provides a straightforward means to capture the error in the model’s parameters and to incorporate ideological placements of political actors who are placed by some respondents, but not by others (e.g. a respondent’s congressional representatives) (Hare et al., 2015).

As I discuss in the article, however, the method proposed by Aldrich and McKelvey (1977) is limited because it does not account for the fact that typical ideological placement scales are ordinal, with response categories that range, for example, from “Extremely liberal” to “Extremely conservative.” Furthermore, the method does not incorporate respondents’ ideological self-placement as an outcome in the model itself, which leads to poor estimates of ideology. To remedy this, I propose a heterogeneous ordinal AM scaling model. The model accounts for the ordinal nature of ideological placement data; the fact that some political actors are more difficult for respondents to place than others; and incorporates ideological self-placement as an outcome in the model itself. The upshot is that the model permits researchers to generate more valid and accurate estimates of ideology among the electorate and political actors.

As an empirical application, I apply the method to recent iterations of the Cooperative Congressional Election Study (CCES). CCES data are useful because they are large-scale (n ≈ 40-60,000) and contain survey responses regarding both national-level political actors such as political parties and Presidential candidates, and state-level actors such as House and Senate representatives. I first examine differences between the standard linear and proposed ordinal AM scaling model, which shows that legislators on the ideological extremes are perceived to be more extreme than the linear AM scaling model would suggest. In particular, Senator Bernie Sanders, in each of the last four mid-term and Presidential years is estimated to be far more extreme than his colleagues in the senate.

To investigate this result further, I examine to what degree Sanders is ideologically out-of-step with the U.S. electorate as a whole. This is an important question because in the aftermath of Hillary Clinton’s loss to Donald Trump during the 2016 U.S. Presidential election, some have suggested that, had Sanders won the Democratic Party primary elections, he would have also won the Presidential election. However, as application of the proposed method shows, Sanders is perceived to be substantially more liberal than the vast majority of the national electorate, with only 6% of it to his left. This compares with Clinton and Trump, who sit roughly at the median of Democratic and Republican voters respectively. Although not conclusive, this finding suggests that Sanders would have faced an uphill battle against perceptions that he was far out
Third article  In the final paper, I turn from a focus on large-scale survey data and the opportunities that they provide for methodological development to show how large-scale survey data can open up multiple avenues of investigation in the context of a natural experiment. The article, co-authored with Charles Breton, examines public opinion regarding Syrian refugees in the days immediately before and after the 2015 Islamic State terrorist attacks in Paris. Because the attacks were perpetrated by a Syria-based terrorist organization, they brought together two salient issues in modern international politics: the fight against terrorism and the large-scale resettlement of refugees.

To investigate the effects of the attacks on public opinion, we take advantage of a unique and exceptionally large-scale post-election survey (n = 18,600) that was first fielded less than 48 hours before the Paris terrorist attacks. The survey includes a wide-ranging set of questions regarding attitudes toward refugees as potential threats to security, culture, and the economy; support for the admittance of Syrian refugees to Canada; and political mobilization concerning the issue of refugee resettlement. The survey was fielded in a second wave less than 48 hours after the attacks, and then sent in subsequent waves each day for two weeks thereafter. The size of the data permit us to investigate the effects of a major terrorist attack on (1) public support for refugees, (2) attitudes and emotions toward refugees themselves, and (3) political mobilization regarding resettlement. Furthermore, because the survey was fielded in large daily waves for two weeks after the attacks, the article is the first to investigate, with precision, the duration of the effects of terrorism on public opinion.

Because the set of questions and survey data are large, our findings are many. First, our analysis shows that public attitudes toward refugee resettlement policy was resilient in the face of the attacks. In particular, we demonstrate that the attacks caused only a moderate 4 percentage point increase in opposition to refugee resettlement. Furthermore, we show that the attacks did not affect public sympathy for refugees themselves. Instead, we show that the attacks caused an increase is public concern over refugees as a security and cultural threat, and anxiety over the presence of Syrian refugees in Canada. Using recent advancements in methods for the investigation of causal mechanisms, we show, moreover, that it was through these changes in threat perceptions and emotions that led to the increase in opposition to Syrian refugee resettlement. We illustrate, furthermore, that the attacks increased political mobilization around the issue of resettlement, but did so asymmetrically: the attacks led to increased mobilization among those opposed to resettlement, but not among those supportive of resettlement. Finally, due both to the large size of the data and its having been fielded each day for two weeks after

---

4Charles Breton is currently a post-doctoral fellow in the Department of Political Science at the University of British Columbia.
the attacks, we show that the effects of the attacks on public opinion were decidedly short-lived. Within roughly 1-2 weeks, attitudes and emotions toward Syrian refugees had returned to their pre-attack levels. This finding has substantial implications for our understanding of the window of opportunity open to political entrepreneurs following large-scale terrorist attacks and for explaining the often short-lived political efforts to enact policy changes following large-scale acts of violence.

In the final chapter, I summarize the contributions of this dissertation and conclude by discussing the potential for traditional survey data and methods to complement the increasing use of 'Big data' for the study of public opinion and political behavior.
Chapter 2

The Statistical Analysis of Misreporting on Sensitive Survey Questions

2.1 Introduction

When survey respondents are asked directly about sensitive issues and behaviors, they frequently do not provide truthful responses. Would you be willing to support an African-American candidate for President? Do you believe women are as competent as men to serve in public office? Have you ever bribed a public official? Did you vote? Sensitive questions in survey research are many and play an integral role in numerous research programs. Respondents who misreport their beliefs and behaviors in order to adhere to social expectations, however, can be a substantial source of measurement error.

To reduce this measurement error — social desirability bias — methodologists have developed a variety of indirect questioning techniques to elicit truthful survey responses. Recent work on these techniques has led to the development of statistical methods to model survey responses to indirect questions within a multivariate regression context. Methodologists have developed such methods for, among others, the randomized response technique (Gingerich, 2010; Blair, Imai and Zhou, 2015), endorsement experiment (Bullock, Imai and Shapiro, 2011; Blair, Imai and Lyall, 2014), and list experiment (Corstange, 2009; Imai, 2011; Blair and Imai, 2012; Glynn, 2013; Imai, Park and Greene, 2015). These methods now enable researchers to investigate the causes and consequences of sensitive attitudes, beliefs, and behaviors in much more sophisticated ways than previously possible.

The development and availability of these methods has not, however, been accompanied by a substantial increase in our understanding of misreporting generally or how the predictors of misreporting on sensitive survey questions vary with the attitude, belief, or behavior of interest to researchers. Methodological work in this area has instead focused primarily on developing
methods to model indirect question responses, to test research design assumptions, and to
increase statistical efficiency. Misreporting on survey questions — the fundamental problem
that indirect questioning techniques are designed to solve — is typically treated as a bias to
be eliminated rather than the object of interest itself.\footnote{One notable, if unique, exception to this is the literature on voter validation, in which survey responses about
whether respondents have voted are verified using publicly available voting records (Silver, Anderson and Abramson, 1986; Karp and Brockington, 2005; Fullerton, Dixon and Borch, 2007; Ansolabehere and Hersh, 2012).} There is thus a substantial gap in the
literature for methods to model, in a multivariate framework, the predictors of misreporting
on sensitive survey questions. Filling this gap is important from a methodological perspective,
but also substantively: explaining variation in survey misreporting is integral to understanding
manifold aspects of public opinion and the practice of politics more generally (Noelle-Neumann,
1993; Kuran, 1998; Berinsky, 1999).

To address this gap, this article proposes a method to enable researchers to model the
predictors of misreporting on sensitive survey questions. It develops a maximum likelihood
estimator to permit researchers to model, within a multivariate regression context, the joint
survey response to a list experiment and direct question simultaneously. In doing so, it enables
researchers to model whether survey respondents provide one response to an indirect question
that is designed to elicit the truth, but answer otherwise when asked to reveal their response
openly to a direct question. The upshot is that the proposed method provides a means for
researchers to empirically examine why people misreport their attitudes toward a wide variety
of politically and socially sensitive issues. To facilitate its application, the method is provided for
researchers as open-source software.

\section*{2.2 Proposed Methodology}

In this section, I build on the method introduced by Imai (2011) by proposing a method to
model survey responses to a list experiment and to a direct question simultaneously within
a multivariate regression setting. The aim of the proposed method is to enable researchers
to model the probability that a respondent provides one response to a sensitive item in a list
experiment, but answers otherwise when asked to reveal that answer as a response to a direct
question.

To simplify exposition, the case addressed in this section is that in which responding affir-
matively to the sensitive question is to give the socially unacceptable response. In the online
appendix, an analogous method is presented for the case in which responding affirmatively to
the sensitive question is to give the socially acceptable response.
2.2.1 The list experiment setup

To clarify notation and to demonstrate concretely how a list experiment works, I use as a running example the list experiment conducted by Kuklinski, Cobb and Gilens (1997). In this study, respondents were divided at random into control and treatment groups and those in the control group assigned the following question and list of items:

Now I am going to read you three things that sometimes make people angry or upset. After I read all three, just tell me how many of them upset you. I don’t want to know which ones, just how many.

• The federal government increasing the tax on gasoline
• Professional athletes getting million-dollar contracts
• Large corporations polluting the environment

Respondents in the treatment group were presented with the same question and a list of items that included those received by the control group in addition to the sensitive item of interest:

• The federal government increasing the tax on gasoline
• Professional athletes getting million-dollar contracts
• Large corporations polluting the environment
• A black family moving in next door

The logic underlying a list experiment such as this is both simple and powerful. Suppose first that a random sample of \( N \) respondents \( i = 1, \ldots, N \), is drawn from a population of interest and respondents are assigned at random into treatment and control groups, where \( T_i \in \{0, 1\} \) denotes treatment status. Those in the control group \( (T_i = 0) \) are then assigned a list of items. The number of items in a typical control list is 3 or 4, which I denote by the integer \( J \). In the list experiment described above, for example, the control list contains 3 items. Those in the treatment group \( (T_i = 1) \) are assigned the same list of \( J \) items in addition to the sensitive item, which in the example above is “A black family moving next door.” Each respondent therefore receives \( J + T_i \) items in total. Finally, respondents are asked how many of the items in the list would make them upset, how many they agree with, or an analogous question that is consistent with a “how many” question formulation.

The reason that list experiments are able to elicit truthful responses is because respondents are not asked about each of the \( J + T_i \) items individually. This design property is powerful because the response to the sensitive item for any respondent in the treatment group cannot be known unless all \( J + 1 \) items are answered affirmatively, or unless none of them are. The goal of the list experiment is therefore to ensure that survey responses cannot be deconstructed.
by survey enumerators and researchers per item. This question design property provides respondents with the opportunity to express their beliefs safe in the knowledge that researchers cannot determine their response to the sensitive item (or any other item).

To see how analysis of a list experiment proceeds, let $Y_i \in \{0, \ldots, J + T_i\}$ denote a respondent’s answer to the list experiment question. This variable can be decomposed into the sum of two latent variables as follows:

$$Y_i = \begin{cases} Y_i^* & \text{if } T_i = 0 \\ Y_i^* + Z_i^* & \text{if } T_i = 1, \end{cases}$$

(2.1)

where $Y_i^* \in \{0, \ldots, J\}$ is the latent response to the $J$ control items, and $Z_i^* \in \{0, 1\}$ is the latent response to the sensitive item. As is clear from Equation 2.1, responses to the list experiment differ depending on treatment status only by the response to the sensitive item. Imai (2011) demonstrates that, given a set of assumptions, the difference in the mean of $Y$ for the treatment and control groups provides an unbiased estimate of the proportion of the population that would truthfully answer affirmatively to the sensitive item.

For much research in the social sciences, it will also frequently be of interest to model the predictors of responding affirmatively to the sensitive item. Imai (2011) therefore develops and implements a maximum likelihood estimator to enable the sensitive item in a list experiment to be modeled within a multivariate regression context. The method that I propose below builds on the framework introduced by Imai (2011), which I detail further in the online appendix (see also Blair and Imai, 2012).

### 2.2.2 The direct question and misreporting

To identify the response pattern consistent with misreporting, I begin by introducing an assumption that I use in conjunction with those for the standard list experiment model (Imai, 2011, 408-409). To demonstrate this assumption concretely, I return to the list experiment conducted by Kuklinski, Cobb and Gilens (1997) for illustration. Respondents in this study were not asked the sensitive question directly, but for the sake of concreteness we can assume that it would have been asked as follows:

**Would you be upset if a black family moved in next door?**

One drawback to the list experiment is that per-item protection is jeopardized for respondents in the treatment group who respond 0 or $J + 1$. It is recommended that researchers therefore devise items to avoid these floor and ceiling responses by, for example, using control items that are negatively correlated (Glynn, 2013) and/or testing these items in a pilot study when possible.

These assumptions are the following: (1) that $Y_i^*$ and $Z_i^*$ are independent of $T_i$ (through randomization); (2) that respondents do not misreport their response to the sensitive item in the list experiment; and (3) that $Y_i^*_{T_i=0} = Y_i^*_{T_i=1}$. In brief, this third assumption states that responses to the $J$ control items are unaffected by treatment assignment: that receiving the sensitive item does not change respondents’ answers to the control items. For further details regarding these assumptions, see Imai (2011, 408-409).
Let $D_i \in \{0,1\}$ denote an individual’s response to the direct question, with response categories “No” ($D_i = 0$) and “Yes” ($D_i = 1$). To denote “misreporting”, let $U_i^* \in \{0,1\}$ be an indicator variable representing whether the response to the direct question is different from that given to the sensitive item in the list experiment: $U_i^* = 1(Z_i^* \neq D_i)$, such that an individual is said to be misreporting if $U_i^* = 1$.

**Monotonicity assumption** The monotonicity assumption states that individuals who hold the socially undesirable belief may or may not misreport it when asked to express that belief openly. Those who do not hold the socially undesirable belief are assumed to state so truthfully.\(^4\) This assumption follows naturally from our general understanding of social desirability bias, such that those who do not hold a socially undesirable belief are not expected to feel social pressure to misreport it by answering in a socially undesirable manner. More formally, this assumption can be stated as follows:

$$U_i^* = \begin{cases} 
0 & \text{if } Z_i^* = 0 \\
0 \text{ or } 1 & \text{if } Z_i^* = 1.
\end{cases} \quad (2.2)$$

In our running example, this assumption means that only those who would be upset if an African-American family were to move next door may or may not claim otherwise when asked to reveal that attitude openly. Respondents who would not be upset, on the other hand, are assumed to respond truthfully. Put differently, if this were not the case, it would mean that some respondents who are not prejudiced would misreport by claiming to be prejudiced nonetheless.\(^5\)

**2.2.3 Respondent types**

Given the monotonicity assumption, we can identify three response patterns to the list experiment and direct question ($Z_i^*$, $D_i$). Substantively, these define the following respondent types:

1. those who hold the sensitive belief and misreport it when asked directly ($Z_i^* = 1$, $D_i = 0$);
2. those who hold the sensitive belief and do not misreport it ($Z_i^* = 1$, $D_i = 1$), and
3. those who do not hold the sensitive belief and do not misreport it ($Z_i^* = 0$, $D_i = 0$). These response patterns and their respective descriptions are presented in Table 2.1. By contrast, the standard list experiment model (Imai, 2011) identifies two respondent types: those who hold the socially unacceptable belief and those who do not. The method introduced here therefore differs in that it identifies two sub-types of respondents among those who provide the socially unacceptable

---

\(^4\)This assumption is analogous to the “one-sided lying” assumption used by Gingerich et al. (2016, 137) for the crosswise design and the monotonicity assumption by Aronow et al. (2015).

\(^5\)In practice, this assumption may also be violated if respondents incorrectly answer the direct or list experiment question due to inattention. It is therefore recommended that researchers employ screener questions when possible, as in survey research more generally (Berinsky, Margolis and Sances, 2014). An example screener question specific to the list experiment is provided in the empirical application section further below.
Table 2.1: Respondent types defined by the response \((Z^*_i, D_i)\)

<table>
<thead>
<tr>
<th>Type</th>
<th>Sensitive</th>
<th>Direct</th>
<th>Misreport</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misreport sensitive</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Respondent holds the sensitive belief but misreports it when asked directly.</td>
</tr>
<tr>
<td>Truthful sensitive</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>Respondent holds the sensitive belief and states so truthfully when asked directly.</td>
</tr>
<tr>
<td>Non-sensitive</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Respondent does not hold the sensitive belief and states so truthfully when asked directly.</td>
</tr>
</tbody>
</table>

Note that the respondent type for the response \((Z^*_i = 0, D_i = 1)\) is undefined by the monotonicity assumption.

response: those who admit openly to holding the socially unacceptable belief, and those who misreport it.

2.2.4 Maximum likelihood estimator

To model these responses within a regression-based framework, I define three sub-models to model the joint distribution \((Y^*_i, Z^*_i, U^*_i)\). First, the sub-model used to model the response to the sensitive item, \(Z^*_i\), is defined as follows:

\[
g(x; \delta) = \Pr(Z^*_i = 1 | X_i = x; \delta), \tag{2.3}
\]

where \(X_i\) is a vector of covariates and \(\delta\) is a vector of parameters to be estimated.

The second sub-model models the probability that a respondent’s answer to the sensitive item in the list experiment question is different from that given to the direct question:

\[
j(x, z, t; \gamma) = \Pr(U^*_i = 1 | X_i = x, Z^*_i = z, T_i = t; \gamma), \tag{2.4}
\]

where \(\gamma\) is a vector of parameters to be estimated. Note that, by the monotonicity assumption, \(j(x, 0, t; \gamma) = 0\). In other words, the probability that a respondent misreports if that respondent does not hold the socially unacceptable belief is 0 by assumption.

In the setup described here, survey respondents are presumed to answer the direct question some time after providing a response to the list experiment question. Although researchers may also ask the direct question before the list experiment, it is recommended that the list experiment question be asked first. Doing so permits researchers both to straightforwardly test whether receiving the treatment list affects the response to the direct question, and if such a relationship exists, to account for this by modeling it in the misreport sub-model (Equation 2.4).\(^6\)

\(^6\)The coefficient for the treatment indicator in the misreport sub-model indicates whether those who received the
Theoretically, such a relationship might exist because those in the treatment group recall their previous response to the sensitive item in the list experiment and provide the same response to the direct question for reasons of cognitive ease or to be consistent on principle. To avoid reliance on modeling this relationship, it is recommended that researchers aim to minimize or eliminate this potential effect through conscientious survey design. It is suggested, therefore, that the list experiment and direct questions be separated from one another in the survey by other sets of unrelated questions when possible.

Lastly, the third sub-model models the response to the control items, $Y^*$, as follows:

$$h(y|x, z, u; \psi) = \Pr(Y_i^* = y|X_i = x, Z_i^* = z, U_i^* = u; \psi),$$

where $\psi$ is a vector of parameters to be estimated. For implementation, binomial regression models with a logistic link are used for each of these sub-models.

Given the models $g(x; \delta)$, $j(x, z, t; \gamma)$, and $h(y|x, z; u; \psi)$, I now derive the observed-data likelihood. Due to the likelihood function’s length, the components of the likelihood for each response pattern are provided in Table 2.2. The observed-data model likelihood is the product of the relevant individual likelihoods as given in the last column of the table.

### 2.2.5 Model optimization

For optimization, I implement an expectation-maximization (EM) algorithm (Dempster, Laird and Rubin, 1977). The EM algorithm requires the complete-data likelihood function, which is the likelihood function for the case in which the latent variables $Z_i^*$, $Y_i^*$, and $U_i^*$ are fully observed. This function is given as follows:

$$L(\delta, \gamma, \psi; \{T_i, Y_i, X_i, Z_i^*, U_i^*\}_{i=1}^n) =$$

$$\prod_{i=1}^n \left\{ \left(1 - g(X_i; \delta)\right) \left(1 - j(X_i, 0, T_i; \gamma)\right) h(Y_i|X_i, 0, 0; \psi) \right\}^{1(Z_i^* = 0 \land U_i^* = 0)}$$

$$\times \left\{ g(X_i; \delta) \left(1 - j(X_i, 1, T_i; \gamma)\right) h(Y_i - T_i|X_i, 1, 0; \psi) \right\}^{1(Z_i^* = 1 \land U_i^* = 0)}$$

$$\times \left\{ g(X_i; \delta) j(X_i, 1, T_i; \gamma) h(Y_i - T_i|X_i, 1, 1; \psi) \right\}^{1(Z_i^* = 1 \land U_i^* = 1)}.$$

The EM algorithm is iterative, consisting of two steps. The expectation step calculates the conditional expectation of the response $(Z_i^*, D_i)$, given the model parameters $\delta$, $\gamma$, and $\psi$ from the most recent iteration of the maximization step. The weights calculated from this step

---

7Note that this potential problem is not avoided by asking the direct question first. In this case, one will similarly be concerned that providing a response to the direct question affects responses to the list experiment question, potentially leading to a violation of the “no liar” assumption (Imai, 2011, 408-409).
### Table 2.2: Observed-data likelihood

<table>
<thead>
<tr>
<th>Observed variables</th>
<th>Latent variables</th>
<th>Observed-data likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_i$</td>
<td>$Y_i$</td>
<td>$D_i$</td>
</tr>
<tr>
<td>0</td>
<td>$Y_i$</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>$Y_i$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0 or 1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>$J+1$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>$Y_i &gt; 0$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 2.3: Weights, $w_i^{(j)}$, as calculated in the E-step

- **$w_i^{(\text{non-sensitive})}$**

$$w_i^{(\text{non-sensitive})} = \frac{(1-g(x_i;\delta))(1-j(x_i,0,t_i;\gamma))h(y_i|x_i,0,0;\psi)}{(1-g(x_i;\delta))(1-j(x_i,0,t_i;\gamma))h(y_i|x_i,0,0;\psi) + g(x_i;\delta)(1-j(x_i,1,t_i;\gamma))h(y_i-t_i|x_i,1,0;\psi) + g(x_i;\delta)j(x_i,1,t_i;\gamma)h(y_i-t_i|x_i,1,1;\psi)}$$

- **$w_i^{(\text{truthful sensitive})}$**

$$w_i^{(\text{truthful sensitive})} = \frac{g(x_i;\delta)(1-j(x_i,1,t_i;\gamma))h(y_i-t_i|x_i,1,0;\psi)}{(1-g(x_i;\delta))(1-j(x_i,0,t_i;\gamma))h(y_i|x_i,0,0;\psi) + g(x_i;\delta)(1-j(x_i,1,t_i;\gamma))h(y_i-t_i|x_i,1,0;\psi) + g(x_i;\delta)j(x_i,1,t_i;\gamma)h(y_i-t_i|x_i,1,1;\psi)}$$

- **$w_i^{(\text{misreport sensitive})}$**

$$w_i^{(\text{misreport sensitive})} = \frac{g(x_i;\delta)j(x_i,1,t_i;\gamma)h(y_i-t_i|x_i,1,1;\psi)}{(1-g(x_i;\delta))(1-j(x_i,0,t_i;\gamma))h(y_i|x_i,0,0;\psi) + g(x_i;\delta)(1-j(x_i,1,t_i;\gamma))h(y_i-t_i|x_i,1,0;\psi) + g(x_i;\delta)j(x_i,1,t_i;\gamma)h(y_i-t_i|x_i,1,1;\psi)}$$
represent the posterior predicted probabilities that respondents are each of the three types as presented in Table 2.1. The expressions used to calculate these weights are given in Table 2.3.

The maximization step then computes the parameters $\delta$, $\gamma$, and $\psi$ using the most recent values of the weights $w_i^{(j)}$ from the E-step that maximize the complete-data log-likelihood function, which is given as follows:

\[
Q(\delta, \gamma, \psi; \{Y_i, X_i, T_i, w_i^{(j)}\}_{i=1}^n) = \\
\sum_{i=1}^n w_i^{(\text{non-sensitive})} \left\{ \log (1 - g(X_i; \delta)) + \log (1 - j(X_i, 0, T_i; \gamma)) + \log h(Y_i|X_i, 0, 0; \psi) \right\} \\
+ w_i^{(\text{truthful sensitive})} \left\{ \log g(X_i; \delta) + \log (1 - j(X_i, 1, T_i; \gamma)) + \log h(Y_i - T_i|X_i, 1, 0; \psi) \right\} \\
+ w_i^{(\text{misreport sensitive})} \left\{ \log g(X_i; \delta) + \log j(X_i, 1, T_i; \gamma) + \log h(Y_i - T_i|X_i, 1, 1; \psi) \right\}
\]  

(2.7)

The EM algorithm iterates between the E-step and M-step until convergence. Because the algorithm is not guaranteed to converge to a global maximum and can depend on the initialization of the starting parameters, the algorithm can be run multiple times using randomized starting values and the parameters from the run that results in the highest log-likelihood used as the model solution. Following optimization, standard errors are calculated through numerical approximation.

### 2.3 Monte Carlo Simulations

To investigate the properties of the proposed method and to ensure its ability to estimate population parameters as intended, I conduct three Monte Carlo simulation studies. The first study examines how the proposed method compares to the method introduced by Imai (2011), which does not incorporate a direct question response (henceforth referred to as the “standard estimator”). In the first study, a comparison is made with respect to root mean squared error (RMSE) for populations containing different proportions of those who misreport. The second study compares the proposed estimator to the standard estimator in terms of RMSE in a model with covariates. Lastly, I show how RMSE in the misreport sub-model decreases as the proportion of those holding the sensitive belief and those who misreport increases, both of which are important practical concerns for researchers.

---

8All simulations in this section were made on the supercomputer Colosse, managed by Calcul Québec and Compute Canada. The operation of this supercomputer is funded by the Canada Foundation for Innovation (CFI), the ministère de l’Économie, de la science et de l’innovation du Québec (MESI) and the Fonds de recherche du Québec - Nature et technologies (FRQ-NT).
These figures present a comparison by root mean squared error (RMSE), bias, and confidence interval coverage of the standard list experiment estimator (dashed line) (Imai, 2011) to the proposed estimator (solid line) in populations containing different proportions of those who misreport.

2.3.1 Simulation study 1

In this study, I demonstrate through simulation how researchers can increase efficiency in the sensitive-item sub-model by incorporating information from responses to a direct question. This is an important secondary benefit to the proposed method because it lowers the bias-variance tradeoff inherent in using a list experiment in place of a direct question: using both a list experiment and direct question simultaneously provides the benefits of lower bias without sacrificing the same degree of efficiency as using a list experiment alone.\textsuperscript{9}

To conduct the study, I simulate data from a list experiment containing 4 control items, with the intercept and sensitive-item parameters in the control-items sub-model set to 0 (on the logistic scale). Responses to a direct question are simulated from an intercept-only misreport sub-model where the population proportion of those who misreport is varied from 0.1 to 0.9 at 0.1 intervals. With this setup, 10,000 datasets are simulated, each containing 3,000 respondents. Models for the proposed and standard estimators are then fit to each dataset.

Results from these models are presented in Figure 2.1. As shown in the first panel, the proposed estimator substantially improves on the standard estimator in terms of RMSE across the range of populations containing different proportions of those who misreport. These gains in efficiency result from the additional information provided by the direct question response: respondents who openly admit to holding the socially unacceptable belief through the direct question are, by the monotonicity assumption, assumed to provide the same response to the sensitive item in the list experiment, reducing estimation uncertainty. In the second and third

\textsuperscript{9}Efficiency gains through the use of direct question responses in concert with list experiment responses have also been demonstrated by Aronow et al. (2015).
panels of Figure 2.1, both the standard and proposed estimators appear unbiased, with confidence interval coverage close to the nominal rate.

### 2.3.2 Simulation study 2

In this study, I examine the properties of the proposed and standard methods in the more typical scenario in which a researcher wishes to model responses to the list experiment and direct question with covariates. To do so, I simulate two variables for use as model covariates: the first covariate is a binary variable drawn from a binomial distribution with probability 0.5; the second, a continuous variable drawn from a univariate normal distribution with mean and variance 1. The parameters for the intercept and the continuous and binary covariates are set respectively to \((-0.25, 0.25, -0.5)\) for the control-items sub-model; \((0, 0.25, -0.5)\) for the sensitive-item sub-model; and \((0.5, -0.5, 0.25)\) for the misreport sub-model. The parameter for the sensitive item in the control-items sub-model \((Z^*)\) is set to 0.5 and that for whether someone is misreporting \((U^*)\) is set to 0 to permit estimator comparison. Given the parameters as defined above, 10,000 datasets are simulated with covariates drawn from the appropriate distributions for samples of size 2,000, 3,000 and 5,000. The proposed and standard estimators are then fit to each simulated dataset.

Figure 2.2 presents results from this procedure for the sensitive-item sub-model. Figures for the control-items and misreport sub-models are presented in the online appendix. As Figure 2.2 shows, in line with results from the first simulation study, RMSE is lower for the proposed estimator compared to the standard estimator for all parameters. Parameter estimates appear unbiased and confidence interval coverage is close to the nominal rate, improving as expected with sample size. As this study shows, researchers who are only interested in the sensitive item sub-model can nevertheless benefit from the incorporation of information from a direct question in the form of lower error in parameter estimates.

### 2.3.3 Simulation study 3

In the final simulation study, I show how error in parameter estimates in the misreport sub-model varies as a function of both the proportion of the population that hold the sensitive belief and the proportion that misreport it. Both of these properties will be important considerations for applied researchers. First, as the simulations will show, investigation into misreporting requires substantial sample sizes when very few members of a population hold the socially unacceptable belief. This occurs because the misreport sub-model is used to distinguish between those who misreport and those who tell the truth within the sub-population that holds the sensitive belief: when very few respondents hold the sensitive belief, there will be little data from which to precisely estimate model parameters. Second, the simulations show that when
These figures show a comparison of the standard estimator (black) (Imai, 2011) and proposed estimator (white) by root mean squared error (RMSE), bias, and confidence interval coverage for the sensitive item sub-model of a list experiment with covariates. Results for the control-items and misreport sub-models are available in the online appendix.

very few members of a population misreport, estimation error in the misreport sub-model will be higher compared to populations containing higher rates of misreporting. To demonstrate these properties, I run two complementary simulation studies, varying the proportions of the population that hold the sensitive belief in the first, and the proportion of those who misreport in the second.

To conduct these studies, I simulate 10,000 datasets with sample sizes of 2,000, 3,000, and 5,000 respondents. For the first part of the study, the proportion of the population that hold
These figures show root mean squared error (RMSE), bias, and confidence interval coverage for two Monte Carlo simulation studies that examine the properties of the misreport sub-model. Panel A presents results for the misreport sub-model from a simulation study in which the proportion of those holding the sensitive belief is varied from 0.1 to 0.9 at 0.1 intervals; Panel B presents results from the misreport sub-model in which the proportion of those who misreport is varied from 0.1 to 0.9 at 0.1 intervals.

In the first study, therefore, interest is in error as it varies with the proportion of the population that hold the sensitive belief; in the second, interest is in error as it varies with the proportion of population that misreport.

Results from the first and second parts of the simulation study are presented in panels A and B of Figure 2.3. In the first panel (A), RMSE in estimates of the population proportion of those who misreport is shown to decrease as the proportion of those in the population who
hold the sensitive belief increases. Similarly, in the second panel (B), RMSE is decreasing in the proportion of those who misreport. In both cases, therefore, error in parameter estimates are shown to be larger when either very few members of a population hold the sensitive belief, or when very few members of a population misreport it. In each case, estimates appear slightly biased when very few members of a population hold the sensitive belief or misreport it, but appear unbiased elsewhere, with confidence interval coverage close to the nominal rate. The upshot is that when the misreport sub-model is of interest to researchers, care should be taken in selecting a sensitive item during the survey design stage: when researchers expect that very few respondents will hold the socially unacceptable belief of interest or few will misreport it, larger sample sizes will be required to estimate parameters in the misreport sub-model with precision.

2.4 Empirical Application

In this section, I demonstrate how the proposed method can be used to answer substantive questions that have until now remained difficult to investigate empirically. As an application, I examine the question of whether those on the ideological left are more likely to misreport their prejudices than those on the right.

2.4.1 Does the left misreport their prejudices more than those on the right?

Since the 1990s, the term “political correctness” has frequently been used to describe adherence to the set of social norms that prohibit the expression of prejudiced beliefs and the use of stigmatizing language (Hughes, 2010, 46). These social norms are now relatively pervasive across the Western world and apply to a wide variety of beliefs regarding gender, race, ethnicity, religion, and sexuality. As a consequence, those who hold beliefs that society deems prejudiced are, in effect, expected either to misrepresent them or to refrain from voicing them altogether.

The strength of the social norms governing expressions of real or perceived prejudiced beliefs appear, however, to vary strongly along ideological lines. This has been highlighted recently by a variety of public debates and commentaries regarding the validity and effects of these norms. For example, political and social commentators have questioned whether those on the left are unduly apprehensive in their criticisms of illiberal practices by some new immigrants and refugees (The Telegraph, 2016); whether university students are being “coddled” by the use of sanitized language and the suppression of sensitive topics (Lukianoff and Haidt, 2015); whether politically correct social and political norms are “perverting” the values of liberalism (Chait, 2015); whether liberal students are overly sensitive to controversial subjects (Schlosser, 2015; Taub, 2015); whether anti-racism protesters should be sheltered from media attention and criticism (Starr, 2015; Friedersdorf, 2015); and whether political appointments that achieve
equality in gender representation are being made for the sake of appearances rather than merit (Coyne, 2015). Opinions on these issues, furthermore, have recently been expressed by Western political leaders (e.g. Obama, 2015; Cameron, 2015; Cross, 2015). As prominent commentators on the issue have argued, the episodes that spark these debates have “created a culture in which everyone must think twice before speaking up, lest they face charges of insensitivity, aggression, or worse” (Lukianoff and Haidt, 2015). As another commentator writes, difficulty in determining what is or is not insensitive language or behavior has led to the elaboration of complicated sets of norms and terminology within communities on the political left (Chait, 2015).

To investigate whether the strength of these norms varies along ideological lines, I apply the proposed methodology to a large-scale list experiment and direct question regarding beliefs about whether women and men are equally competent in politics. Substantively, this question is likely to be highly sensitive, in particular due to the under-representation of women in political office worldwide (Inter-parliamentary Union, 2015) and discrimination against women in, among other areas, encouragement to serve in politics, political socialization, and elected and politically appointed office (Fox and Lawless, 2014). Empirically, recent studies suggest that women who run for office are in fact more likely to be competent than men (Lawless and Pearson, 2008) and that those who are elected to office are as or more effective lawmakers than their male counterparts (Jeydel and Taylor, 2003; Volden and Wiseman, 2011; Anzia and Berry, 2011). However, setting aside the empirical question of whether gender differences in competence indeed exist, we can nevertheless presume that those who disagree that women are as competent in politics as men will feel strong pressure to claim otherwise. In the following, I test whether this pressure — as measured by whether respondents misreport their beliefs concerning gender equality — varies along ideological lines.

2.4.2 Data and survey design

To test the hypothesis that those on the left are more likely than those on the right to misreport their beliefs about women's and men's competence in politics, I use survey data from a list experiment and direct question that were embedded in a Canadian post-election study fielded after the 2015 federal election. The study was fielded by the research firm Vox Pop Labs to all

---

10 For example, in a speech regarding college access and affordability, U.S. President Obama said the following:

I’ve heard of some college campuses where they don’t want to have a guest speaker who is too conservative. Or they don’t want to read a book if it has language that is offensive to African Americans, or somehow sends a demeaning signal towards women. And I’ve got to tell you, I don’t agree with that either. I don’t agree that you, when you become students at colleges, have to be coddled and protected from different points of views (Obama, 2015).

11 The reason for this, Anzia and Berry (2011) argue, is that sex-based candidate selection leads to higher quality candidates who are implicitly required to meet higher standards to run for or be elected into political office.
of the members of its national online panel, resulting in an exceptionally large sample size (n = 24,020). Because the survey was sent to all members of the research firm’s panel, the sample was not pre-stratified to match population characteristics. To address differences in sample and population characteristics, survey weights from the data are used to post-stratify estimates.\textsuperscript{12}

To gauge attitudes toward gender equality in politics, respondents were asked whether they agreed or disagreed with the statement “women are as competent as men in politics.” Respondents were first asked to provide an answer regarding this statement through a list experiment question and later in the survey asked to provide an answer openly in response to a direct question.

The list experiment question included four control items and was asked as follows:

How many of the following do you agree with?

- There should be more funding for the arts
- The government should raise taxes on gasoline
- Corporations are taxed too much
- Unions have too much power

Respondents in the treatment group received the same question and a list that included both the control items and the sensitive item:

How many of the following do you agree with?

- There should be more funding for the arts
- The government should raise taxes on gasoline
- Corporations are taxed too much
- Women are as competent as men in politics
- Unions have too much power

The direct question was asked later in the survey as follows:

Do you agree or disagree with the following statement?

- Women are as competent as men in politics

The list experiment and direct question were separated by a large set of unrelated questions to obviate potential response bias, whereby the direct question response is not independent of treatment status.

\textsuperscript{12}Weighting variables include gender, age, education, mother tongue, region, and partisanship.
Because list experiments are cognitively taxing, a question was included in the survey to filter out inattentive respondents. Rather than screen for inattentiveness in the survey as a whole (Berinsky, Margolis and Sances, 2014), a question was included to address this concern for the list experiment directly. The screener question was displayed immediately after the list experiment question and asked respondents to identify, among a new list, the topic of one of the control items from the previous page (see online appendix for complete question text). Seven percent of respondents either did not answer this question correctly or answered “don’t know.” These respondents were dropped from the dataset analyzed in the results section.

Before analyzing the results, a series of tests were run on the data to check for violations of the list experiment’s design assumptions. These tests do not find strong evidence of any violations, the results of which are detailed in the online appendix.

### 2.4.3 Results

I begin by examining the data descriptively. Figure 2.4 presents the estimated proportions of agreement with the statement “women are as competent as men in politics” for data from the list experiment and direct questions. As the figure shows, the proportion of the population that agree with the statement as estimated using data from the list experiment (81.8%) is substantially lower than that as estimated using data from the direct question (96.6%). Put differently, although a relatively large proportion of the population does not believe that women are as competent as men in politics (18.2%), only a small fraction of the population appears willing to state so openly (3.4%). The overall strength of the social norm prohibiting the expression of beliefs inconsistent with gender equality, in other words, is substantial.
To examine the predictors of agreement with the statement and to test the hypothesis that those on the left side of the ideological spectrum are more likely than those on the right to misreport the belief concerning gender equality in politics, I fit models using both the proposed and standard estimators to the data. The variable of interest, ideology, is measured by self-placement on a political ideology scale with response categories 0 through 10, where 0 represents right-wing and 10 represents left-wing. In both models, a set of socio-demographic controls are included as covariates. Results from these models are presented in Table 2.4.

I begin by examining the sensitive-item sub-models in Models 1 and 2. These models present coefficients for the predictors of a belief that women are as competent as men in politics. As these models show, there is strong evidence that the more individuals are to the left politically, the more likely they are to believe that men and women are equally competent (p < 0.001). Being a woman, being younger, and having a university education are also positively associated with believing that women and men are equally competent in politics.

To demonstrate the relationship with political ideology graphically, I simulate coefficients from Model 2 (King, Tomz and Wittenberg, 2000) and calculate the predicted probabilities of agreement with the statement for each value on the ideology scale, holding values of the control variables at their observed values (Hanmer and Kalkan, 2013). Results from this procedure are presented in the left panel of Figure 2.5. As the figure shows, the predicted probability of a belief that women are as competent as men in politics increases the more one is to the left politically. For those on the far right (ideology = 0), the predicted probability of holding this belief is 0.76 on average; for those on the far left (ideology = 10), this probability is 0.96.

I now address the hypothesis that among those who do not believe in gender equality, those on the left are more likely than those on the right to misreport it. To do so, I examine the results from the misreport sub-model of Model 2 in Table 2.4. This sub-model presents the parameters associated with providing one response to the sensitive item in the list experiment, but responding otherwise when asked to reveal that response openly to the direct question. As the results from this model show, there is strong evidence (p < 0.001) that being on the left side of the ideological spectrum is positively associated with misreporting one's belief about gender equality in politics.

To demonstrate this relationship graphically, I simulate coefficients from the model and calculate predicted probabilities of misreporting across the range of the ideological self-placement variable, holding values of the control variables at their observed values. Results from this procedure are shown in the right panel of Figure 2.5. As the graph shows, the probability of

---

13Question wording for the ideological self-placement scale is as follows: In politics people sometimes talk of left and right. Where would you place yourself on the scale below, where 0 is left and 10 is right? To ease interpretation in the discussion of the results, the values of this variable were reversed through recoding so that high values indicate self-placement on the left side of the ideological spectrum.
misreporting increases steeply the more one is to the left side of the ideological spectrum. For those on the far right (ideology = 0), the predicted probability of misreporting one's belief that women and men are not equally competent in politics is 0.58 on average; for those on the far left (ideology = 10), however, this probability is 0.88, a 29 percentage point difference.

To summarize, although the evidence strongly suggests that those on the left are more likely to agree that women are as competent as men in politics, among those who do not believe this,
Figure 2.5: Predicted probabilities of support for and misreporting of a belief that “women are as competent as men in politics”

The left panel shows predicted probabilities of support for the statement “Women are as competent as men in politics” across the range of the ideological self-placement variable. The right panel shows the predicted probabilities of misreporting this belief among those who believe that women are not as competent as men. 95% prediction intervals are shown in gray, as calculated from 5,000 Monte Carlo simulations.

those on the left are also substantially more likely to misreport it.

2.5 Conclusion

Indirect measurement techniques have recently gained substantial attention from researchers and are being applied with increasing frequency to address a wide variety of research questions. This research has been greatly strengthened by the recent development of a research program to develop methodological tools to enable responses to indirect questions to be investigated with much greater sophistication than was previously possible (Gingerich, 2010; Bullock, Green and Ha, 2010; Imai, 2011; Blair and Imai, 2012; Blair, Imai and Lyall, 2014; Imai, Park and Greene, 2015; Blair, Imai and Zhou, 2015). These methodological advances have focused on the development of methods to permit researchers to model responses to sensitive statements, to test research design assumptions, and to increase statistical efficiency. Misreporting on sensitive questions, however, has typically been treated as a nuisance or bias in need of elimination. This article highlights the substantive and methodological importance of turning the problem of measurement error due to misreporting on its head, focusing attention on misreporting as a methodologically and substantively important object of interest in and of itself. As demonstrated through empirical application, the method proposed in this article has the potential to answer research questions that are of considerable political relevance and scientific interest. The method thus provides an important and useful new tool for the analysis of a critical but under-analyzed
dimension of research into sensitive beliefs, attitudes, and behaviors.
Chapter 3

How Should We Measure Citizen Ideology with Ideological Placement Scales?

3.1 Introduction

How can we compare the ideology of political actors and ordinary citizens on the same scale? This question is fundamental to our understanding of electoral politics and democratic representation, but is challenging due to the methodological difficulty of measuring politicians, parties, institutions, and citizens on a common ideological scale (Jessee, 2016).

Recently, research that addresses this measurement problem has grown substantially. The methodological strategies are twofold. First, researchers can use survey questions fielded to citizens that can be matched to legislators’ roll-call votes or to survey responses from legislators themselves. Jessee (2009), for example, examines spatial voting using survey questions that ask respondents whether they agree with policies voted on by legislators, and Jessee and Malhotra (2013) scale citizens and Supreme Court Justices by asking respondents how they would decide a set of cases adjudicated by the court (see also, Bafumi and Herron, 2010; Malhotra and Jessee, 2014). Shor and McCarty (2011), on the other hand, use surveys of citizens containing questions that have been answered by legislators themselves in a separate survey (see also, Shor and Rogowski, Forthcoming).1

A more straightforward and accessible approach, however, is to use survey data containing questions that ask respondents for perceptions of their own ideology and that of political

---

1The survey data from legislators that are used by Shor and McCarty (2011) are from the Project Vote Smart National Political Awareness Test (NPAT). These data are very useful for researchers, but the survey’s response rate has unfortunately fallen over time.
actors. These data have the virtue of simplicity in their collection and are available in major election studies such as the American National Election Study (ANES), Cooperative Congressional Election Study (CCES), and European Election Study (EES). The drawback to these data, however, is that because they rely on citizen's perceptions of ideology, they can suffer from substantial perceptual biases. On a typical ideological placement scale, what might be considered “liberal” or “conservative” for one respondent might mean something much different for another (Brady, 1985).

To overcome this problem, Aldrich and McKelvey (1977) proposed, in their classic article, a statistical model to adjust each respondents’ ideological placements to estimate the ideology of citizens, politicians, parties, and institutions on a common scale. With the growth in large-scale national and multi-national surveys, Aldrich-McKelvey (AM) scaling is now being increasingly put to good use to examine differences between citizens and political representatives. Researchers have used AM scaling to compare differences in the ideology of political actors across countries (Saiegh, 2009; Lo, Proksch and Gschwend, 2014; Bakker et al., 2014; Saiegh, 2015); to examine how well the perceived ideology of state- and district-level political representatives align with the ideology of the citizens they represent (Hollibaugh, Rothenberg and Rulison, 2013; Ramey, 2016); to investigate why perceptions of self-placed ideology differ between individuals (Hare et al., 2015); and to compare to measures of ideology from alternative sources (Tausanovitch and Warshaw, 2017). Hare et al. (2015), furthermore, implement AM scaling in a Bayesian framework, which allows for missing data to be incorporated with ease, and uncertainty in all parameters to be calculated as part of the estimation process.²

As this article demonstrates, however, we can substantially improve the accuracy and credibility of AM-scaled ideological estimates of citizens and political actors. The model proposed in this article does so in two ways. First, the model explicitly accounts for the ordinality of typical ideological placement survey questions.³ In most well-known election surveys, such as the ANES and recent versions of the CCES, political ideology is asked on a seven-point likert scale, ranging from “Extremely liberal” to “Extremely conservative.”⁴ Because methods such as AM scaling models are motivated by a desire to correct for perceptual biases, they arguably should also account for the fact that response scales will be treated by respondents as ordinal, with differences between categories perceived to vary from category to category. As this article demonstrates, accounting for the ordinality of the data is particularly important in estimating the ideology of respondents and candidates who are at the extreme ends of the ideological scale.

²A further benefit to a Bayesian approach is that, with the availability of flexible and powerful Bayesian inference engines such as Stan (Carpenter et al., 2017) and JAGS (Plummer, 2016), one can further minimize error in parameter estimates through regularization by fitting AM scaling models with multilevel components.
³King et al. (2004) previously has noted this drawback to existing AM scaling models. With the growth in computational power, these models have become increasingly easy for applied researchers to fit.
⁴The European Election Study, by contrast, measures ideology on an 11-point scale, a more fine-grained, but nonetheless ordinal scale.
Perceptions of the Republican Party as “Extremely conservative” rather than “Conservative” can say something substantially different about one’s perceptions of ideology than perceiving the Republican Party as “Conservative” rather than “Somewhat conservative.”

Second, the model permits incorporation of ideological self-placement into the estimation framework itself. Unlike in the standard AM scaling model set up, in which respondent ideology is calculated in a post-estimation step, the method that I propose estimates respondent ideology by treating it as a latent variable to be estimated simultaneously with the ideological positions of political actors. As shown using large-scale survey data, this can drastically improve estimates of the ideology for some respondents, and substantially improve it for others.

3.2 AM Scaling and Differential Item Functioning

The aim of AM scaling is to address the fact that response scales on ideological placement scales are frequently perceived differently between individuals (Aldrich and McKelvey, 1977; Brady, 1985; Palfrey and Poole, 1987). On a typical ideological scale, what is considered “Conservative” for one respondent might be considered “Moderate” for another. For example, while someone on the ideological left might consider the Republican Party to be extremely conservative, someone on the right might consider the party to be moderate. Thus, although there is a ‘true’ ideological position for the party, whether that position should be considered ideologically moderate, extreme, or otherwise can differ greatly between individuals. Differences in perceptions on questions such as this belong to a family of measurement error problems called differential item functioning (DIF) (King et al., 2004; King and Wand, 2007) and can pose a significant problem for survey research when respondents hold substantially different understandings of a response scale.

To demonstrate this problem and its consequences more concretely, suppose that we observe the ideological placements given by two respondents, both for themselves (self placement) and for the Democratic and Republican parties. Now suppose that each respondent indicates that he or she is an ideological moderate, as shown graphically on the first row of Figure 3.1. The problem that Aldrich and McKelvey (1977) highlight is that although these ideological self-placements are observationally equivalent, the respondents’ ‘true’ positions may nonetheless differ due to differences in their understandings of the meaning of the label moderate (and ‘liberal’ and ‘conservative’). The goal of AM scaling is to correct for these perceptual differences to render these placements and those of the parties on a common ideological scale.

The basic idea motivating AM scaling is that respondents’ perceptions of the ideological scale can be inferred by the ideological placement of the same items, such as political parties and candidates that are placed by some or all respondents and act as ‘bridges’ across respondents. These shared items are important because each item has a unique ‘true’ position, but which is
Figure 3.1: AM scaling stylized example

This figure presents a stylized example to show how two respondents who both self-identify as political moderates — as shown on Row (1) — can have much different political ideologies due to interpersonal variation in perceptions of the ideological scale. Row (2) shows how the respondents vary in their classification of the Democratic and Republican parties. To place respondents on a common scale, Row (3) shifts each scale to have the same center/location and Row (4) stretches/shrinks each scale to have the same degree of party polarization. In practice, AM scaling accounts for the placement of many items, as well as error in item placement.

shifted left or right and scaled outward or inward by respondents based on their understanding of the scale. For example, as shown on the second row of Figure 3.1, it is clear that each respondent perceives the ideology of the Democratic and Republican Parties at different places on the scale: the first respondent perceives the parties as both farther left and more polarized than the second. The aim of AM scaling is to shift the placements of each respondent’s items to have the same location (row 3) by adding a respondent-specific constant to each item’s position. This has the effect of shifting Respondent 1’s placements shown on row 2 to the right, and shifting Respondent 2’s placement shown on row 2 to the left (shifted positions shown on row 3). AM scaling also accounts for respondents who see more or less polarization in the political landscape by stretching ideological placements inward or outward. This has the effect of shrinking inward Respondent 1’s placements on row 3 and stretching outward Respondent 2’s placements shown on row 3 (stretched/shrunk positions shown on row 4). As a consequence, even though in our example each respondent placed themselves as moderate on the original scale, we can see that each respondent’s ideology is substantially different from each other on the common ideological scale.
3.3 Statistical model

In the two-item case of our stylized example, these calculations are trivial. In general, however, respondents will typically place multiple items in a survey and do so with error. The aim of AM scaling is to model this process to estimate the relative latent ideological positions of each item and in a post-estimation step to calculate the ideological positions of respondents relative to each item and to each other. The result is that all items and respondents are placed on a common ideological scale that can be used to make valid comparisons between individuals and political actors.

In the standard set up for an AM scaling model, survey respondents \( i = 1, \ldots, N \) are asked to place items \( j = 1, \ldots, J \) on an ideological scale, where the response categories typically take the form of likert response options, ranging from, for example, “Extremely liberal” to “Extremely conservative.” The response to each item, \( y_{ij} \), is then modeled as a linear transformation of the ‘true’ position of an item, \( \zeta_j \), through a respondent-specific shift parameter \( \alpha_i \) and a respondent-specific scale parameter \( \beta_i \):

\[
y_{ij} \sim \text{Normal}(\alpha_i + \beta_i \zeta_j, \sigma_j),
\]

where \( \sigma_j \) denotes error in the placement of item \( j \). The likelihood function is therefore:

\[
L(y; \alpha, \beta, \zeta, \sigma) = \prod_{i=1}^{N} \prod_{j=1}^{J} \text{Normal}(y_{ij} | \alpha_i + \beta_i \zeta_j, \sigma_j).
\]

In a Bayesian framework, this basic model is straightforward to fit by placing reasonable priors on the model parameters, and has the benefit that uncertainty in the parameters is calculated as part of the estimation process (see Hare et al., 2015). Finally, following model fitting, one can generate estimates of respondent ideology through the following calculation:

\[
\hat{\theta}_i = \frac{y_{i,\text{self}} - \alpha_i}{\beta_i},
\]

where \( y_{i,\text{self}} \) is the ideological self-placement of respondent \( i \). This calculation ‘reverses’ the data generating process by shifting and scaling respondents’ perceived positions onto a common scale (as shown graphically on rows 3 and 4 of Figure 3.1).

Estimates of the ideological positions of parties and candidates using this model have been

\[5\] In the ANES (since 1972) and CCES (since 2010), response categories given to respondents are “(1) Extremely liberal,” “(2) Liberal,” “(3) Somewhat liberal,” “(4) Moderate,” “(5) Somewhat conservative,” “(6) Conservative,” and “(7) Extremely conservative.”

\[6\] Hare et al. (2015) extend this model further by allowing error in item placement to vary at the individual level, such that \( y_{ij} \sim \text{Normal}(\alpha_i + \beta_i \zeta_j, \sigma_i \sigma_j) \). For simplicity, this extension is avoided here, but is relatively straightforward to include in the model.
shown to have strong convergent validity with ideology estimated both through expert surveys (Lo, Proksch and Gschwend, 2014) and congressional roll-call data (e.g. Hare et al., 2015; Ramey, 2016). Recent extensions of this model have sought to permit the use of ideological placement data that span different geographies to compare differences in a (common) ideological landscape across countries (e.g. Bakker et al., 2014; Lo, Proksch and Gschwend, 2014; Hare et al., 2015; Ramey, 2016).

Despite these promising results, researchers have yet to account for (1) the fact that typical ideological placement scales are ordinal, and (2) the fact that ideological self-placement has not been included as an outcome as part of model fitting to improve parameter estimates. To address this, I introduce an ordinal AM scaling model that incorporates observed respondent self-placement in the model and models respondents’ ideology on a common scale as a latent variable.

Specifically, I propose a heteroskedastic ordinal logistic AM scaling model that estimates separate scale parameters for each item. The model is defined as follows:

\[
\begin{align*}
\Pr(y_{ij} = 1) &= \text{logit}^{-1}\left(\frac{\kappa_1 - \alpha_i - \beta_i \zeta_j}{\sigma_j}\right) \\
\Pr(y_{ij} = k) &= \text{logit}^{-1}\left(\frac{\kappa_k - \alpha_i - \beta_i \zeta_j}{\sigma_j}\right) - \text{logit}^{-1}\left(\frac{\kappa_{k-1} - \alpha_i - \beta_i \zeta_j}{\sigma_j}\right) \\
&\vdots \\
\Pr(y_{ij} = K) &= 1 - \text{logit}^{-1}\left(\frac{\kappa_K - \alpha_i - \beta_i \zeta_j}{\sigma_j}\right),
\end{align*}
\]

where \( k \in \{1, \ldots, K\} \) denotes the \( K \) response categories to the ideological placement scale; \( \kappa_k \) denotes each of the \( K - 1 \) cut points; \( \text{logit}^{-1}(\cdot) \) denotes the inverse logistic function; and \( \sigma_j \) denotes the scale parameter. Note that the scale parameters \( \sigma_j \) vary by item to account for heterogeneity in the accuracy of respondents’ placement of political actors. This is theoretically informed by the assumption that respondents are likely to have difficulty placing some items more than others, such as candidates who are not well known. These scale parameters are analogous to the parameters \( \sigma_j \) in the linear AM scaling model, as shown in Equation 3.1. If these scale parameters were set to be constant across items, the model would reduce to the canonical form of the ordered logistic regression model.\(^7\)\(^8\)

To incorporate respondents’ ideological self-placement into the model, I model perceived

\(^7\)As in the canonical ordinal regression model set up, the model here is defined such that the relationship between \( \zeta_j \) and \( Y_{ij} \) is determined by \( \beta_i \), and does not vary with the response category \( k \) (parallel regression assumption). A more complex model may be possible, but is heavily limited by the fact that the \( \beta_i \) already vary at the individual level, with little data to estimate many parameters per respondent.

\(^8\)As I show in the supplementary material, allowing the scale parameters to vary by item changes little in the estimated ideology of respondents and items, and strengthens a similar finding for the linear model demonstrated by Palfrey and Poole (1987). Thus, although the benefit of allowing the scale to vary per item is theoretically satisfying as a better approximation to the data-generating process, the empirical benefits of this more complicated and more computationally demanding ordered model are relatively minor.
self-placement, $y_{i,\text{self}}$, as a function of the scale and shift parameters $\alpha_i$ and $\beta_i$, and a latent variable $\theta_i$ which denotes the unobserved position of each respondent in common ideological space:

$$
\begin{align*}
\Pr(y_{i,\text{self}} = 1) &= \logit^{-1}\left(\frac{\kappa_1 - \alpha_i - \beta_i \theta_i}{\sigma_{\text{self}}}\right) \\
\Pr(y_{i,\text{self}} = k) &= \logit^{-1}\left(\frac{\kappa_k - \alpha_i - \beta_i \theta_i}{\sigma_{\text{self}}}\right) - \logit^{-1}\left(\frac{\kappa_{k-1} - \alpha_i - \beta_i \theta_i}{\sigma_{\text{self}}}\right) \\
&\vdots \\
\Pr(y_{i,\text{self}} = K) &= 1 - \logit^{-1}\left(\frac{\kappa_{K-1} - \alpha_i - \beta_i \theta_i}{\sigma_{\text{self}}}\right).
\end{align*}
$$

(3.5)

For estimation in a Bayesian framework, I place priors on each group of parameters, and set constraints where necessary to identify the model. In particular, the shift parameters $\alpha$, stretch parameters $\beta$, and latent respondent ideology $\theta$, are each given common distributions respectively: $\alpha_i \sim \text{Normal}(0, \sigma_\alpha)$, $\beta_i \sim \text{Normal}(\mu_\beta, \sigma_\beta)$, $\theta_i \sim \text{Normal}(\mu_\theta, \sigma_\theta)$, where uniform priors are placed on the hyperparameters $\mu_{(\cdot)}$ and $\sigma_{(\cdot)}$. The item parameters $\zeta_j$ are given weakly informative priors, $\zeta_j \sim \text{Normal}(0, 10)$, and the scale parameters $\sigma_j$ are given uniform priors.10

To identify the model, we need to address the problem of reflection invariance (Bafumi et al., 2005), which results from the fact that the likelihood is invariant to multiplication of the parameters $\beta$ and $\zeta$ by -1, resulting in a bimodal posterior. In other words, we need to fix the direction of the scale such that high values indicate either “Liberal” or “Conservative.” Although there are a number of ways to permit model identification (see Bafumi et al., 2005, 176-178), I opt to fix two items at the values $-1$ and $+1$ (Clinton, Jackman and Rivers, 2004).11 Note that because the AM scaling problem aims to identify a common ideological space that is relative, the choice of item parameters to fix can be made arbitrarily. To make summary of the results easier in the empirical section, these constraints are applied to the positions of the Democratic and Republican Parties, which are set to $-1$ and $+1$ respectively.

Lastly, we must constrain the scale parameters for one of the items and for respondent self-placement. The need for the former constraint is typical of ordered logistic regression models, where the scale parameter is typically set to 1. The latter necessity results from the fact that there is a single self-placement per respondent, which prevents the error in self-placement from being estimated. To identify the model, I fix the scale of the self-placement scale $\sigma_{\text{self}}$ to be equivalent to that of the Democratic and Republican parties, under the assumption that respondents will place these three actors in relative space with (roughly) equivalent error, such that $\sigma_{\text{self}} = \sigma_{\text{Dem}} = \sigma_{\text{Repub}} = 1$. The scale parameters for all other items are estimated from the data relative to this constraint.

---

9 All models in this letter are fit using the Bayesian inference engine Stan (Carpenter et al., 2017).
10 Specifically, the uniform priors used are improper uniform($-\infty, \infty$).
11 We could also restrict the signs of these parameters to be of different signs, or constrain one to be greater than the other. In practice, however, when the total number of items is large, identifying the model this way can occasionally lead to problems in sampling.
3.4 Empirical Comparison and Application

Before examining the properties of the proposed ordinal model, I begin by comparing ideology estimates from two linear AM scaling models, one in which ideological self-placement is calculated in a post-estimation step as in the typical AM scaling model set up, and one in which ideology is estimated as a latent variable simultaneously with other model parameters, as described above. This comparison permits us to examine differences in estimates that result from the inclusion of self-placement as a component of the standard AM scaling model. To investigate this, I use replication data from Hare et al. (2015), who first introduced the Bayesian AM scaling model and fit a linear model to data from the 2004, 2008, and 2012 American National Election Studies. Because the model proposed by Hare et al. (2015) is not perfectly comparable to the proposed model due to slight differences in model specification, I fit a comparable linear AM scaling model that is equivalent to the proposed model, but which does not incorporate information from respondent self-placement. As demonstrated in the Supplementary Material, the ideology estimates provided in the replication data by Hare et al. (2015) also show a substantively equivalent pattern to the results presented below.

To demonstrate differences between the models, Figure 3.2 presents histograms of respondents’ estimated ideology, using the median as a point estimate for each respondent. As is clear from the figure, estimates from the proposed model, which allows ideological self-placement to inform the model parameters, appears to have higher face validity, with estimates confined to a much more reasonable latent ideological space than the model that does not. This is expected because, as we can see in Equation 3.3, values of $\beta_i$ that approach zero lead to extreme values of $\theta_i$. Without additional information from ideological self-placement to inform the model parameters, these estimates can be extreme. In introducing the Bayesian AM scaling model, Hare et al. (2015) note this problem explicitly, and use the median rather than the mean for their point estimates to avoid these estimates being heavily affected by outliers. As we can see in the third column of Figure 3.2, the models provide somewhat similar estimates in the center of the distribution, but these estimates diverge greatly in the tails.

Furthermore, when examining draws from the posterior itself rather than a summary statistic such as the median for point estimates, ideology estimates from the model that does not incorporate self-placement are unreasonable on their face, with values of $\theta_i$ for some respondents in the hundreds and thousands. In the model that incorporates self-placement, by contrast, estimates are confined to a range of roughly $-4$ to $+4$. This is important because to examine the distribution of ideology in the population, one is typically not interested in the distribution of

---

12Specifically, Hare et al. (2015) allow for error in item placement to vary at the individual level, a more complex model that I do not pursue here. The authors also use slightly different priors and constraints for identification.

13As they write, “in [Equation 3.3] that $\lim_{\beta_i \to 0} = \infty$, so we use the median value as the point estimate of the respondent positions since the median will be more robust to long tails produced when posterior draws of $\beta_i$ are very close to zero.
Figure 3.2: Comparison of AM Models including and excluding self-placement

Excl. self-placement | Incl. self-placement | Comparison

2012

2008

2004

point estimates for respondents, but in the posterior itself.

To further demonstrate the difference in these models, I examine how well estimates from each model predict Presidential vote choice in ANES data from 2004, 2008, and 2012. I fit logistic regression models in which the outcome is voting for the Republican Presidential candidate in each of the three elections, and latent ideology is the sole predictor. I then calculate Brier scores to measure predictive performance. The ordinal AM scaling model is also included for completeness. I examine the ordinal model in greater detail in the section that follows. Results from these calculations are presented in Table 3.1. As we can see, for each dataset, the models that include self-placement outperform both the raw ideological self-placement response given by respondents and those from the models that calculate respondent ideology in
Table 3.1: Predictive performance of AM Scaling Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw self-placement</td>
<td></td>
<td>0.149</td>
<td>0.157</td>
<td>0.144</td>
<td>0.092</td>
<td>0.097</td>
<td>0.112</td>
<td>0.120</td>
<td></td>
</tr>
<tr>
<td>Hare et al. (2015)</td>
<td></td>
<td>0.139</td>
<td>0.154</td>
<td>0.103</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>✓</td>
<td>0.145</td>
<td>0.148</td>
<td>0.113</td>
<td>0.067</td>
<td>0.064</td>
<td>0.090</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>Ordinal</td>
<td>✓</td>
<td>0.124</td>
<td>0.125</td>
<td>0.098</td>
<td>0.061</td>
<td>0.055</td>
<td>0.070</td>
<td>0.074</td>
<td></td>
</tr>
</tbody>
</table>

A post-estimation step. The AM scaling model that incorporates self-placement, in other words, appears to have both much stronger face validity and stronger predictive performance.

3.4.1 Political ideology during the mid-term and Presidential elections

Having shown the benefits of including ideological self-placement in the standard AM scaling model, I now examine the differences between the linear and proposed ordinal AM scaling models by applying them to data from the Cooperative Congressional Election Study (CCES). CCES data are useful for researchers who study political ideology because they are both large-scale and contain responses from respondents regarding a variety of sub-national actors. In particular, respondents are asked about the senators and governor of their state and the house representative in their congressional district. Respondents therefore do not ideologically place every congressional representative. However, because all respondents place the same national political actors, these items act as ideological ‘bridges’ to allow sub-national actors to be mapped on to a common scale. As a result, researchers have used AM scaling to examine the relationship between the political ideology of citizens and their congressional and gubernatorial representatives (e.g. Hollibaugh, Rothenberg and Rulison, 2013; Hare et al., 2015; Ramey, 2016).\(^ \text{14} \)

To examine political ideology during recent congressional and presidential election years, I fit a linear AM scaling model and the proposed ordinal AM scaling model\(^ \text{15} \) to each separate CCES data set from 2010, 2012, 2014, and 2016, where ideological placement is measured on a 7-point likert scale.\(^ \text{16} \) The items included in the models are those for all political actors that are placed by all respondents (e.g. national political parties and leaders) and all senators and governors serving at the time of the survey. Details regarding the items included in the model

\(^ {14} \) The American National Election Study, by contrast, asks primarily about a small number of national candidates and parties.

\(^ {15} \) In both models, ideological self-placement is incorporated as described in the Statistical Model section.

\(^ {16} \) The response categories provided to respondents in the CCES are “(1) Extremely liberal,” “(2) Liberal,” “(3) Somewhat liberal,” “(4) Moderate,” “(5) Somewhat conservative,” “(6) Conservative,” and “(7) Extremely conservative.”
Figure 3.3: Comparison of linear and ordinal AM scaling

To demonstrate differences in ideology as estimated by the linear and ordinal AM scaling models, Figure 3.3 shows the distribution of respondent ideology (median point estimates) for each election year. There are two notable differences in these distributions. First, the distribution of ideology as estimated from the ordinal model is substantially wider in each year compared to estimates from the linear model. The ordinal model, in other words, suggests that citizens perceive their political ideology in more extreme ways relative to the Democratic and Republican parties than would be concluded if we assumed that respondents perceive the

---

17 These samples include respondents who did not provide a response to the question regarding ideological self-placement: although no measure of common-scale ideology can be estimated for such respondents, their placement of political actors nevertheless aid in estimation and therefore need not be excluded.

18 The Democratic and Republican Party positions are fixed in each model at −1 and +1 respectively for model identification, as noted in the Statistical Model section.
ideological placement scale linearly. Second, this difference appears to be asymmetric during the 2010, 2012, and 2014 elections, in which those on the right appear more conservative in the ordinal model relative to the Republican and Democratic Parties. This is a particularly important characteristic for the 2010 election, which followed the emergence of the Tea Party movement, and during which we should expect conservatives were inclined to see themselves as distant from the Republican Party than in previous years. The ordinal model, in this case, thus captures important features of ideology that are masked by the simpler linear model.

If the proposed ordinal AM scaling model suggests that the U.S. electorate is more polarized than it appears under the more restrictive linear model, what about the landscape of political actors? To answer this question, I examine the ideological placement of senators during the 2010, 2012, 2014, and 2016 elections. Figure 3.4 presents point estimates of the ideology of each senator as estimated by the linear and ordinal AM scaling models. Each panel makes clear that

---

Figure 3.4: AM Scaled Senator Ideology (CCES)

---

42
although there are few differences in ideology for the vast majority of senators in the middle of the distribution, there are substantial differences in perceptions of senators on the ideological extremes. These include, for example, well-known conservative Republicans Jim DeMint, Tom Coburn, Mike Lee, and Ted Cruz. But on the left, Bernie Sanders is consistently perceived as considerably more extreme than his senate colleagues in each of the four Presidential and mid-term election years examined. For example, with the positions of the Democratic and Republican parties fixed, Sanders’ ideology in 2016 is estimated to be -1.54 under the linear model, but a substantially more extreme -2.04 under the ordinal model.

3.4.2 Was Bernie Sanders ideologically out of step during the 2016 Presidential campaign?

That Bernie Sanders is estimated to be even more extreme in the proposed ordinal model raises questions about his ideological appeal as a national candidate during the 2016 Presidential election. Despite losing to Hillary Clinton in the 2016 Democratic primary elections, Sanders proved to be an unexpectedly strong challenger. Indeed, in the aftermath of Clinton’s subsequent defeat to Donald Trump in the Presidential election, some have suggested that Sanders would have been a more formidable presidential candidate than Clinton in the national context (Bump, 2016; Flegenheimer and Alcindor, 2016; Budowsky, 2017).

Investigating this counter-factual in its entirety is a difficult challenge. However, we can help shed light on one potentially major hurdle faced by Sanders by using AM scaling to examine the ideology of the U.S. electorate relative to Sanders and to the parties’ presidential nominees as measured on a common ideological scale. We can ask, in other words, whether Bernie Sanders was ideologically out of step with the national electorate during the 2016 Presidential campaign, and, if so, to what degree.

To investigate this, I examine the distribution of ideology among members of the electorate and compare it to the estimated ideological placement of Sanders, the parties, and the Presidential candidates. To show the posterior graphically, Figure 3.5 presents overlapping densities for ten draws from the posterior of respondent ideology, which allows us to capture uncertainty in the distribution of voter ideology during each election. The panels in the first column of Figure 3.5 show the distribution of ideology among all respondents who indicated that they voted in the given election year; panels in the second column show the distribution of ideology among all respondents who voted for the Republican or Democratic Presidential candidate in the most recent Presidential election. In the 2010 and 2014 mid-term election years, the distributions of voters shown are, respectively, those who voted for Obama or McCain in the 2008 election, and those who voted for Obama or Romney in the 2012 election. The Democratic and Republican parties are shown at −1 and +1, which, recall, were set as constraints for model identification. Lastly, the labeled vertical lines show the median placement of each political
As we can see in the panels in the first column in Figure 3.5, Obama, Romney, Clinton, and Trump each are estimated to be relatively close to the ideology of their respective political parties. Whereas Obama is estimated to be slightly ideologically to the left of the Democratic Party in each year, Romney, Clinton, Trump are perceived as more moderate than their respective parties. This is also the case, as we can see in the second column of Figure 3.5, in comparison to Democratic and Republican voters, where the President and Presidential candidates each are estimated to be close to the median Democratic and Republican voter. By contrast, it is clear that Bernie Sanders is substantially to the ideological left of each party and to the left of the Democratic voters in each of the four election years shown.

To quantify these differences, I estimate the proportion of all voters and the proportion of Democratic and Republican voters who are more ideologically liberal in each election year than each candidate and party. Results are presented in Table 3.2. The results demonstrate that Bernie Sanders was substantially more ideologically liberal than the vast majority of voters and the vast majority of those who have voted for the Democratic Presidential candidate. Whereas 26% of voters in 2016 were more ideologically liberal than Clinton and 69% of voters more liberal than Trump, only 6% were more liberal than Sanders. Among co-partisans, the results are similarly stark. Among Democratic voters, 41% were more liberal than Clinton, but only 10% more liberal Sanders. As Table 3.2 shows, furthermore, these results are consistent across election years. We cannot necessarily conclude from these results that Sanders’ chances as a Presidential candidate would have been heavily affected by his perceived ideological position. Ahler and Broockman (Forthcoming) make the case, for example, that candidates on the ideological extremes can be appealing to moderate voters. What is clear, however, is that absent a sharp ideological shift to the center, Sanders would have faced an uphill battle against perceptions of his being substantially out of step with the vast majority of voters.

3.5 Conclusion

The development of statistical methods to measure the political ideology of citizens, politicians, parties, and institutions on a common scale has helped advance our understanding of democratic representation, the ideological fit of candidates to their constituents, and the contours of the ideological landscape more generally. Some researchers have sought to develop and apply methods to data from constituents whose responses can be matched to analogous data from political representatives (e.g. Jessee, 2009; Bafumi and Herron, 2010; Shor and McCarty, 2011; Jessee and Malhotra, 2013; Malhotra and Jessee, 2014; Shor and Rogowski, Forthcoming). These are impressive contributions, but the data they often rely on are generally difficult to

---

19 This includes in 2016, in which Obama’s position (-1.04) is not shown.
Figure 3.5: Posterior of political ideology for voters and political actors

**All Voters**

2010

Sanders

2012

Sanders

2014

Sanders

2016

Sanders

**Democratic and Republican Voters**

2010

Obama

2012

Obama

Romney

2014

Obama

2016

Clinton

Trump
Table 3.2: Proportion of voters and co-partisans who are more liberal than candidates and parties

<table>
<thead>
<tr>
<th></th>
<th>CCES 2010</th>
<th></th>
<th>CCES 2012</th>
<th></th>
<th>CCES 2014</th>
<th></th>
<th>CCES 2016</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Co-partisans</td>
<td>All</td>
<td>Co-partisans</td>
<td>All</td>
<td>Co-partisans</td>
<td>All</td>
<td>Co-partisans</td>
</tr>
<tr>
<td>Bernie Sanders</td>
<td>6</td>
<td>11</td>
<td>7</td>
<td>13</td>
<td>8</td>
<td>16</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Hillary Clinton</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>21</td>
<td>41</td>
<td>26</td>
<td>41</td>
</tr>
<tr>
<td>Donald Trump</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>69</td>
<td>43</td>
</tr>
<tr>
<td>Barack Obama</td>
<td>18</td>
<td>36</td>
<td>22</td>
<td>39</td>
<td>19</td>
<td>37</td>
<td>21</td>
<td>38</td>
</tr>
<tr>
<td>Mitt Romney</td>
<td>73</td>
<td>48</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Democratic Party</td>
<td>20</td>
<td>39</td>
<td>22</td>
<td>40</td>
<td>21</td>
<td>40</td>
<td>22</td>
<td>40</td>
</tr>
<tr>
<td>Republican Party</td>
<td>65</td>
<td>41</td>
<td>75</td>
<td>51</td>
<td>69</td>
<td>47</td>
<td>75</td>
<td>54</td>
</tr>
</tbody>
</table>

collect, are unavailable, or require the development of methods specific to each unique data set in order to be analyzed. There is thus considerable utility in the pursuit of techniques that use standard survey-based ideological placement scales — frequently available in large-scale election studies — to measure the ideology of citizens and political actors on a common scale. Such data are both simple to collect, and, assuming that much of citizens’ political behavior is driven by perceptions of the ideology of political actors rather than their ‘true’ ideological positions, offer the opportunity to examine a wide array of research questions.

20 This article has advanced the use of ideological placement data by proposing a method to account for both the ordinality of standard ideological placement data and the ability of self-placement responses to ground estimates of survey respondents’ latent ideology on a common scale. The method produces estimates that have higher face validity and stronger predictive performance than those from the standard AM scaling model, and reveal important features of the ideological distribution of the electorate and political actors that are masked by the standard model. Furthermore, given the growth in the use and speed of Bayesian inference engines, the model can be estimated with relative ease.

As Jessee (2016, 1123) notes, the decision to use policy-based measures of ideology as estimated from, for example, item-response theory models, or to use a perception-based measure, such as estimated from ideological self-placement scales, should be driven by the research question of interest to researchers. As Hare et al. (2015) suggest, there is also good reason to examine why the ideology of political actors as estimated from roll-call data differ from the ideology of those same actors as perceived by the electorate. NOMINATE estimates of the ideology of senators from roll-call data, for example, do not place Bernie Sanders as the most liberal senator, despite his being perceived, by a wide margin, as the most liberal senator by the electorate.
Chapter 4

Does International Terrorism affect Public Attitudes toward Refugees? Evidence from a Large-scale Natural Experiment

Paris changes everything.

— Markus Söder
Bavarian Finance Minister

4.1 Introduction

On November 13, 2015, nine heavily armed gunmen and suicide bombers affiliated with the terrorist organization Islamic State perpetrated a series of attacks in the heart of Paris, killing 130 civilians and injuring more than 350. The attacks were the largest in the West in more than a decade and they intensified concerns among world leaders and the public about the threat of international terrorism. They also coincided, however, with the Syrian refugee crisis, the largest refugee crisis since the end of the Second World War. Perpetrated by a Syria-based terrorist organization, the Paris attacks would tightly intertwine two of the most salient issues in modern international politics: the fight against terrorism and the large-scale resettlement of refugees.

In the aftermath of the attacks, politicians across Europe and North America sought to mobilize public opposition to refugee resettlement by invoking the perceived threat posed by Syrian refugees to national security. Within a week of the attacks, 30 U.S. governors had voiced opposition to resettlement in their states and the U.S. House of Representatives had passed a
bill to suspend the section of the refugee program concerning Syrian and Iraqi refugees (Healy and Bosman, 2015). In Canada, Prime Minister Justin Trudeau postponed a plan to welcome 25,000 Syrian refugees by the end of the year (Canadian Broadcasting Company, 2015a). In Germany, pressure mounted on Chancellor Angela Merkel to end her government’s open-arms refugee policy. Perceptions of the effects of the attacks on attitudes and policy were neatly summarized by Bavaria’s Finance Minister: “Paris,” he said, “changes everything” (Aust, Malzahn and Vitzhum, 2015).

The Paris attacks and their political aftermath highlight a series of important, but as yet unanswered questions concerning refugees and the effects of large-scale terrorist attacks on attitudes, emotions, and behaviors: What effects does international terrorism have on public attitudes and emotions toward refugees fleeing from countries associated with terrorism? To what degree does terrorism mobilize the public to pressure political representatives to restrict refugee policy? Are these effects long- or short-lived?

To answer these questions, this article presents results from a natural experiment to examine the effects of the Paris terrorist attacks on attitudes and emotions toward Syrian refugees. We use data containing an extensive set of questions concerning Syrian refugees that were asked in a large-scale national survey (n = 18,763) fielded in Canada — a major recipient of Syrian refugees — less than 48 hours before the Paris attacks, and subsequently fielded to large samples of respondents each day for three weeks thereafter. To our knowledge, these data provide the first opportunity to examine, with precision, the effects of major terrorist attacks on attitudes and emotions toward refugees; their effect on issue-based political mobilization; and the duration of these effects.

This article makes four major contributions to the literature. First, by examining the effects of terrorism on public attitudes and emotions toward refugees, we broadly investigate an issue that remains understudied despite its substantial international importance. While there is a growing body of research that investigates the causal effects of terrorist attacks on public opinion, this research has focused heavily on attitudes and behaviors related to voting and partisanship, and on the effects of terrorism in the vicinity of attacks (Bonanno and Jost, 2006; Berrebi and Klor, 2008; Gould and Klor, 2010; Getmansky and Zeitzoff, 2014; Kibris, 2010; Bali, 2007; Montalvo, 2011; Hersh, 2013).1 Little is known, however, about the effects of large-scale attacks on attitudes toward outgroups such as refugees and the effects on publics that are not the direct targets.2 Indeed, large-scale terrorist attacks are frequently aimed at international audiences and can have wide-ranging consequences for policy-making.

Second, we address empirical limitations in the literature on terrorism and public opinion. In contrast to studies that examine real-world terrorist attacks, much research has investigated

---

1One notable recent exception leverages variation in terrorism across time to examine the effects of terrorism on political tolerance in Israel (Peffley, Hutchison and Shamir, 2016).
2Important exceptions include Legewie (2013) and Finseraus and Listhaug (2013), noted further below.
the link between perceptions of the threat of terrorism and an extensive range of attitudes and emotions using cross-sectional surveys and survey experiments (e.g. Davis and Silver, 2004; Huddy et al., 2005; Merolla and Zechmeister, 2009; Hetherington and Suhay, 2011; Malhotra and Popp, 2012; Renshon, Lee and Tingley, 2015). It remains an open question, however, whether empirical relationships demonstrated in these studies carry over to real-world terrorist events. Using a natural experiment and rich survey data, we examine these relationships in the context of large-scale terrorist attacks and investigate their role as causal mechanisms in affecting opposition to refugee resettlement.

Third, although research into real-world attacks has investigated their effect on voter turnout (Hersh, 2013; Robbins, Hunter and Murray, 2013; Getmansky and Zeitzoff, 2014), no research, to our knowledge, has examined more immediate forms of mobilization. We help fill this gap in the literature by examining the effect of the Paris attacks on the public’s willingness to contact political representatives regarding refugees and resettlement. The evidence we provide suggests that growth in vocal opposition to refugee resettlement in the aftermath of the attacks was more likely due to changes in mobilization than it was due to changes in attitudes: although the effect of the attacks on opposition to refugee resettlement was moderate, their effect on political mobilization was substantial and strongly favored the political opponents of resettlement. The upshot is that large-scale terrorist attacks can result in asymmetrical public pressure on political representatives to implement policy changes, even if the underlying distribution of policy preferences do not meaningfully change. This result has important consequences for our understanding of democratic representation.

Lastly, this article is the first to clearly illustrate the short-term dynamics of public opinion in the aftermath of large-scale terrorist attacks. In past research, investigation of real-world terrorist attacks has relied on survey samples pooled across many days or weeks before and after attacks (e.g. Finseraas and Listhaug, 2013; Legewie, 2013; Schüller, 2016). Samples pooled across wide time intervals, however, can obscure rates of decay. We overcome this problem by using data that include exceptionally large daily samples across a three-week period immediately following the Paris attacks. We show that although the effects of the attacks are both clear and immediate, they were decidedly short-lived. This finding has important implications for our understanding of the size of the window of opportunity open to political entrepreneurs in the aftermath of terrorism and explaining the rapid birth and death of political efforts to enact policy changes following large-scale acts of violence.

4.2 The Syrian Refugee Crisis and the Paris Attacks

By the end of 2015, more than 4 million refugees had fled the civil war in Syria (United Nations, 2015). The resulting refugee crisis led to intense debate in Western democracies, particularly
in Europe. The intensity of this debate in Europe, however, at first contrasted starkly with the limited attention given to the issue in North America. This changed in September 2015 when a photo was widely published in the international media of Alan Kurdi, a Syrian child who had drowned in an effort to seek refuge in Europe with his family and whose body was pictured washed ashore on a beach in Turkey. Calls came from across Europe and North America for increased international support for refugees. In the U.S., where 2,500 refugees had been admitted since the beginning of the Syrian civil war, the Obama administration proposed to accept 10,000 more refugees by the end of 2015. In Canada, Alan Kurdi’s death occurred in the midst of a federal election campaign, and the Liberal Party, which would eventually form government, committed to accept 25,000 Syrian refugees by the end of 2015.

Two months later, the Paris attacks appeared to drastically change the international political landscape. The attacks resulted in the deaths of 130 civilians and represented the deadliest attacks in the West since the bombings in Madrid more than a decade earlier. In the wake of the attacks, politicians across Europe and North America publicly expressed strong opposition to Syrian refugee resettlement. In Europe, anti-immigration parties seized on the opportunity to argue for restrictive refugee policies. In the U.S., House Speaker Paul Ryan called for a “pause” in the U.S.’s plan to accept more refugees. “Our nation has always been welcoming,” he remarked, “but we cannot let terrorists take advantage of our compassion” (Werner, 2015). In Canada, Saskatchewan Premier Brad Wall became the voice of opposition to resettlement. Wall, who two months earlier had expressed willingness to increase his initial pledge to resettle Syrian refugees, called for the federal government to postpone its plan to admit 25,000 Syrian refugees (Canadian Broadcasting Company, 2015b).

The reactions of politicians and commentators suggested that concerns about national security had become the driver of policy positions on refugee resettlement. The attacks in Paris, it appeared, had led political leaders and the public to replace sympathy for Syrian refugees with heightened fears over national security.

4.3 Terrorism and Attitudes toward Refugees

Do large-scale terrorist attacks affect the public’s emotions and attitudes toward refugees? To answer this question, we examine the effects of the Paris terrorist attacks on emotions toward and beliefs about Syrian refugees, and the public’s policy preferences regarding resettlement.

4.3.1 Terrorism and emotion

In the aftermath of terrorist attacks in which refugees are perceived to be indirectly linked to the perpetrators — whether by religion, nationality, or otherwise — two competing emotions are expected to drive attitudes toward refugees and preferences over refugee policy. On the one hand,
terrorism may increase perceptions that terrorists will infiltrate refugee flows, raising anxiety over the possibility of future terrorist attacks close to home. On the other hand, increased perceptions of a link between refugees and terrorism may decrease sympathy for refugees themselves.

**Anxiety**  In general, studies that investigate emotional responses to terrorism find a positive relationship between perceptions of the threat of terrorism and anxiety (Huddy et al., 2005; Huddy, Feldman and Weber, 2007; Merolla and Zechmeister, 2009). These findings are complemented by research in psychology that demonstrates that anxiety is frequently linked to situations that individuals perceive to be out of their immediate control (Lerner and Keltner, 2001; Tiedens and Linton, 2001).

Anxiety also figures prominently in explanations of policy preferences. For example, higher levels of anxiety are associated with preferences for isolationist foreign policies (Huddy et al., 2005) and restrictive immigration policies (Brader, Valentino and Suhay, 2008; Renshon, Lee and Tingley, 2015). There is not strong evidence, however, that anxiety concerning the threat of terrorism is associated either with support for restrictive immigration or harsh anti-terrorism policies (Huddy et al., 2005). Yet, empirical investigation into these relationships has not been conducted in the context of real-world terrorist attacks. Indeed, it would be surprising if anxiety toward the presence of outgroups, such as refugees, were not affected by terrorism, or if anxiety were not a predictor of opposition to refugee resettlement. Whether real-world terrorist attacks increase anxiety regarding refugees and resettlement, in other words, remains an open empirical question.

**Sympathy**  The focus in the literature on the ‘negative’ determinants of attitudes toward outgroups, such as anxiety, prejudice, and threat perception has come at the relative neglect of ‘positive’ determinants, such as sympathy or affect. This neglect is unfortunate because ‘positive’ emotions may be important predictors of attitudes toward outgroups. In one of the few studies to address this, for example, survey respondents with higher levels of empathy who are informed of the difficult conditions faced by prospective immigrants are shown to be less likely to favor restrictions on immigration than those not informed of such conditions (Newman et al., 2013; see also Haubert and Fussell, 2006). Furthermore, the reasons for immigration and asylum-seeking have been demonstrated to play a substantial role in preferences over individual migration applications, with those seeking entrance for reasons of persecution being much more likely to be chosen for admittance than those seeking to improve their economic conditions (Hainmueller and Hopkins, 2015; Bansak, Hainmueller and Hangartner, 2016).

One possible reason underlying sympathy for refugees is that refugees may be perceived to face dire situations as a result of factors outside of their control. Outgroups that are perceived to lack control over outcomes in their lives, for example, have been shown to elicit compassion.
(Weiner, 2006; Gill and Andreychik, 2007), whereas those who face hardships due to internal factors (e.g. lack of effort) elicited hostility (e.g. Aarøe and Petersen, 2014; Harell, Soroka and Iyengar, 2017). Terrorist attacks may affect the balance of these factors by implicating refugees as members of a society perceived responsible for breeding and exporting violence. By shifting perceptions of responsibility onto refugees themselves, terrorism may therefore decrease public sympathy for refugees overall.

4.3.2 Terrorism and threat perception

Prior to the Paris attacks, opposition to Syrian refugee resettlement frequently centered on the potential threat posed by refugees to Western culture and the economy. Such concerns are central to explanations in the literature regarding attitudes toward immigrants (for a recent review, see Hainmueller and Hopkins, Forthcoming). However, in contrast to debates over immigration, the growth of the Syrian civil war and the expansion of Islamic State pushed the issue of refugee resettlement toward concerns that Syrian refugees posed a substantial threat to national security.

Security threat Even before the Paris attacks, politicians and commentators expressed concerns that Syrian refugee resettlement would pose a threat to national security. For example, when a series of controversial vetting procedures for Syrian refugees were introduced by the Canadian government, the Prime Minister emphasized the need to keep the “country safe and secure” (Bailey, Strezhnev and Voeten, 2015), with similar concerns expressed by politicians across the United States and Europe.

This “securitization” of migration is not a new phenomenon (Messina, 2014), and the claim that migration flows can pose a threat to national security is not without empirical support (Bove and Böhmelt, 2016). Yet even if refugee flows were not empirically linked to terrorism, we can expect that terrorism will affect perceptions of refugees as members of a society, culture, and/or religion thought to bear responsibility. These changes in perceptions may arise from the inherent uncertainty in determining which individuals pose a security threat. This may lead terrorism to be perceived as a group-based threat (Huddy and Feldman, 2011), where those responsible are defined in homogenizing terms (Rothgerber, 1997), causing an increase in negative attitudes toward members of an outgroup, broadly defined (e.g. LeVine and Campbell, 1972). Accordingly, we expect that the Paris attacks, which linked Syrian refugees with the perpetrators by virtue of their shared nationality and religion, increased perceptions of refugees as a threat to national security.

Cultural threat The perceived differences between Western culture and that of Syrian refugees were used before the attacks by opponents of resettlement to represent refugees as a threat to
Western norms and practices. Hungarian Prime Minister Viktor Orban exemplified this strain of thought. In closing Hungary’s borders to Syrian refugees, he justified the action by claiming that Hungarians “do not want to see a significant minority among [themselves] that [have] different cultural characteristics and background[s]. ... We would like to keep Hungary as Hungary.” (Traub, 2015).

In addition to representing a physical threat, Islamic terrorism is also perceived as an attack on the West’s values and culture. As such, terrorism is expected to exacerbate perceptions of cultural differences with refugees. Threats to a group’s identity or culture such as these, often called ‘symbolic’ threats (e.g. Kinder and Sears, 1981; Stephan, Ybarra and Bachman, 1999), have been shown to increase the strength of ingroup identity; to raise the salience of ingroup-outgroup differences; and to generate outgroup hostility (Tajfel, 1982; Brewer, 2001). Negative attitudes resulting from symbolic threats have, moreover, been shown to have consequences for policy preferences regarding the groups associated with those threats (Sniderman, Hagendoorn and Prior, 2004; Newman, Hartman and Taber, 2012; de Rooij, Goodwin and Pickup, 2015). We expect, therefore, that the Paris attacks increased perceptions of Syrian refugees as culturally incompatible with Western society and, consequently, a potential threat to national culture.

4.4 Terrorism and Political Mobilization

Policy preferences, and the emotions and perceptions that underlie them, represent an important area of public opinion with potential to be affected by terrorism. Yet whether public preferences translate into pressure on legislators may be as consequential for policy as are changes in preferences themselves. Does terrorism mobilize the public to apply pressure on political representatives? And, if so, are opponents of resettlement more politically mobilized than supporters following terrorist attacks?

To begin, there is good reason to expect that terrorism is politically mobilizing because changes in emotions and threat perceptions have both been demonstrated to increase political participation (e.g. Marcus, Neuman and MacKuen, 2000; Cho, Gimpel and Wu, 2006; Valentino et al., 2011). For example, increased perceptions of threat in the form of undesired policy changes are shown to mobilize issue publics (Miller and Krosnick, 2004); increases in anxiety, to stimulate mobilization (Marcus, Neuman and MacKuen, 2000; Brader, Valentino and Suhay, 2008); and terrorist attacks, to increase voter registration and turnout (Cho, Gimpel and Wu, 2006; Hersh, 2013; Robbins, Hunter and Murray, 2013).

There is also reason, however, to expect that terrorism will result in higher mobilization among some groups more than others. Hersh (2013) finds, for instance, that relatives and neighbors of victims of the 9/11 attacks were more likely than otherwise similar individuals to vote in future elections, and Cho, Gimpel and Wu (2006) find that Arab-Americans mobilized by
way of voter registration in response to the debate over and implementation of the Patriot Act.

For the case examined herein, the Paris attacks provided opponents of refugee resettlement with the political opportunity to highlight the potential risks of liberal resettlement policy. The attacks therefore provided opportunities for advocates on one side of the issue to argue for a preferred policy. But terrorism can simultaneously close off opportunities for others. For supporters of resettlement, the Paris attacks appeared to undermine, albeit indirectly, claims that Syrian refugees would pose no or few threats to national security. As a consequence, the attacks are expected to have left the level of mobilization among supporters of resettlement unchanged or to have been de-mobilizing.

4.5 Effect Duration

Lastly, one of the least examined but most theoretically and politically important considerations in the analysis of terrorism is effect duration. Effects that decay rapidly and those that are long-lasting have substantially different political implications. As Gaines, Kuklinski and Quirk (2007, 2) write, “determining the rates of decay of various treatment effects and deriving the political implications could be one of the most informative tasks” of research design and analysis. Yet, despite its fundamental importance, effect duration is rarely investigated. Among the few studies that do investigate duration — typically those in the context of political communication survey experiments — effects tend to disappear rapidly (Luskin, Fishkin and Jowell, 2002; Druckman and Nelson, 2003; Mutz and Reeves, 2005; Gerber et al., 2011; Hill et al., 2013).

Large-scale terrorist attacks, however, have the potential to bring about much more long-lasting emotional and attitudinal responses. Real-world events may also generate more persistent effects than those manipulated in a laboratory or survey-experimental setting. Hersh (2013) shows, for example, that increases in conservatism caused by the 9/11 attacks among relatives and neighbors of victims were observable more than ten years later.

The absence of large-scale survey data to track public opinion with precision immediately following terrorist attacks, however, has meant that the duration of their effects are unknown. On the one hand, the Paris attacks were substantial in their scale, with clear links to the issue of migration and refugee resettlement, suggesting the potential for long-lasting effects. On the other hand, research into the duration of effects on public opinion in other domains has demonstrated their rapid decay. Whether the effects of large-scale terrorist attacks on public opinion are long- or short-lasting thus remains an open question, with important political implications.
4.6 Hypotheses

Given the foregoing discussion, our empirical expectations are the following. First, we hypothesize that the Paris attacks affected emotions in two ways:

\[ H_1: \text{The Paris attacks increased anxiety regarding the presence of Syrian refugees.} \]

\[ H_2: \text{The Paris attacks decreased sympathy for Syrian refugees.} \]

Second, we hypothesize that the attacks affected public perceptions of Syrian refugees as threats to security and culture:

\[ H_3: \text{The Paris attacks increased perceptions of Syrian refugees as threats to national security.} \]

\[ H_4: \text{The Paris attacks increased perceptions of Syrian refugees as threats to national culture.} \]

As a result of these effects, we expect the following:

\[ H_5: \text{The Paris attacks decreased support for Syrian refugee resettlement.} \]

Lastly, we hypothesize that the attacks mobilized the public to advocate for policy change, but with different outcomes for supporters and opponents of resettlement:

\[ H_6: \text{The Paris attacks increased political mobilization concerning refugee resettlement policy.} \]

\[ H_7: \text{Following the Paris attacks, mobilization will be higher among opponents of resettlement than among supporters.} \]

4.7 Research Design

To test the foregoing hypotheses, we rely on the fortunate timing of a large-scale post-election study that contains an extensive range of questions concerning Syrian refugees that was fielded to a national online panel of respondents in Canada by the research firm Vox Pop Labs.\(^3\) The questions were asked as part of a $3 \times 2$ factorial survey experiment module that was designed

\(^3\)The post-election survey was sent to all respondents in the research firm's national online panel, of which roughly a quarter received the questions regarding Syrian refugees. Respondents were therefore not purposefully selected to match the socio-demographic characteristics of the national population. Estimates presented in the results section are therefore statistically adjusted through regression and survey weighting as appropriate to the context. For further details, see the Supplementary material.
to test whether physical proximity to resettled Syrian refugees and refugees’ religious affiliation affected support for resettlement and willingness to contact a political representative about the issue.\textsuperscript{4} The survey further captured 21 indicators to build indexes measuring feelings of sympathy for Syrian refugees; anxiety regarding Syrian refugee resettlement; and perceptions of Syrian refugees as a potential threat to national security, culture, and the economy. The first wave of the survey was fielded on November 11, 2015. The terrorist attacks in Paris occurred less than 48 hours later.

As a result of the attacks, we modified the fielding schedule of subsequent survey waves. The second wave was fielded to a large sample of respondents less than 48 hours after the attacks and, to anticipate examination of effect duration, fielded to further independent samples of respondents each day for 18 days thereafter. In sum, the data constitute an exceptionally large three-week repeated cross-sectional survey (n = 18,763) with the largest samples collected immediately before and after the attacks.

\textbf{4.7.1 The Paris attacks as natural experiment}

Due to the timing of the survey and the exogeneity of the attacks with respect to survey fielding, we treat the Paris terrorist attacks as an as-if randomly assigned treatment to respondents surveyed within a short time interval (within 2 days) of the attacks. The large daily samples collected thereafter are leveraged to examine effect duration. Because treatment assignment is a function of time, we first address two primary concerns regarding the credibility of our research design.

The first potential concern is that events or processes other than the Paris attacks may have contributed to differences in survey responses between the immediate pre- and post-attack periods. Theoretically, any differences in outcomes across this short fielding interval could be decomposed into the sum of the effect of the attacks and that due to other causes. There is good reason, however, to expect that differences in attitudes and emotions due to other events or processes are implausible. First, the attacks occurred during a period outside of any meaningful news cycle concerning either Syrian refugees or the Syrian conflict in general. Second, as noted variously throughout the literature, public opinion in such cases is slow-moving (e.g. \textcite{PageShapiro, Zaller, DruckmanLeeper}). Finally, as we demonstrate graphically in the Results section, once the effects of the attacks return to the pre-attack baseline, there appears to be little if any change across time. We assume therefore that any differences between the immediate pre- and post- periods result from the occurrence of the attacks, and that if this assumption were violated, that any bias in our estimates would, at most, be slight.

\textsuperscript{4}Respondents were assigned at random to conditions specifying the religion of Syrian refugees and their potential place of resettlement. For the purposes of this study, these factors are not examined herein, and will be analyzed in a companion paper.
The second and potentially larger threat to causal inference is differential survey non-response. Although the survey was fielded in each wave to randomly selected sets of respondents, and therefore by design the treatment is independent of survey receipt, the composition of the treatment and control samples may differ due to differences in survey non-response. There is reason, however, to reject this as a plausible explanation of differences in observed outcomes. First, the survey invitation itself referred to the post-election survey as such, and did not reference either the Paris attacks or refugees, minimizing the possibility of selection as a consequence of survey content. Second, we conduct a series of pre-treatment covariate balance checks. These checks demonstrate that the immediate pre- and post-attack samples are extremely similar in composition and none of a variety of pre-treatment covariates predict membership in the treatment or control groups. Furthermore, a likelihood-ratio test finds no strong evidence \( p = 0.74 \) that the full set of pre-treatment covariates jointly differentiate between respondents in the immediate pre- and post-attack samples (for further details, see the Supplementary Material). We nevertheless statistically adjust for potential differences in the pre- and post-attack samples by including pre-treatment covariates in regression models as appropriate to each outcome of interest. In only one model do the adjusted and unadjusted statistical tests differ, which we discuss explicitly at the relevant point of the text.

### 4.7.2 Survey design

The survey was designed to capture two primary outcomes: (1) support for Syrian refugee resettlement and (2) willingness to contact a political representative regarding resettlement. It further captured 21 separate indicators measuring sympathy for and anxiety toward Syrian refugees, and perceptions of Syrian refugees as a threat to security, culture, and the economy. To avoid burdening respondents with large sets of questions, respondents were assigned at random to one of two survey branches containing either questions regarding emotions (sympathy and anxiety), or those regarding threat (security, cultural, and economic).

All respondents read a short paragraph stating that the government was considering admitting more refugees from Syria. Respondents who were assigned to the emotion branch were then asked to specify, on a 0 to 10 scale,\(^5\) the degree to which they felt the following emotions toward Syrian refugees: sympathy, indifference, compassion, sadness, and distress. Responses to these five items were summed to form a sympathy index \( \alpha = 0.79 \). Respondents were then asked a similar question regarding their anxieties concerning Syrian refugee resettlement for the followings items: anxiety, upset, worry, anger, fear, pride, and hope. Responses to the first five of

---

\(^5\)Terminal ends of the scale were labeled ‘Not at all’ and ‘A great deal’.
these items were summed to form an anxiety index ($\alpha = 0.93$).\(^6\)

Respondents who were assigned to the threat branch of the survey were asked three sets of three questions to measure beliefs concerning the degree to which Syrian refugee resettlement posed a threat to security, culture, and the economy. To create a security threat index, respondents were asked whether they believed that some refugees would have links to terrorism; whether refugees would pose a threat to national security; and whether refugees’ presence would lead respondents to fear for their safety. A six-category likert scale, from “Strongly disagree” to “Strongly agree”, was used as the response scale. Responses were then summed to form a security threat index ($\alpha = 0.91$). To measure perceptions of refugees as a cultural threat, respondents were asked whether they believed that Syrian refugees would integrate well into society; whether their values would conflict with those of Canadians; and whether their presence would benefit national culture ($\alpha = 0.81$). For the final index, respondents were asked whether Syrian refugees’ presence would be economically costly; whether refugees would help grow the economy; and whether refugees would increase competition for jobs ($\alpha = 0.64$). All indexes are standardized to have mean zero and unit variance in the pre-attack period, and estimated effects are therefore presented in standardized units.

Lastly, all respondents were asked whether they favored Syrian refugee resettlement on a 6-category likert scale (“Strongly disagree” to “Strongly agree”), and whether they would consider contacting their Member of Parliament regarding refugee resettlement (“No”/“Yes”).

4.8 Results

Before testing each hypothesis statistically, we present each measure graphically across time. As we will see, these graphs tell an exceptionally clear story concerning how emotions and attitudes shifted (or did not shift) between the immediate pre- and post-attack period, and how they changed over time.

4.8.1 Emotions and attitudes toward refugees and resettlement

Anxiety and sympathy Figure 4.1 presents the raw data for the indexes measuring anxiety concerning resettlement and sympathy for refugees. To ease visual interpretation, changes across time in this figure and subsequent ones are approximated with a second-degree polynomial

\(^{6}\)Following Brader, Valentino and Suhay (2008), we label this index “anxiety”, but it may also be interpreted as measuring “negative affect” more generally. Some studies also suggest that anxiety and anger are measurably distinct (Druckman and McDermott, 2008; Huddy, Feldman and Cassese, 2007; Petersen, 2010). Thus as a robustness check, we demonstrate in the Supplementary Material that removal of the ‘anger’ indicator from the “anxiety” index does not affect the results.

\(^{7}\)In factor analysis, the loadings for pride and hope were low, and thus these indicators were excluded from the index.
Table 4.1: Emotions and perceptions of threat OLS regression results

<table>
<thead>
<tr>
<th></th>
<th>Anxiety (1)</th>
<th>Sympathy (2)</th>
<th>Security threat (3)</th>
<th>Cultural threat (4)</th>
<th>Economic threat (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paris attacks</td>
<td>0.265***</td>
<td>-0.033</td>
<td>0.391***</td>
<td>0.190***</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.050)</td>
<td>(0.054)</td>
<td>(0.051)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Female</td>
<td>0.050</td>
<td>0.263***</td>
<td>0.022</td>
<td>-0.122*</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.048)</td>
<td>(0.052)</td>
<td>(0.049)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Age 30-39</td>
<td>0.089</td>
<td>-0.075</td>
<td>0.236**</td>
<td>0.119</td>
<td>0.256**</td>
</tr>
<tr>
<td></td>
<td>(0.088)</td>
<td>(0.080)</td>
<td>(0.090)</td>
<td>(0.085)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Age 40-49</td>
<td>-0.003</td>
<td>0.195*</td>
<td>0.093</td>
<td>0.013</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.087)</td>
<td>(0.098)</td>
<td>(0.092)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>0.030</td>
<td>0.209**</td>
<td>0.126</td>
<td>0.102</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.074)</td>
<td>(0.081)</td>
<td>(0.075)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>Age 65+</td>
<td>0.107</td>
<td>0.171*</td>
<td>0.212*</td>
<td>0.140</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.076)</td>
<td>(0.085)</td>
<td>(0.080)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>College</td>
<td>0.176*</td>
<td>-0.192*</td>
<td>0.032</td>
<td>-0.042</td>
<td>-0.078</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.078)</td>
<td>(0.084)</td>
<td>(0.079)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>University degree</td>
<td>-0.189*</td>
<td>0.169*</td>
<td>-0.322***</td>
<td>-0.326***</td>
<td>-0.354***</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.068)</td>
<td>(0.074)</td>
<td>(0.070)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Francophone</td>
<td>0.120</td>
<td>-0.100</td>
<td>0.025</td>
<td>-0.010</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.084)</td>
<td>(0.100)</td>
<td>(0.094)</td>
<td>(0.092)</td>
</tr>
<tr>
<td>Other language</td>
<td>0.056</td>
<td>-0.085</td>
<td>0.016</td>
<td>0.004</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.077)</td>
<td>(0.087)</td>
<td>(0.081)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Ontario</td>
<td>0.221</td>
<td>-0.265*</td>
<td>0.158</td>
<td>0.259**</td>
<td>0.285**</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.111)</td>
<td>(0.107)</td>
<td>(0.100)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>Quebec</td>
<td>0.120</td>
<td>-0.424***</td>
<td>0.282*</td>
<td>0.519***</td>
<td>0.357**</td>
</tr>
<tr>
<td></td>
<td>(0.136)</td>
<td>(0.125)</td>
<td>(0.129)</td>
<td>(0.121)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>West</td>
<td>0.169</td>
<td>-0.288**</td>
<td>0.043</td>
<td>0.250*</td>
<td>0.292**</td>
</tr>
<tr>
<td></td>
<td>(0.121)</td>
<td>(0.112)</td>
<td>(0.109)</td>
<td>(0.102)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Political ideology</td>
<td>0.199***</td>
<td>-0.140***</td>
<td>0.213***</td>
<td>0.187***</td>
<td>0.191***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.985***</td>
<td>0.626***</td>
<td>-0.992***</td>
<td>-0.917***</td>
<td>-0.942***</td>
</tr>
<tr>
<td></td>
<td>(0.148)</td>
<td>(0.135)</td>
<td>(0.143)</td>
<td>(0.134)</td>
<td>(0.131)</td>
</tr>
</tbody>
</table>

N 1,688 1,685 1,799 1,800 1,802

Standard errors in parentheses. * p < .05; ** p < .01; *** p < .001
regression line. As expected, levels of anxiety concerning refugee resettlement increase sharply in the immediate aftermath of the Paris terrorist attacks. In the days that follow, however, the effect of the attacks on anxiety rebounds rapidly. Within roughly a week and a half after the attacks, levels of anxiety concerning refugee resettlement return to their pre-attack level.

To examine the effect of the attacks on anxiety statistically, we regress the anxiety index on an indicator variable representing the post-attack period and a set of pre-treatment covariates using data collected within 2 days of the attacks. Regression results are presented in Model (1) of Table 4.1. Consistent with the first hypothesis, and as is visually evident in Figure 4.1, there is strong evidence (\(p < 0.001\)) that the attacks caused a substantial increase in the public’s anxiety over Syrian refugee resettlement.

Turning to the second panel of Figure 4.1, we observe a similarly clear but substantively different picture with respect to the effect of the attacks on sympathy for refugees. In contrast to the pronounced increase in anxiety observable in the first panel, there is no clear indication that the attacks meaningfully affected sympathy for Syrian refugees themselves. Regression results presented in Model (2) of Table 4.1 bear this out: differences in sympathy for Syrian refugees between the immediate pre- and post-attack period is neither large nor is there strong evidence (\(p = 0.52\)) that the attacks affected public sympathy for refugees.

**Security and cultural threat**  Figure 4.2 presents the raw data for the security and cultural threat indexes. As can be see in both panels of Figure 4.2, respondents’ beliefs regarding whether refugees pose a security and cultural threat to the country both increase in the immediate aftermath of the Paris terrorist attacks, with a particularly sharp increase in perceptions of refugees as a security threat. Similar to the pattern of change in anxiety following the attacks,
however, these changes in threat perceptions rebound rapidly, returning to their pre-attack levels roughly one to two weeks after the attacks.

To examine the effects of the terrorist attacks on threat perceptions statistically, we fit a regression model to each threat index using data collected within 2 days of the attacks. Results from these models are presented in the third and fourth columns of Table 4.1. The results provide strong evidence that both the perceived threat posed by refugees to security (p < 0.001) and to culture (p < 0.001) increase in the immediate aftermath of the terrorist attacks.

Support for refugee resettlement To investigate the effect of the Paris terrorist attacks on support for refugee resettlement, we begin by presenting the raw data in Figure 4.3. As expected, there is a sharp decrease in support for refugee resettlement immediately following the attacks. In a now-familiar pattern, in the days that follow, attitudes toward refugee resettlement rebound rapidly. Within roughly a week after the attacks, support for resettlement rises to its pre-attack level, increasing slightly further before remaining relatively constant during the final week that data were collected.

To examine the effect of the attacks on support for refugee resettlement statistically, we fit an ordinal regression model to the data collected within two days of the attacks. Results from the model are presented in Table C.3. They show strong evidence (p < 0.001) that the attacks caused a decrease in support for refugee resettlement.

To calculate the magnitude of this effect on the original scale, we use parameter estimates from the fitted model to calculate the probability of support for resettlement for each individual.

Notes:

8Results for the economic threat index are shown for completeness (for further detail, see the Supplementary material).
in the dataset, first by setting the Paris attacks indicator variable to 0 and then to 1, after which we calculate an average of the differences in these probabilities. Results from this procedure are presented in Figure 4.4. As the figure shows, the Paris attacks are estimated to have resulted in a 4.4 percentage point decrease in support for refugee resettlement (agree vs. disagree), a relatively moderate effect given the scale of the attacks.

### 4.8.2 Evidence of causal mechanisms

The absence of a meaningful effect of the Paris attacks on sympathy for Syrian refugees suggests that the decrease in support for resettlement may have been driven primarily by increased perceptions of refugees as a security threat and anxiety over resettlement, and not due changes to the perceived deservedness of refugees themselves. To investigate this further, we briefly examine the mechanisms involved in the decrease in support for resettlement by turning to recent advancements in methods for the study of causal mechanisms (Imai et al., 2011; Imai, Tingley and Yamamoto, 2013).

**Definition and estimation** A causal mechanism represents the process through which a causal variable affects an outcome (Imai et al., 2011, 765). The basic set up for a single mechanism can...
Table 4.2: Support for refugee resettlement ordinal logistic regression results

<table>
<thead>
<tr>
<th>Support for resettlement</th>
<th>Coef</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paris attacks</td>
<td>-0.301***</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Female</td>
<td>0.190**</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Age 30-39</td>
<td>-0.215</td>
<td>(0.116)</td>
</tr>
<tr>
<td>Age 40-49</td>
<td>0.089</td>
<td>(0.126)</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>0.102</td>
<td>(0.105)</td>
</tr>
<tr>
<td>Age 65+</td>
<td>-0.026</td>
<td>(0.109)</td>
</tr>
<tr>
<td>College</td>
<td>-0.197</td>
<td>(0.108)</td>
</tr>
<tr>
<td>University degree</td>
<td>0.435***</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Francophone</td>
<td>-0.263*</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Other language</td>
<td>-0.292**</td>
<td>(0.110)</td>
</tr>
<tr>
<td>Ontario</td>
<td>-0.537***</td>
<td>(0.160)</td>
</tr>
<tr>
<td>Quebec</td>
<td>-0.918***</td>
<td>(0.180)</td>
</tr>
<tr>
<td>West</td>
<td>-0.539***</td>
<td>(0.161)</td>
</tr>
<tr>
<td>Political ideology</td>
<td>-0.384***</td>
<td>(0.016)</td>
</tr>
</tbody>
</table>

N 3,531

Cut-point parameter estimates not shown (see Supplementary Material). *p < .05; **p < .01; ***p < .001

be diagrammed as follows:

\[ T_i \rightarrow M_i \rightarrow Y_i \]

where \( T_i \) denotes treatment status, \( M_i \) denotes the mechanism of interest, and \( Y_i \) denotes the outcome.\(^9\) The goal of causal mechanisms analysis is to decompose the effect of a treatment into the effect which operates through a causal mediator of interest \((T_i \rightarrow M_i \rightarrow Y_i)\) and that which operates through other channels \((T_i \rightarrow Y_i)\).\(^{10}\)

To calculate the average causal mediation effect (ACME), we use the two-step procedure

\[ \tau_i = Y_i(t, M_i(1)) - Y_i(t, M_i(0)), \]  

where \( M_i(t) \) denotes the potential value of the mediator for individual \( i \) under treatment status \( t \in \{0, 1\} \), and \( Y_i(t, m) \) denotes the potential value of the outcome when the treatment status and mediator are set to \( t \) and \( m \) respectively. As Equation 4.1 indicates, the causal mediation effect represents the difference in the effect of the treatment, holding the value of the treatment constant, and manipulating the value of the mediator as would be realized under conditions \( t = 1 \) and \( t = 0 \). For further details, see Imai et al. (2011) and Imai, Tingley and Yamamoto (2013).
Figure 4.4: Estimated effect of terrorist attacks on public support for refugee resettlement

![Effect size (%-points)](image)

In the first step, two regression models are fit to the data: first, the mediator is modeled as a function of the treatment and pre-treatment covariates; second, the outcome is modeled as a function of the mediator, treatment, and pre-treatment covariates. In the second step, the fitted models are used to predict support for refugee resettlement, first by using the predicted values of the mediator under the treatment condition, and then under the control condition, holding all other variables constant. We then calculate the average in the difference in these two predicted outcomes to estimate the average causal mediation effect.

In the following, we examine anxiety, sympathy, and security and cultural threat as mechanisms. Because two independent branches of the survey were used to capture the two sets of mechanisms (threat and emotion), the models are fit to the relevant subset of the sample collected within 2 days of the attacks. First-stage regression models are those presented in Table 4.1 and those for the second stage are provided in the Supplementary Material. Because causal mediation analysis relies on the sequential ignorability assumption (see Imai et al., 2011; Imai, Tingley and Yamamoto, 2013), which is generally considered a strong assumption, we conduct sensitivity analyses for each causal mediation estimate. The sensitivity analysis shows that our primary results are robust to the omission of a relatively strong pre-treatment confounder (see

---

11 All results provided in this section were generated using the statistical package mediation (Tingley et al., 2014).
12 This assumption requires first that, conditional on pre-treatment covariates, the treatment is independent of the potential outcome and potential mediator, and second, that the mediator is independent of the potential outcome conditional on pre-treatment covariates and observed treatment status. To make this assumption plausible, we include controls for socio-demographics and political ideology.
Table 4.3: Estimates of causal mechanisms

A. Emotion branch

<table>
<thead>
<tr>
<th>Causal mechanism</th>
<th>ACME</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>-3.3</td>
<td>(-4.7, -2.0)</td>
</tr>
<tr>
<td>Sympathy</td>
<td>-0.4</td>
<td>(-1.9, 1.1)</td>
</tr>
</tbody>
</table>

N = 1,828

B. Threat branch

<table>
<thead>
<tr>
<th>Causal mechanism</th>
<th>ACME</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security threat</td>
<td>-5.9</td>
<td>(-7.6, -4.3)</td>
</tr>
<tr>
<td>Cultural threat</td>
<td>-2.8</td>
<td>(-4.2, -1.3)</td>
</tr>
</tbody>
</table>

N = 1,789

Supplementary Material).\(^{13}\)

**Causal mechanism results** Table 4.3 presents estimates of the effect of the Paris attacks that operate through each mechanism. As the table shows, the model provides evidence that the Paris attacks decreased support for refugee resettlement by increasing the public’s anxiety about the presence of refugees. The attacks are estimated to have caused a 3.3 percentage point decrease (95% CI: -4.7, -2.0)\(^{14}\) in support for refugee resettlement by increasing anxiety about the presence of refugees. There is little evidence, on the other hand, that the attacks decreased support for refugee resettlement by decreasing sympathy for refugees themselves.

Turning to the threat measures, we find strong evidence that the effect of the attacks on support for refugee resettlement operated through respondents’ concerns about security. The Paris attacks are estimated to have caused a 5.9 percentage point decrease (95% CI: -7.6, -4.3)\(^{15}\) in support for refugee resettlement by increasing the public’s concerns over perceptions of the security threat posed by refugees. An important caveat is that although both anxiety and security appear to be substantial mechanisms through which the Paris attacks operated on public attitudes toward refugee resettlement, these mechanisms may be strongly linked within a longer causal chain, whereby terrorism affects security concerns which in turn affects anxiety or

---

\(^{13}\)As we note in the Supplementary Material, one cannot, however, rule out the potential for post-treatment confounding. As we discuss below, we address this concern directly when examining cultural threat as a causal mechanism.

\(^{14}\)Confidence intervals are calculated through non-parametric bootstrapping.

\(^{15}\)Note that this estimate is larger than the estimated (total) average treatment effect of the attacks. This is because the estimate for the direct effect (that not operating through security concerns) is positive, although not significantly so. Changes operating through the security threat channel, in other words, explain the entirety of the effect on support for resettlement.
vice versa. However, because the survey was not designed to untangle this relationship and used two independent branches to measure emotion and threat, we cannot investigate this more complex relationship further. Untangling this and other similar relationships is a difficult, but important empirical question that we leave for future research.

Finally, we estimate the average causal mediation effect of the Paris attacks through cultural threat perceptions. To do so, we use a model that relaxes the assumption of causal independence between security and cultural threat. Theoretically, it is unlikely that perceptions of cultural threat are causally independent of the effect that the attacks have on security concerns. We therefore posit a model whereby the Paris attacks cause an increase in perceptions of security threat, which in turn affect perceptions of cultural threat. To fit this model, we use the estimation procedure proposed by Imai and Yamamoto (2013) that allows one to account for post-treatment confounding. Using this model, the Paris attacks are estimated to have caused a 2.8 percentage point decrease (95% CI: -4.2, -1.3) in support for refugee resettlement by increasing the public’s perceptions of refugees as a threat to national culture.

4.8.3 Political mobilization

Lastly, we investigate whether the Paris terrorist attacks politically mobilized the public around the issue of refugee resettlement and whether this effect resulted in differences in mobilization among supporters and opponents of resettlement. To do so, we examine, as a quasi-behavioral measure, respondents’ expressed willingness to contact a Member of Parliament (MP) regarding resettlement.

The raw data for this outcome are presented in Figure 4.5. Unlike data shown in previous figures, the effect of the attacks on the public’s willingness to contact a political representative concerning refugee resettlement is less visually apparent. To examine the effect of the attacks on political mobilization statistically, we fit a logistic regression model to the data collected within two days of the attacks, including pre-treatment covariates. Results from this model are presented in Model (1) in Table 4.4. The model provides evidence that the Paris attacks increased the probability of respondents expressing willingness to contact their MP regarding Syrian refugee resettlement ($p = 0.048$). We note, however, that although the regression-adjusted

---

Diagrammatically, this multi-mediator set up can be shown as follows:

\[
W_i \rightarrow M_i \rightarrow Y_i ,
\]

where $W_i$ denotes a post-treatment confounder (security) that causally affects both the mediator of interest $M_i$ (cultural threat) and the outcome $Y_i$ (support for refugee resettlement).

If we assume, by contrast, that perceptions of security threat is not a post-treatment confounder, the estimated average causal mediation effect of cultural threat is $-3.4$ (95% CI: $-6.9, -2.1$) percentage points.
Figure 4.5: Willingness to contact Member of Parliament across time

Each point represents a two-day average, with 90% confidence intervals.

estimate is statistically significant, evidence from the unadjusted difference (as is evident in Figure 4.5) is weaker ($p = 0.14$).

More consequential for policy and our understanding of political responses to international terrorism, however, is whether terrorism can lead to asymmetry in political mobilization among supporters and opponents of a given policy. Because attitudes toward refugee resettlement is a post-treatment variable, we do not aim to estimate the effect of the attacks on mobilization among those who opposed or supported resettlement before the attacks. Instead, we examine how willingness to contact a political representative about the issue differed between opponents and supporters of refugee resettlement immediately before and after the attacks.

We begin by fitting two models to the data. The first model is fit to evaluate the difference in willingness to contact a political representative between opponents and supporters of refugee resettlement among those surveyed immediately before the attacks; the second, among those surveyed immediately afterward. Results from these models are presented in Models (2) and (3) of Table 4.4. As Model (2) demonstrates, before the attacks there is little evidence that political mobilization regarding refugee resettlement differed between those who opposed or supported refugee resettlement ($p = 0.65$). In the immediate aftermath of the attacks, however, those opposed to resettlement were substantially more likely to express willingness to contact an MP regarding Syrian refugee resettlement ($p < 0.001$). To examine these differences, Model (4) in Table 4.4 is fit to the data collected within two days of the attacks, and includes an interaction term between support for resettlement and the Paris attacks indicator. The model demonstrates that the difference in mobilization between supporters and opponents in the post-attack period is different from that in the pre-attack period ($p < 0.01$).
To illustrate the difference in political mobilization between opponents and supporters of refugees before and after the attacks, we generate predicted probabilities of willingness to contact an MP regarding Syrian refugee resettlement using parameter estimates from the models fit separately to the pre- and post-attack data (Models (2) and (3)). Predicted probabilities are calculated using these models for all respondents in the dataset, holding covariates at their observed values.

Figure 4.6 presents the result of this calculation. As the figure shows, willingness to contact a political representative prior to the attacks is roughly equivalent between opponents and supporters of refugee resettlement: the probability of contacting an MP is slightly higher among opponents (2 percentage points). After the attacks, however, the difference in political mobilization between opponents and supporters is substantial. Among opponents of resettlement, the probability of expressing willingness to contact a political representative regarding refugees is 12 percentage points higher, a substantial difference in political mobilization; among supporters, the difference is 2 percentage points.\(^1\) The difference is thus a substantial 10 percentage points higher among opponents for the model fit to the post-attack data. Large-scale terrorist attacks, in other words, can lead to major differences in political mobilization between opponents and supporters of related policies.

\(^1\) Unadjusted estimates show even larger differences: 0.5 percentage points among supporters of resettlement between the pre- and post-attack periods; 13 percentage points among opponents.
Table 4.4: Willingness to contact MP regarding resettlement logistic regression results

<table>
<thead>
<tr>
<th></th>
<th>(1) Within 2 days of attacks</th>
<th>(2) Before attacks</th>
<th>(3) After attacks</th>
<th>(4) Within 2 days of attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paris attacks</td>
<td>0.157*</td>
<td>0.092</td>
<td>0.515***</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.207)</td>
<td>(0.123)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>Oppose refugees</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.092</td>
<td>0.515***</td>
<td>−0.036</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.207)</td>
<td>(0.123)</td>
<td>(0.087)</td>
<td></td>
</tr>
<tr>
<td>Oppose refugees × Paris attacks</td>
<td></td>
<td></td>
<td></td>
<td>0.608**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.219)</td>
</tr>
<tr>
<td>Female</td>
<td>0.187*</td>
<td>0.171</td>
<td>0.207*</td>
<td>0.194*</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.141)</td>
<td>(0.092)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Age 30-39</td>
<td>0.144</td>
<td>−0.182</td>
<td>0.262</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.242)</td>
<td>(0.158)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>Age 40-49</td>
<td>0.446**</td>
<td>0.527*</td>
<td>0.427*</td>
<td>0.454**</td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.261)</td>
<td>(0.170)</td>
<td>(0.142)</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>0.454***</td>
<td>0.231</td>
<td>0.553***</td>
<td>0.455***</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.218)</td>
<td>(0.143)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>Age 65+</td>
<td>0.519***</td>
<td>0.387</td>
<td>0.579***</td>
<td>0.517***</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.229)</td>
<td>(0.147)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>College</td>
<td>0.022</td>
<td>0.100</td>
<td>−0.012</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.226)</td>
<td>(0.150)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>University degree</td>
<td>−0.090</td>
<td>0.071</td>
<td>−0.085</td>
<td>−0.032</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.198)</td>
<td>(0.133)</td>
<td>(0.110)</td>
</tr>
<tr>
<td>Francophone</td>
<td>−0.498***</td>
<td>−0.243</td>
<td>−0.695***</td>
<td>−0.508***</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td>(0.248)</td>
<td>(0.181)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>Other language</td>
<td>−0.345**</td>
<td>−0.500*</td>
<td>−0.300*</td>
<td>−0.359**</td>
</tr>
<tr>
<td></td>
<td>(0.122)</td>
<td>(0.226)</td>
<td>(0.147)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Ontario</td>
<td>−0.136</td>
<td>0.033</td>
<td>−0.215</td>
<td>−0.154</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.300)</td>
<td>(0.183)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Quebec</td>
<td>−1.364***</td>
<td>−1.347***</td>
<td>−1.342***</td>
<td>−1.391***</td>
</tr>
<tr>
<td></td>
<td>(0.187)</td>
<td>(0.345)</td>
<td>(0.226)</td>
<td>(0.188)</td>
</tr>
<tr>
<td>West</td>
<td>−0.130</td>
<td>0.167</td>
<td>−0.259</td>
<td>−0.147</td>
</tr>
<tr>
<td></td>
<td>(0.156)</td>
<td>(0.305)</td>
<td>(0.184)</td>
<td>(0.157)</td>
</tr>
<tr>
<td>Political ideology</td>
<td>−0.051**</td>
<td>−0.125***</td>
<td>−0.054**</td>
<td>−0.076***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.034)</td>
<td>(0.021)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.015</td>
<td>0.165</td>
<td>0.117</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.371)</td>
<td>(0.233)</td>
<td>(0.206)</td>
</tr>
<tr>
<td>N</td>
<td>3,535</td>
<td>1,128</td>
<td>2,397</td>
<td>3,525</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. * p < .05; ** p < .01; *** p < .001
4.9 Discussion and Conclusion

Although research into the effects of terrorism has grown in recent years, few studies have examined the effects of real-world acts of terrorism on public opinion and behavior internationally. Filling this gap in the literature is important because the audiences targeted and affected by terrorism, and the policy implications, are frequently international.

This study helps fill this gap by examining the effects of the 2015 Paris terrorist attacks on public attitudes, emotions, and mobilization concerning refugees and refugee resettlement. The Paris attacks are a critical case not only because they were the largest in the West in over a decade, but also because they coincided with the Syrian refugee crisis, the largest international refugee crises since the end of the Second World War. Unlike previous attacks in the West, whose policy implications centered primarily on anti-terrorism and national security policy generally, the Paris case is defined by the intersection of national security policy and concerns over international migration. The Paris case represents one of the clearest examples of overlap between national security concerns tied to terrorism and a set of domestic and international policies consequential for a wide range of actors. The implications of this study are therefore many.

Attitudes and emotions To begin, the results suggest that public preferences over policy are highly resilient, even in the face of what were substantial, dramatic, and widely covered terrorist attacks. Although the attacks increased opposition to refugee resettlement as hypothesized, the magnitude of this effect was a moderate 4 percentage point decrease in support for resettlement.

This estimate is particular to a single country (Canada), but it is unlikely that its magnitude differed substantially across Europe or in the U.S. First, as in Europe and the U.S., the Syrian refugee crisis was a sensitive and highly political issue in Canada. As elsewhere, there was substantial debate about refugee resettlement, and the Paris attacks led to concern over the threat posed by refugees to national security and a vocal political backlash against resettlement. Second, the resilience of public opinion to the attacks appears to be mirrored elsewhere in policy. For example, in Germany and Sweden — two major recipients of refugees in Europe — resettlement policy changed little in the aftermath of the attacks despite heated debate. In France, the government promised to accept 30,000 Syrian refugees shortly after the attacks, despite their having occurred in the French capital (Tharoor, 2015). In the United States, the Obama administration maintained its plan to accept 10,000 refugees by the end of 2016.

Our findings also demonstrate that the Paris attacks increased the public’s anxiety toward refugees and increased perceptions of refugees as a threat to security and culture. We show, furthermore, that the decrease in support for resettlement operated through each of these channels. One limitation of this study, however, is that the data do not lend themselves well to differentiating between threat perceptions and emotions as competing mechanisms (e.g. Brader,
Contrary to expectations, we find little evidence that the attacks affected sympathy for Syrian refugees themselves. Combined with the clear effects of the attacks on threat perceptions and anxiety, this suggests that although terrorism affects both attitudes and emotions toward refugees and resettlement, it does not necessarily lead the public to paint with a broad brush those who share a national, religious, and/or ethnic identity with the perpetrators. Terrorism may increase security concerns and anxiety about a minority of refugees, but concerns over security appear to increase absent any changes in beliefs in the deservedness of refugees as a whole.

**Political mobilization** If the effect of the attacks on attitudes toward resettlement policy was moderate, what explains the appearance, in a wide range of countries, of a substantial increase in vocal opposition to refugees and refugee resettlement in political discourse? Our findings suggest that while attacks terrorism may shift public attitudes on policy only moderately, they can lead to large differences in mobilization among politically important sub-groups of the population. This is suggested by the fact that, whereas before the attacks, supporters and opponents of resettlement were roughly equally likely to contact their MP concerning refugee resettlement, the difference in this form of political mobilization widened substantially in the immediate aftermath of the attacks: among opponents of resettlement, willingness to express attitudes concerning refugees to a political representative increased by 12 percentage points; among supporters, only 2.

These differences in political mobilization highlight a critical complication for the democratic process. In the aftermath of terrorism, the signal received from the public by political representatives will in part be the product of the mobilizing effect of terrorism. If mobilization differs between supporters and opponents of a given policy in the aftermath of an attack, however, then as a measure of public opinion, this signal can be misleading. Theoretically, the growth in *expressions* of opposition to refugees can increase even if the underlying *attitudes* shift little or not at all. Absent attention to information distinguishing expressions of political preference from changes in attitudes, political mobilization may be mistaken for increases in opposition to or support for highly consequential policies. This suggests that observers and political leaders should exercise caution in interpreting increases in expressions of opposition to or support for policies in the aftermath of terrorism or potentially similar acts of violence in other domains, such as mass shootings.
Duration  Perhaps the most striking finding from this study is the speed with which the observed effects of the Paris attacks rebound. With large samples collected daily for nearly three weeks after the attacks, we show that despite the severity of the attacks, their effects are surprisingly short-lived. These patterns of decay, furthermore, are both clear and similar across multiple indicators.

The rapid decay in the effects of the attacks raises a series of important questions, and opens up multiple avenues for future research. Because this is the first study to track the effects of a major terrorist attack with precision, whether the effects of comparable events cause similar dynamics is unknown. Although research in political communication also shows that effects can disappear rapidly (e.g. Druckman and Nelson, 2003; Gerber et al., 2011; Hill et al., 2013; Luskin, Fishkin and Jowell, 2002; Mutz and Reeves, 2005), it is relatively surprising that the effects of a massive terrorist attack follow a similar pattern. This decay may be due to rapidly decreasing news exposure\(^{19}\) or be particular to the actions of politicians in the aftermath of the attacks. It is nevertheless important to highlight the scope of the Paris attacks. Because the attacks were the deadliest in the West in more than a decade, these effects, compared to those of other cases, should be those most likely to be substantial and persistent. The results suggest, however, that large-scale events may not have the large and long-term effects on public opinion that are often presented as fact by politicians and political commentators. For public attitudes and emotions toward refugees, the Paris attacks did not “change everything” for long.

\(^{19}\)For descriptive evidence that suggests this, see the Supplementary Material.
Chapter 5

Conclusion

5.1 Introduction

With the growth in, novelty of, and current focus in the social sciences on ‘Big Data’ — data whose size is typically many orders of magnitude larger than typical surveys — it is easy to ignore the research possibilities afforded by the relatively modest increases in survey sample sizes. In the foregoing articles, I make methodological and substantive contributions to the literature on public opinion by capitalizing on the possibilities opened up by the availability of these data. As this dissertation shows, recent increases in survey sizes and computational resources can open up methodological and empirical opportunities for researchers to ask and answer important questions in public opinion research that in the past were exceptionally difficult to address.

Because each article in this dissertation stands as an independent contribution to the public opinion literature that it addresses, in this concluding chapter I discuss how these articles together contribute to the study of large-scale survey data and how they fit into current trends in the quantitative social sciences. Lastly, I suggest how we might begin to find links between traditional survey-based modes of data collection and newer forms of public opinion, such as those from social media.

5.2 Motivation and contributions

The motivation for this dissertation has been to explore the ways in which the growth in the size of traditional survey data and the increasing availability of computational resources has made possible the investigation of research questions that were previously difficult to answer. What explains why some people misreport their attitudes toward sensitive survey questions and others do not? How can we measure perceptions of the ideology of local-level political representatives and citizens on a common ideological scale? What are the effects of large-scale terrorist attacks
on attitudes toward refugee resettlement and how long do these effects last? These questions differ in their substance, but their answers require taking advantage of large-scale public opinion data and, in the former two cases, the development of new methods to take advantage of these data. In this section, I suggest the benefits of large-scale survey data demonstrated by this dissertation and consider the importance for researchers to thoroughly explore the benefits such data offer.

5.2.1 Bigger data, smaller groups

The first benefit of large-scale survey data shown by this dissertation is that they permit researchers both to examine small groups and to use smaller groups in aggregate to answer previously intractable questions.

The most obvious way that large-scale survey data allow researchers to examine small groups is by providing sufficient samples of proportionally small socio-demographic groups for analysis. Ghitza and Gelman (2013), for example, use large-scale survey data\footnote{The authors use data from the National Annenberg Election Survey and aggregated Pew Research pre-election polls.} to examine vote choice and turnout in the 2000, 2004, and 2008 U.S. presidential elections. These data, which due to their size contain many respondents per state and per socio-demographic group, permit the authors to examine how the relationship between income and vote intention varies across states, and how the relationship between race and voter turnout varies across age groups. The authors demonstrate that when such data are combined with a relatively complex model, they can provide insight into the electorate in ways that would be impossible with smaller samples.

Each of the three articles presented in this dissertation, on the other hand, demonstrate less obvious ways in which we can think about the benefits of large-scale survey data to study small groups.

In the first article, for example, the goal was to examine misreporting among respondents who hold a socially unacceptable belief. In this case, the population of interest is a sub-group: those who hold the sensitive belief. Ideally, one would collect survey data from only these respondents. The problem, however, is that we cannot \textit{a priori} know who belongs to this group. In fact, the purpose of indirect questioning techniques is to encourage respondents to answer sensitive survey questions honestly by providing a context in which researchers cannot identify whether any given individual holds or does not hold that belief. In aggregate, these data are useful for estimating desired quantities of interest, but indirect questioning techniques nevertheless prevent individuals within the sub-group of sensitive-belief holders from being identified. In this case, we must collect data from a substantially larger population, even if the population of interest is much smaller. The growth in the availability of large-scale survey data allows researchers to investigate these types of problems and should encourage researchers to
development new methods to take advantage of them as a result.

In the second article, large-scale survey data allow research into the ideology of hundreds of sub-national actors by dividing the measurement problem up among smaller groups. In particular, surveys such as the Cooperative Congressional Election Study ask respondents to ideologically place (1) a small group of political actors such as national-level politicians, who are placed by all respondents, and (2) a small subset of local-level actors such as the governor or senators in a respondent's state. As a consequence, of the hundreds of actors who are placed in aggregate, each respondent is asked to place relatively few. As I demonstrate in the second article, the benefit of this approach is that researchers can use the actors who are placed by all respondents as ‘bridges’ in an AM scaling model. This permits the many hundreds of political actors, each of whom is placed by only a subset of respondents, to be placed on a common scale with precision. As I show in the empirical section of the article, this allows researchers to examine how an important state-level actor (Bernie Sanders) compares ideologically to the national electorate. The benefit of these data is not only that we gain precision in our estimates, but also that they allow for a feasible survey design. The alternative would, in effect, be to ask all respondents to place hundreds of actors within a single survey. These types of problems, in other words, require large-scale survey data both for reasons of precision and to make the survey itself feasible.

Finally, in the third article, my co-author and I demonstrate how large-scale surveys with samples split into large time series cross-sections can answer one of the least examined but most theoretically and politically important considerations in the social sciences: the duration of causal effects. As we argue, rapidly decaying effects and those that are long-lasting can have considerably different political implications. If the effects of major events on public opinion decay rapidly, for example, the opportunity for political entrepreneurs to enact policy changes in response to such events is likely to be similarly short. Gaines, Kuklinski and Quirk (2007, 2), in their critique of the analysis of survey experiments, highlight the importance of examining effect duration. “[D]etermining the rates of decay of various treatment effects and deriving the political implications,” they write, “could be one of the most informative tasks” of research design and analysis.

Due to the availability of large-scale time series cross-sectional data, in the third article we are able to demonstrate the duration of the effects of a major terrorist attack on attitudes and emotions toward refugees and refugee resettlement policy. The article shows that although the Islamic State terrorist attacks in Paris affected a variety of the public’s attitudes and emotions, each of the effects was decidedly short-lived. Within roughly two weeks of the attacks, perceptions of security and cultural threat, anxiety regarding the presence of Syrian refugees, and support for Syrian refugee resettlement all return to their pre-attack levels. Unlike previous work, which aggregates public opinion data across weeks in the aftermath of a terrorist attack (e.g.
Finseraas and Listhaug, 2013; Legewie, 2013), the size of the daily samples in the Paris attacks data permits us to show their effects and the pattern of their decay with exceptional clarity.

### 5.2.2 Bigger data, bigger models

The second benefit of large-scale survey data shown by this dissertation is that such data allow for the development of more complex models — models that would be severely limited or of no use without large-scale data. Bigger data, in other words, permit bigger models.

In some cases, more complex models can be useful because they can lead to more accurate forecasts. For instance, in one of the few examples of research that uses ‘Big Data’ in a survey context, Wang et al. (2015) demonstrate how survey responses from hundreds of thousands of X-Box video game users can be used for election forecasting. Because these data are highly non-representative, the authors set up a complex multilevel model to statistically adjust for bias in sample composition. The result is a highly accurate forecast of the 2012 U.S. presidential election.

However, the benefit of large-scale survey data is most pronounced when they permit the development of models to tackle problems that would be almost impossible to tackle otherwise. In the first article of this dissertation, for example, I address a question that is difficult, if not impossible, to address without large-scale survey data: what explains why some respondents misreport their responses to sensitive survey questions and others do not? This question is difficult to address for two related reasons. First, survey respondents often misreport their beliefs about sensitive issues due to a fear of violating social norms regarding beliefs that are considered socially unacceptable. Second, to examine the predictors of misreporting, researchers need to know both the ‘true’ beliefs of respondents and if they are unwilling to state them openly.

In the first article, I propose a method to permit researchers to investigate variation in the misreporting of sensitive attitudes by combining a list experiment and a standard direct question. The article develops a regression-based method to model whether respondents provide one response to the sensitive item in a list experiment, but answer otherwise when asked to respond openly to an analogous direct question. The model is demanding on the data, however, for two reasons: (1) parameters in the misreporting sub-model are, in effect, estimated only using the data from the sub-group of respondents who hold the sensitive belief, and (2) membership in this sub-group is unknown and itself must be estimated from the data.

These two problems are related to the size of available survey data because indirect measurement techniques help address social desirability bias at the cost of statistical efficiency. Indirect measurement techniques, in other words, require substantially larger sample sizes to achieve the same degree of efficiency as traditional direct questions (Rosenfeld, Imai and Shapiro, 2016, 797-798). The need for larger survey data for indirect measurement techniques is also not only one of precision. As Blair and Imai (2012, 67) find, more complicated models for indirect
questioning techniques can result in the complete separation of covariates when the data are too few, leading to the need to include further information to identify their model. Examination of misreporting is also compounded by the fact that interest is in modeling who misreports among those who hold the sensitive belief, an often small sub-group within the larger sample. The availability of larger-scale survey data thus make the problem of investigating misreporting on sensitive survey questions tractable.

In the second article, on the other hand, I develop a model that is both demanding on the data and highly computationally intensive. The goal of the article is to adjust for perceptual biases in survey respondents’ understanding of ideological scales. A political actor who is considered ‘Extremely conservative’ by one respondent might, for example, be considered ‘Moderate’ by another. The contribution of the article is the development of a Bayesian Aldrich-McKelvey (AM) scaling model that accounts for both the ordinality of standard ideological placement scales and incorporates respondents’ self-placement into the model itself. Furthermore, the model addresses the fact that respondents have more difficulty placing some political actors than others by introducing a heterogeneous ordinal AM scaling model.\(^2\)

The model is both demanding on the data and computationally intensive because, as noted above, when applied to data such as the Cooperative Congressional Election Study, the model is constructed to estimate the ideology of hundreds of political actors at both national and sub-national levels. Furthermore, the model estimates separate stretch and shift parameters for each individual, which are drawn from common distributions (i.e. a multilevel model). Given the size of the CCES data, each of the models presented in the article using these data contain over one hundred thousand parameters and take roughly a day to fit. The benefits of the ability to fit these types of models to large-scale data, however, are large. The model proposed in the second article, when applied to CCES data, permits researchers to capture important features of the distribution of ideology in the population and, as noted in the previous section, to compare the ideology of the national electorate to that of local-level actors.

5.3 Survey research and the growth of ‘Big Data’

Although this dissertation has demonstrated the benefits of using large-scale survey data, one of the most pressing questions in contemporary political science is how researchers should use public opinion and political behavioral data that are large, unstructured, and highly unrepresentative. In other words, ‘Big data’. In this final section, I discuss the need to bridge the gap between survey data and these new sources of data.

Some of the most innovative recent work that uses these new sources of data, such as audio

\(^2\)As noted by King et al. (2004), models to account for the ordinality of the data were avoided by Aldrich and McKelvey (1977) for lack of computational resources.
(e.g. Dietrich, Enos and Sen, 2016; Dietrich and Juelich, 2017), video (e.g. Dietrich, 2017; Voigt et al., 2017), and text (see Grimmer and Stewart, 2013; Wilkerson and Casas, 2017) require the development of new methods for their analysis, and new ways of thinking about how they reflect relevant populations of interest (Ruths and Pfeffer, 2014). Other ‘Big data’ sources, however, such as those from social media, have the potential to be used in concert with more traditional survey data, both to validate and improve measures of political phenomena. This is important because obtaining valid measurements from ‘Big Data’ is much less straightforward than obtaining such measurements from traditional survey data in which question design is under the control of researchers. Linking traditional survey data and social media data (e.g. surveying a small subset of social media users) has two potential benefits: (1) it would provide a means to rigorously validate existing methods developed for these data, and (2) it has the potential to provide new ways to improve the validity of measurements through the development of new methods.

By way of example, one of the most prominent problems in the study of social media is how to measure political ideology. In traditional political science research, measurement strategies for ideology are of course many. The best known of these are NOMINATE scores for legislators using Congressional roll-call data (Poole and Rosenthal, 1985), and factor analysis and item-response theory models using survey data (Jackman, 2001; Clinton, Jackman and Rivers, 2004; Ansolabehere, Rodden and Jr., 2008). As descriptive efforts, the development of these methods has proven extremely fruitful in advancing theoretical development Grimmer and Stewart (2013). Work on the origins of political polarization (McCarty, Poole and Rosenthal, 2006, 2009), for example, grew out of the finding that the U.S. Congress has become increasingly polarized over the last half century (Poole and Rosenthal, 1984).

It is unsurprising, then, that much recent work in political science that uses ‘Big data’ has focused on the measurement of political ideology. Many studies, for example, now examine the ideology of Facebook or Twitter users, or use these measurements to test theories regarding political behavior online (e.g. Bond and Messing, 2015; Boutyline and Willer, 2017; Barberá, Jost, Nagler, Tucker and Bonneau, 2015; Barberá, 2015).

Survey data are particularly useful to validate methods to measure ideology from social media data because the methods are typically analogous. For example, current methods in political science for the measurement of the ideology of social media users are based on who users follow or endorse: user ideology is inferred from how they place themselves in social media space relative to political actors (Barberá, 2015; Bond and Messing, 2015). The survey analogue to this is the ideological placement scale, a scale addressed in this dissertation. The validity of methods to measure the ideology of social media users would be best assessed with comparison to such survey responses. One can expect that measurement strategies for the study of political attitudes and behaviors using other sources of data will find ready analogues in survey research.

Lastly, although comparisons of estimates from social media data to those from survey
data are important for assessing validity, there may also be benefits from using survey data to ground ideology estimates from social media data. This is possible because ideological scaling methods for social media users typically rely on simultaneously estimated ideology parameters for political actors. One possible strategy is to fix the ideology parameters of political actors in social media space to those estimated from survey data (or roll-call data). This would effectively estimate the ideology of social media users on a common scale with respondents of, say, a large-scale representative survey. This could potentially lead to more valid measures of ideology and also allow for meaningful comparisons between social media users and the general population.

In sum, although this dissertation has outlined some of the potential research opportunities enabled by recent increases in survey samples sizes, these benefits can and should be assessed in tandem with, rather than independently of, new sources of large-scale or ‘Big’ data.
Appendix A

The Statistical Analysis of Misreporting on Sensitive Survey Questions:
Appendix

A.1 Standard list experiment model

The regression-based model for list experiments proposed by Imai (2011) models the response \((Y_i^*, Z_i^*)\) through the use of two sub-models. The first sub-model models the probability of an affirmative response to the sensitive item, \(Z_i^*\):

\[
g(x; \delta) = \Pr(Z_i^* = 1 | X_i = x; \delta),
\]

(A.1)

where \(X_i\) is a vector of covariates and \(\delta\) is a vector of parameters to be estimated. The second sub-model models the probability of an individual’s response to the control items, \(Y_i^*\):

\[
h(y|x, z; \psi) = \Pr(Y_i^* = y | X_i = x, Z_i^* = z; \psi),
\]

(A.2)

where \(X_i\) is a vector of covariates, \(Z_i^*\) is the latent response to the sensitive item, and \(\psi\) is a vector of parameters to be estimated.

To derive the likelihood function, note that there are four distinct response types to the list experiment. First, the response to the control items, \(Y_i^*\), is fully observed for respondents in the control group: those assigned to the control group are only asked to provide a response to the control items, and therefore \(Y_i^* = Y_i\) if \(T_i = 0\). Second and third, for respondents in the treatment group who answer \(Y_i = 0\) or \(Y_i = J + 1\), the response to the sensitive item \(Z_i^*\) is fully observed: responding affirmatively to none of the \(J + 1\) items \((Y_i = 0)\) or responding affirmatively to all of them \((Y_i = J + 1)\) indicates that all individual items, including the sensitive item, were
answered 0 or 1 respectively. Lastly, among those who are assigned to the treatment group and whose response is greater than 0 and less than $J + 1$, responses to both the sensitive item and control items are latent. The observed-data likelihood can be derived from recognition of these response types, as shown in Table A.1. The observed-data likelihood is given by the following:

$$L(\delta, \psi; \{T_i, Y_i, X_i\}_{i=1}^{n}) = \prod_{i=1}^{n} \left\{ g(X_i; \delta) h(Y_i|X_i, 1; \psi) + (1 - g(X_i; \delta)) h(Y_i|X_i, 0; \psi) \right\}^{1-T_i} \times \left\{ g(X_i; \delta) h(Y_i-1|X_i,1;\psi) \right\}^{1(Y_i=J+1)T_i} \times \left\{ (1 - g(X_i; \delta)) h(Y_i|X_i, 0; \psi) \right\}^{1(Y_i=0)T_i} \times \left\{ g(X_i; \delta) h(Y_i-1|X_i,1;\psi) + (1 - g(X_i; \delta)) h(Y_i|X_i, 0; \psi) \right\}^{1(0<Y_i<J+1)T_i} \times \left\{ (1 - g(X_i; \delta)) h(0|X_i,0; \psi) \right\}^{1(Y_i=0)T_i}. \quad (A.3)$$

For optimization, Imai (2011, 410-411) proposes and implements an expectation-maximization (EM) algorithm (Blair et al., 2016), details of which can be found in Imai (2011) and Blair and Imai (2012).

Similar to the goal of the main article, Blair and Imai (2012) also propose examining social desirability bias by fitting two separate list-experiment and direct-question regression models and using the fitted models to generate predicted probabilities to compare social desirability bias across groups (for details, see Blair and Imai, 2012, 54 & 60-62). This procedure is different from that outlined in the main article in that it does not seek to explicitly model inconsistency in respondents’ answers.

### A.2 Proposed method for alternative case

In the following, I lay out the proposed method for the case in which not answering affirmatively to the sensitive item is to provide the socially unacceptable response. For this case, we can identify three response patterns to the list experiment and direct question that define the response ($Z_i^*, D_i$). Substantively, these responses define the following respondent types: (1) those who hold the sensitive belief and misreport it when asked directly ($Z_i^* = 0, D_i = 1$); (2)
those who hold the sensitive belief and do not misreport it \((Z^*_i = 0, D_i = 0)\); and (3) those who
do not hold the sensitive belief and do not misreport it \((Z^*_i = 1, D_i = 1)\). These response patterns
and their respective descriptions are presented in Table A.2 (analogous to Table 1 in the main
article).

The individual likelihoods for each response pattern are provided in Table A.3, where the
observed-data model likelihood is the product of the relevant individual likelihoods as given in
the last column of the table. Models \(g(\cdot), j(\cdot),\) and \(h(\cdot)\) remain as defined in the main body of
the article. In this alternative case, note that by the monotonicity assumption, \(j(x, 1, t; \gamma) = 0\).

For the EM algorithm, the complete-data likelihood function is given by the following:

\[
L(\delta, \gamma, \psi; \{T_i, Y_i, X_i, Z^*_i, U^*_i\}_{i=1}^n) = \prod_{i=1}^n \left\{ g(X_i; \delta)[1 - j(X_i, 1, T_i; \gamma)] h(Y_i - T_i|X_i, 1, 0; \psi) \right\}^{1(Z^*_i = 1 \land U^*_i = 0)} \times \left\{ [1 - g(X_i; \delta)] [1 - j(X_i, 0, T_i; \gamma)] h(Y_i|X_i, 0, 0; \psi) \right\}^{1(Z^*_i = 0 \land U^*_i = 0)} \times \left\{ [1 - g(X_i; \delta)] j(X_i, 0, T_i; \gamma) h(Y_i|X_i, 0, 1; \psi) \right\}^{1(Z^*_i = 0 \land U^*_i = 1)}.
\]

The expressions used to calculate the weights for the EM algorithm are given in Table A.4.

The maximization step computes the parameters \(\delta, \gamma,\) and \(\psi\) using the most recent values
of the weights, \(w^{(t)}_i\), from the E-step, that maximize the complete-data log-likelihood function

---

**Table A.2: Respondent types defined by the response \((Z^*_i, D_i)\) for the case in which responding
\(Z^*_i = 0\) is to provide the sensitive response**

<table>
<thead>
<tr>
<th>Type</th>
<th>Sensitive (Z^*_i)</th>
<th>Direct (D_i)</th>
<th>Misreport (U^*_i)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misreport sensitive</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>Respondent holds the sensitive belief but misreports it when asked directly.</td>
</tr>
<tr>
<td>Truthful sensitive</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Respondent holds the sensitive belief and states so truthfully when asked directly.</td>
</tr>
<tr>
<td>Non-sensitive</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>Respondent does not hold the sensitive belief and states so truthfully when asked directly.</td>
</tr>
</tbody>
</table>

Note that the respondent type for the response \((Z^*_i = 1, D_i = 0)\) is undefined by the monotonicity assumption.
Table A.3: Observed-data likelihood for the case in which $Z^* = 0$ is the socially undesirable response

<table>
<thead>
<tr>
<th>Observed variables</th>
<th>Latent variables</th>
<th>Observed-data likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_i$</td>
<td>$Y_i$</td>
<td>$D_i$</td>
</tr>
<tr>
<td>0</td>
<td>$Y_i$</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>$Y_i$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>$0 &lt; Y_i &lt; J + 1$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>$J + 1$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>$J + 1$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>$Y_i &lt; J + 1$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table A.4: Weights $w_i^{(s)}$ as calculated in the E-step for the case in which $Z^* = 0$ is the socially undesirable response

\[
w_i^{\text{non-sensitive}} = \frac{g(x_i; \delta)[1 - j(x_i, 1, t_i; \gamma)] h(y_i - t_i|x_i, 1, 0; \psi) + [1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 0; \psi) + [1 - g(x_i; \delta)] j(x_i, 0, t_i; \gamma) h(y_i|x_i, 0, 1; \psi)}{g(x_i; \delta)[1 - j(x_i, 1, t_i; \gamma)] h(y_i - t_i|x_i, 1, 0; \psi) + [1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 0; \psi) + [1 - g(x_i; \delta)] j(x_i, 0, t_i; \gamma) h(y_i|x_i, 0, 1; \psi)}
\]

\[
w_i^{\text{truthful sensitive}} = \frac{[1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 0; \psi) + [1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 0; \psi) + [1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 1; \psi)}{g(x_i; \delta)[1 - j(x_i, 1, t_i; \gamma)] h(y_i - t_i|x_i, 1, 0; \psi) + [1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 0; \psi) + [1 - g(x_i; \delta)] j(x_i, 0, t_i; \gamma) h(y_i|x_i, 0, 1; \psi)}
\]

\[
w_i^{\text{misreport sensitive}} = \frac{[1 - g(x_i; \delta)] j(x_i, 0, t_i; \gamma) h(y_i|x_i, 0, 1; \psi)}{g(x_i; \delta)[1 - j(x_i, 1, t_i; \gamma)] h(y_i - t_i|x_i, 1, 0; \psi) + [1 - g(x_i; \delta)][1 - j(x_i, 0, t_i; \gamma)] h(y_i|x_i, 0, 0; \psi) + [1 - g(x_i; \delta)] j(x_i, 0, t_i; \gamma) h(y_i|x_i, 0, 1; \psi)}
\]
given as follows:

\[
Q(\delta, \gamma, \psi; \{Y_i, X_i, T_i, w_i^{(1)}\}_{i=1}^n) = \\
\sum_{i=1}^n \left[ w_i^{(\text{non-sensitive})} \left\{ \log g(X_i; \delta) + \log (1 - j(X_i, 1, T_i; \gamma)) + \log h(Y_i - T_i | X_i, 1, 0; \psi) \right\} + \\
w_i^{(\text{truthful sensitive})} \left\{ \log (1 - g(X_i; \delta)) + \log (1 - j(X_i, 0, T_i; \gamma)) + \log h(Y_i | X_i, 0, 0; \psi) \right\} + \\
w_i^{(\text{misreport sensitive})} \left\{ \log (1 - g(X_i; \delta)) + \log j(X_i, 0, T_i; \gamma) + \log h(Y_i | X_i, 0, 1; \psi) \right\}. \right]
\]  

(A.5)

A.3 Screener question

The screener question was asked as follows:

The previous question contained a list of statements.
Which of the following subjects was a part of that list?

- The power of unions
- Gay marriage
- The federal budget
- Don’t know

Respondents who did not answer “The power of unions” (7% of respondents) were removed from the dataset used in the results section.

I also check for the possibility that treatment assignment affects responses to the screener question. It may be the case that receiving a slightly longer list (5 instead of 4 items) could lead respondents to be less attentive because of the added effort and length of time required to complete the question. In both the treatment and control group however, 7% of respondents did not respond to the screener question correctly. The difference between the two groups (0.5 percentage points) is not statistically significant (p = 0.12).

A.4 Control-items and misreport sub-model results from Simulation Study 2

Figures A.1 and A.2 present results from Simulation Study 2 (section 3.2 of the main article). These figures compare the results from the proposed to the standard estimator in terms of RMSE, bias, and coverage of confidence intervals. Although the control-items sub-model is rarely, if ever, of substantive interest to researchers, we can see that the proposed estimator improves on the standard estimator in terms of RMSE. Note that Figure A.2 does not include data from any simulations for the standard estimator because it does not include a misreport sub-model.
Figure A.1: Control-items sub-model for Simulation Study 2

This figure presents data from 10,000 Monte Carlo simulations for root mean squared error (RMSE), bias, and confidence interval coverage for the control-items sub-model from Simulation Study 2.

A.5 List experiment control-items sub-model for empirical application

Table A.5 presents results from the control-items sub-models for the models presented in Table 4 in the main article. The parameters $Z^* = 1$ and $Z^* = 0$ represent indicator variables for whether a respondent answered affirmatively (or not) to the sensitive item. $U^*$ denotes an indicator variable for whether someone who holds the sensitive belief (i.e. $Z^* = 0$) also misreports it.\(^1\)

\(^1\)Note that not responding affirmatively to the statement “Women are as capable as men in politics” is the sensitive response.
This figure presents data from 10,000 Monte Carlo simulations for root mean squared error (RMSE), bias, and confidence interval coverage for the misreport sub-model from Simulation Study 2.

A.6 Tests for design-assumption violations in empirical application data

A series of preliminary checks are run on the data used in the empirical section to test for violations of the list experiment’s design assumptions. I first test for a “design effect,” which refers to a difference in responses to the control items ($Y^*_i$) associated with treatment assignment (for details, see Blair and Imai, 2012, 63-65). The test for the presence of a design effect proposed by Blair and Imai (2012) shows no strong evidence of one ($p = 1$).
Second, I check for violations of the monotonicity assumption. Violations of this assumption are immediately apparent for respondents in the treatment group who respond that they agree with none of the items in the list experiment question, but answer affirmatively to the direct question. Such a response pattern would indicate that the socially unacceptable response is given in the list experiment, but the socially acceptable response is given to the direct question. Three out of 11,133 respondents in the treatment group provide this response pattern and are removed from the dataset. I then compare list experiment responses among those who openly admit to holding the socially unacceptable response. Among the group of respondents who
provide this response (“No”) to the direct question, the difference in the mean response to the list experiment question between treatment and control groups should be 0 in expectation: these respondents will, by the monotonicity assumption, not respond affirmatively to the sensitive item in the list experiment and therefore responses to the list experiment should not depend on treatment assignment.\(^2\) Testing for a difference in the mean response to the list experiment question between the control and treatment group for those who provide the socially unacceptable response to the direct question does not provide strong evidence of an assumption violation (\(p = 0.07\)).

Lastly, I test whether those who are assigned to the treatment group respond systematically differently to the direct question than those in the control group. As noted previously, to avoid this problem, the list experiment and direct questions were separated from each other by a large number of unrelated questions. There is no strong evidence that treatment assignment affects responses to the direct question (\(p = 0.77\)).

\(^2\)This test is analogous to that proposed by Aronow et al. (2015, 50-51) (Placebo Test I), which tests for a difference of 1 between control and treatment groups among those who respond affirmatively to the direct question for the case in which responding affirmatively to the sensitive item is to provide the socially unacceptable response.
Appendix B

How Should We Measure Citizen Ideology with Ideological Placement Scales? Appendix

B.1 Comparing Hare et al. (2015) replication data to the model that incorporates ideological self-placement

In this section, I compare point estimates of respondent ideology from the model proposed by Hare et al. (2015) to those from a linear AM scaling model that incorporates ideological self-placement in the model itself. This comparison is the same as presented in Figure 2 from the article, but uses the model proposed by Hare et al. (2015) using their replication data and model output. These models are not perfectly comparable because, as noted in footnote 10 of the article, the model proposed by Hare et al. (2015) includes an additional parameter, $\sigma_i$, that accounts for error in item placement that might differ at the individual level, a complexity that I do not address. Hare et al. (2015) also use slightly different contraints and priors for identification. Nevertheless, I compare the model proposed by Hare et al. (2015) to one that includes ideological self-placement to show that the results of the models compared in Figure 2 of the article are qualitatively similar to using Hare et al. (2015)’s replication data.

Figure B.1 presents this comparison. As expected, the results are similar to those presented in Figure 2 in the main article: in the model that includes ideological self-placement, estimates of respondent ideology are roughly similar in the middle of the distribution, but diverge greatly in the tails. As with the data presented in Figure 2 of the main article, values of ideology for some respondents in the replication data from Hare et al. (2015) are extreme, with some values of ideology from the posterior in the hundreds and thousands. Therefore, as shown in the main article both graphically in Figure 2 and in terms of predictive performance (as shown in Table 1),
Figure B.1: Comparison of Hare et al. (2015) AM scaling model and model including self-placement

The model that includes self-placement performs substantially better than that which does not.

B.2 Heterogeneous and homogeneous linear and ordinal AM scaling models

Although the models presented in the main article allow the error in item placement to vary per item, it is useful to examine whether estimates from this more flexible (and more theoretically satisfying model) are substantially different empirically. To test this, I fit linear and ordinal AM scaling models to the 2010, 2012, 2014, and 2016 CCES data, but constrain the scale parameters...
to be equal across items, such that $\sigma_{j=1} = \sigma_{j=2} = \ldots = \sigma_{j=J}$. In the linear AM scaling model, I allow this scale parameter, $\sigma$, is estimated from the data. In the ordinal AM scaling model, I set $\sigma$ to 1 for the purpose of model identification — a standard constraint for identifying ordinal regression models.

To examine the heterogeneous and homogeneous AM scaling models, I present a comparison of point estimates of ideology for each political actor and each respondent in Figures B.2 and B.3 respectively. As these figures make clear, estimates of the political ideology of political actors and of respondents are extremely similar between models, with correlations between models that are extremely high. Thus, although allowing error in ideological placement to vary across items is theoretically satisfying, the empirical consequences are nevertheless minimal. This finding supports similar conclusions by Aldrich and McKelvey (1977) and Palfrey and Poole (1987), who find through simulation that allowing error in placement to vary among individuals does not substantially change estimates from the AM scaling model. Given that fitting the more complex model is more computationally demanding, researchers may therefore reasonably choose to avoid this complication with few consequence for their point estimates both among respondents and for political actors.

### B.3 Political actors included in CCES models

In this section, I describe the items included in the AM scaling models that use CCES data. CCES data contain some ideological placement questions that are asked of all respondents, such as national political actors, and some that are asked of a subset of respondents, such as governors, senators, and congressional district representatives. In all models applied to CCES data in the article, the placements included were those for all governors and senators serving at the time of the survey, as well as the items placed by all respondents, which are the following:

<table>
<thead>
<tr>
<th>Survey</th>
<th>Items received by all respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCES 2012</td>
<td>Democratic Party, Republican Party, Barack Obama, Mitt Romney, The Tea Party, Supreme Court</td>
</tr>
<tr>
<td>CCES 2014</td>
<td>Democratic Party, Republican Party, Barack Obama, Hillary Clinton, Ted Cruz, Rand Paul, Jeb Bush</td>
</tr>
<tr>
<td>CCES 2016</td>
<td>Democratic Party, Republican Party, Barack Obama, Hillary Clinton, Donald Trump, Supreme Court</td>
</tr>
</tbody>
</table>
Figure B.2: Comparison of heterogeneous and homogeneous ordinal AM scaling models (political actor ideology $\zeta_j$)
Figure B.3: Comparison of heterogeneous and homogeneous linear AM scaling models (respondent ideology $\theta_i$)
Appendix C

Does International Terrorism affect Public Attitudes toward Refugees? Evidence from a Large-scale Natural Experiment: Appendix

C.1 Survey data and sample covariate balance

Survey data The survey data were collected by the public opinion research firm Vox Pop Labs as part of its 2015 Canadian federal election post-election survey, a survey sent to all members of its national online panel. In total, 64,677 respondents completed the post-election survey, of which roughly 30% (n = 18,763) received questions regarding Syrian refugees.1 Because the survey was sent to all members of the research firm's online panel, respondents were not purposefully selected to match the socio-demographic characteristics of the national population. Estimates presented in the Results section are therefore statistically adjusted through regression and survey weighting to match the population's demographic and geographic distribution as indicated by the 2011 national census. Weighting variables include gender, age, education, mother tongue, and region of residence.

Covariate balance To examine covariate balance, Table C.1 presents descriptive statistics for the sample of survey respondents who were invited and responded to the post-election survey less than two days before the November 13, 2015 Paris terrorist attacks and the sample of respondents who were invited and responded less than two days afterward. The fourth column

---

1The size of the survey was the result of the firm's focus on social and political data, and the need to profile respondents for future social and political studies.
This table shows the mean level of pre-treatment covariates (political ideology, gender, age, education, region, mother tongue) from the survey sample collected less than 2 days before the attacks (n = 1,155) and that collected less than 2 days after the attacks (n = 2,465). The fourth column displays the normalized difference in means between the two samples. Political ideology is an ideological self-placement scale where 0 indicates left-wing and 10 indicates right-wing.

The normalized difference between the two samples for each covariate (Imbens and Rubin, 2015, 310-313), calculated as follows:

$$
\Delta \equiv \frac{\mu_{t=0} - \mu_{t=1}}{\sqrt{(\sigma^2_{t=0} + \sigma^2_{t=1})/2}},
$$

where $\mu_{t=0}$ and $\mu_{t=1}$ denote the sample means calculated from the data collected before ($t = 0$) and after ($t = 1$) the attacks, and $\sigma_{t=0}$ and $\sigma_{t=1}$ denote the respective standard deviations.

As is clear from Table C.1, sample characteristics before and after the Paris attacks appear extremely similar across all observed pre-treatment covariates, showing no evidence of meaningful differences in survey non-response. To test this further, we regress an indicator variable that represents responding to the survey invitation after the attacks on the full set of pre-treatment covariates. As shown in Table C.2, none of the model parameters for these covariates significantly differentiate respondents in the pre- and post-attacks samples, providing further evidence that non-response among the random sample of respondents surveyed after the attacks does not meaningfully differ from that of respondents sampled before the attacks.² Furthermore, a

²Note that a parameter for the variable Survey language: French is included in the regression specification because the proportion of survey invitations to panelists who take surveys in the French language was sampled.
Table C.2: Regression results to examine sample composition before and after the Paris attacks

<table>
<thead>
<tr>
<th></th>
<th>Response after Paris attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>−0.003 (0.076)</td>
</tr>
<tr>
<td>Age 30-39</td>
<td>−0.132 (0.130)</td>
</tr>
<tr>
<td>Age 40-49</td>
<td>0.023 (0.142)</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>−0.095 (0.118)</td>
</tr>
<tr>
<td>Age 65+</td>
<td>−0.025 (0.123)</td>
</tr>
<tr>
<td>College</td>
<td>0.053 (0.123)</td>
</tr>
<tr>
<td>University degree</td>
<td>0.113 (0.108)</td>
</tr>
<tr>
<td>Francophone</td>
<td>0.157 (0.157)</td>
</tr>
<tr>
<td>Other language</td>
<td>−0.024 (0.137)</td>
</tr>
<tr>
<td>Ontario</td>
<td>−0.159 (0.172)</td>
</tr>
<tr>
<td>Quebec</td>
<td>−0.289 (0.229)</td>
</tr>
<tr>
<td>West</td>
<td>−0.039 (0.174)</td>
</tr>
<tr>
<td>Political ideology</td>
<td>0.016 (0.016)</td>
</tr>
<tr>
<td>Survey language: French</td>
<td>−0.578** (0.204)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.964*** (0.208)</td>
</tr>
</tbody>
</table>

This table presents results from a logistic regression model where the outcome variable indicates inclusion in the sample collected within two days after the Paris attacks (compared to the sample collected within 2 days before the attacks). *p < .05; **p < .01; ***p < .001

The likelihood ratio test between a model with and without pretreatment covariates does not provide evidence that the covariates jointly differentiate respondents in the pre- and post-attack samples (p = 0.74).

We note also that the invitation to the survey referred to it as a post-election study, and did not reference either the Paris attacks or refugees. We can therefore plausibly expect that no respondents in the immediate post-attack period would have responded (or not responded) to the survey invitations based on the inclusion of questions regarding refugees. Furthermore, within-survey non-response to the questions regarding refugees is extremely low (~0.5%) and there is no evidence that non-response to the refugee questions varied between the pre- and post-attack periods (p = 0.57). In sum, both theoretically, given the construction of the survey, and empirically, given the comparison of observed pretreatment variables, there is a strong case that responding to the survey before and after the attacks is independent of treatment assignment.

---

3As noted in the main article, the survey questions were embedded in a substantially larger post-election study.
Table C.3: Support for refugee resettlement ordinal logistic regression results

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paris attacks</td>
<td>−0.301</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Female</td>
<td>0.190</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Age 30-39</td>
<td>−0.215</td>
<td>(0.116)</td>
</tr>
<tr>
<td>Age 40-49</td>
<td>0.089</td>
<td>(0.126)</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>0.102</td>
<td>(0.105)</td>
</tr>
<tr>
<td>Age 65+</td>
<td>−0.026</td>
<td>(0.109)</td>
</tr>
<tr>
<td>College</td>
<td>−0.197</td>
<td>(0.108)</td>
</tr>
<tr>
<td>University degree</td>
<td>0.435</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Francophone</td>
<td>−0.263</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Other language</td>
<td>−0.292</td>
<td>(0.110)</td>
</tr>
<tr>
<td>Ontario</td>
<td>−0.537</td>
<td>(0.160)</td>
</tr>
<tr>
<td>Quebec</td>
<td>−0.918</td>
<td>(0.180)</td>
</tr>
<tr>
<td>West</td>
<td>−0.539</td>
<td>(0.161)</td>
</tr>
<tr>
<td>Political ideology</td>
<td>−0.384</td>
<td>(0.016)</td>
</tr>
</tbody>
</table>

κ₁  −5.260         (0.219)
κ₂  −4.549         (0.213)
κ₃  −4.079         (0.210)
κ₄  −3.227         (0.206)
κ₅  −2.006         (0.201)

N  3,531

κ denote cut-point parameter estimates. *p < .05; **p < .01; ***p < .001

C.2 Public support for refugee resettlement regression results

For reasons of space, the ordinal regression results provided in Table 2 of the main article do not include parameter estimates for cut points. The full regression result including cut-point estimates is presented in Table C.3.

C.3 Anxiety and Anger

Because some argue that anger and anxiety are measurably distinct theoretical concepts (Druckman and McDermott, 2008; Huddy, Feldman and Casese, 2007; Petersen, 2010), we also examine the effect of the attacks on anxiety after excluding the indicator ‘anger’ from the anxiety index (including ‘anger’: α = 0.93; excluding ‘anger’: α = 0.93). Figure C.1 shows these data graphically. As we can see, the effect of the attacks and its duration are effectively equivalent on both indexes.
Furthermore, regression results to estimate the effect of the attacks, presented in Table C.4, show substantively similar results.

C.4 Economic threat

For completeness, we now also examine the effect of the attacks on perceptions of Syrian refugees as an economic threat. We expect that the attacks will have little, if any, effect on the public’s perceptions of refugees as a threat to the economy. Theoretically, it is unlikely that large-scale terrorist attacks would provide a meaningful signal to the public concerning the threat to the domestic economy posed by refugee resettlement. If terrorism does increase negative attitudes toward refugees overall, however, the attacks may also affect evaluations of any threat concerning refugees more generally, regardless of their type. To investigate this, we examine the economic threat index concerning Syrian refugee resettlement, and, for comparison, two retrospective indicators of economic evaluations unrelated to refugee resettlement, which should not be affected by the attacks otherwise. The latter indicators are questions regarding (1) evaluations of respondents’ personal financial circumstances over the past year (worse, same, better), and (2) evaluations of the national economy in the past year (worse, same, better) (for complete question text, see Appendix C.8.).

The raw data for the economic threat index is presented graphically in Figure C.2 and those for the retrospective national economic and personal financial evaluations are presented in Figure C.3. As Figure C.2 shows, there appears to be a slight increase in perceptions of Syrian refugee resettlement as a threat to the economy. However, as shown in Table 1 of the main article, the difference in economic threat perceptions between the immediate pre- and post-
Table C.4: Regression results comparing anxiety indexes including and excluding anger indicator

<table>
<thead>
<tr>
<th></th>
<th>Anxiety (excl. ‘anger’)</th>
<th>Anxiety</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>Paris attacks</td>
<td>0.265*** (0.054)</td>
<td>0.271*** (0.055)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.050 (0.053)</td>
<td>0.062 (0.053)</td>
<td></td>
</tr>
<tr>
<td>Age 30-39</td>
<td>0.089 (0.088)</td>
<td>0.089 (0.088)</td>
<td></td>
</tr>
<tr>
<td>Age 40-49</td>
<td>−0.003 (0.095)</td>
<td>−0.004 (0.095)</td>
<td></td>
</tr>
<tr>
<td>Age 50-64</td>
<td>0.030 (0.081)</td>
<td>0.035 (0.081)</td>
<td></td>
</tr>
<tr>
<td>Age 65+</td>
<td>0.107 (0.083)</td>
<td>0.105 (0.083)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>0.176* (0.086)</td>
<td>0.181* (0.086)</td>
<td></td>
</tr>
<tr>
<td>University degree</td>
<td>−0.189* (0.075)</td>
<td>−0.178* (0.075)</td>
<td></td>
</tr>
<tr>
<td>Francophone</td>
<td>0.120 (0.093)</td>
<td>0.138 (0.093)</td>
<td></td>
</tr>
<tr>
<td>Other language</td>
<td>0.056 (0.085)</td>
<td>0.034 (0.085)</td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>0.221 (0.120)</td>
<td>0.223 (0.121)</td>
<td></td>
</tr>
<tr>
<td>Quebec</td>
<td>0.120 (0.136)</td>
<td>0.115 (0.136)</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>0.169 (0.121)</td>
<td>0.163 (0.122)</td>
<td></td>
</tr>
<tr>
<td>Political ideology</td>
<td>0.199*** (0.011)</td>
<td>0.197*** (0.011)</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−0.985*** (0.148)</td>
<td>−0.989*** (0.148)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,688</td>
<td>1,692</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses. * p < .05; ** p < .01; *** p < .001
attack period is not statistically significant. The absence of statistical significance does not, of course, mean that the attacks did not affect perceptions of refugees as an economic threat. However, the estimated magnitude of this effect in standardized units ($\beta = 0.08$) is, as one would expect, substantially smaller than that of anxiety ($\beta = 0.27$), and security ($\beta = 0.39$) and cultural threat ($\beta = 0.19$) (cf. Figures 1 and 2 in main article). Similar to the other indexes, the economic threat index appears to decline, if slightly, to the pre-attack baseline within roughly a week of the attacks.

By comparison, the two retrospective economic indicators show no meaningful difference between the immediate pre- and post-attack period ($p = 0.74$; $p = 0.13$) and demonstrate no clear pattern that would be consistent with an effect that decays in response to the attacks similar to the indicators theoretically related to Syrian refugee resettlement. Although there is more variability in retrospective evaluations of the economy in the weeks after the attacks, we can expect this as a response to economic or related news as a basis for comparison to the previous year's economy.

### C.5 Causal mediation sensitivity analysis

The validity of estimates of average causal mediation effects relies on the sequential ignorability assumption (see Imai, Keele and Yamamoto, 2010; Imai et al., 2011). This assumption states that (1) the potential outcomes of the mediator and of the outcome are independent of treatment status conditional on pre-treatment covariates, and (2) the potential outcome of the outcome is independent of the observed mediator, conditional on treatment assignment and pre-treatment
Although the first assumption holds when treatment assignment is randomized (or is plausibly assigned as-if randomly as in a natural experiment), the second assumption cannot be tested with observed data. Imai, Keele and Yamamoto (2010, 60-62) therefore propose a sensitivity analysis procedure to quantify the robustness of average causal mediation effect estimates to the presence of an unobserved pre-treatment confounder. The procedure allows researchers to test the degree to which the sequential ignorability assumption must be violated before the average causal mechanism effect estimate would be zero.

To examine the robustness of our causal mechanism results to an unobserved pre-treatment confounder, we apply this sensitivity analysis procedure to each causal mechanism examined in the article. Results from this procedure are shown in Figure C.4. Each panel presents estimates of the average casual mediation effect in the presence of a confounder that is correlated with the mediator at levels of correlation $\rho$ from $-1$ to $1$. As the figures shows, the average causal mediation effect estimates for anxiety, security, and culture are each robust to the presence of a strong confounder. In each case, the effects are reversed only when an unobserved confounder is very highly correlated with each mediator of interest: $\rho^{(anxiety)} = -0.64$, $\rho^{(security)} = -0.71$, and $\rho^{(culture)} = -0.75$. In sum, each of the average causal mechanism effect estimates appears highly robust to the presence of unobserved pre-treatment confounders.

A second potential concern is post-treatment confounding, which in applied work is typically left unexamined, but which we account for explicitly in estimating the average causal mediation

---

4 Note that these conditional independence assumptions rely on there not being any post-treatment confounders.

5 Sensitivity analysis results were generated using the library mediation (Tingley et al., 2014) in the statistical package R (R Core Team, 2017). Note that sensitivity analysis in the mediation package is conducted in a linear model framework, whereas average causal mediation effect estimates presented in the article use an ordinal outcome model to permit estimates to be presented in terms of percentage points.
effect for cultural threat. Theoretically, cultural threat is expected to be causally affected by individuals’ perceptions of refugees as a security threat, which itself is affected by the Paris terrorist attacks.\(^6\) Post-treatment confounding can be shown diagrammatically as follows (as

---

\(^6\)As we note in the article, we remain agnostic as to whether anxiety may be also be a post-treatment confounder for threat, and vice versa. However, because the measure of anxiety and security threat were collected in two independent sub-samples of the survey, we cannot examine the causal mediation effects of each while accounting for the other as a possible post-treatment confounder.
noted in fn. 28 in the main article):

\[ W_i - M_i - Y_i, \]

where \( W_i \) denotes a post-treatment confounder (security threat) that causally affects both the mediator of interest \( M_i \) (cultural threat) and the outcome \( Y_i \) (support for refugee resettlement).

Accounting for security as a post-treatment confounder for cultural threat requires assumptions additional to those for identifying the average causal mechanism effect without such confounding. The key untestable assumption is the homogeneous interaction assumption, which states that the interaction between the primary mediator and the treatment is constant across all individuals. In other words, although the magnitude of the relationship between the mediator and outcome can flexibly depend on treatment assignment, the interaction effect is assumed to be constant across all individuals. For further details regarding accounting for post-treatment confounding, see Imai, Tingley and Yamamoto (2013).

Although violations of the homogeneous interaction assumption are not empirically testable, Imai, Tingley and Yamamoto (2013) propose a sensitivity analysis procedure to permit researchers to examine how the average causal mediation effect changes with levels of heterogeneity in the treatment-mediator interaction. In this procedure, the sensitivity parameter \( \rho \) represents the standard deviation of the individual-level treatment-mediator interaction. Because interpretation of \( \rho \) is not necessarily intuitive, Imai, Tingley and Yamamoto (2013) also allow the sensitivity analysis to be understood through a second parameter, \( \tilde{R}^2 \), which represents the proportion of variation in the outcome that would be explained by allowing for heterogeneity in the interaction term.

Results from this procedure are presented in Figure C.5. As the left panel of Figure C.5 shows, the bounds on the average causal mediation effect estimate contain zero when \( \rho = 0.0625 \). This translates to only 2.5% (\( \tilde{R}^2 \)) of the variation in the outcome (second panel of Figure C.5) needing to be explained by heterogeneity in the interaction term before the bounds on the estimate contain zero. The estimate for the cultural threat mechanism is, in other words, fragile to violations of the no-interaction assumption and should therefore be treated cautiously.

### C.6 Causal mediation outcome regression models

Table 2 in the main article provides coefficients and standard errors for the fitted first-stage model for the causal mechanisms results. In Table C.5, we provide the second-stage results. Each model is an ordinal regression model where the outcome is support for refugee resettlement. Each model is specified with a set of pre-treatment control variables as well the causal
Table C.5: Second-stage causal mechanism regression models

<table>
<thead>
<tr>
<th></th>
<th>Anxiety (1)</th>
<th>Sympathy (2)</th>
<th>Security (3)</th>
<th>Culture (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Index</strong></td>
<td>−1.495***</td>
<td>1.797***</td>
<td>−1.751***</td>
<td>−2.283***</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.072)</td>
<td>(0.061)</td>
<td>(0.075)</td>
</tr>
<tr>
<td><strong>Paris attacks</strong></td>
<td>0.112</td>
<td>−0.323**</td>
<td>0.155</td>
<td>−0.141</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.112)</td>
<td>(0.105)</td>
<td>(0.108)</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>0.363***</td>
<td>−0.208</td>
<td>0.280**</td>
<td>−0.004</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.111)</td>
<td>(0.102)</td>
<td>(0.105)</td>
</tr>
<tr>
<td><strong>Age 30-39</strong></td>
<td>−0.030</td>
<td>−0.005</td>
<td>−0.055</td>
<td>−0.236</td>
</tr>
<tr>
<td></td>
<td>(0.175)</td>
<td>(0.183)</td>
<td>(0.178)</td>
<td>(0.183)</td>
</tr>
<tr>
<td><strong>Age 40-49</strong></td>
<td>0.046</td>
<td>−0.364</td>
<td>0.144</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>(0.190)</td>
<td>(0.199)</td>
<td>(0.191)</td>
<td>(0.199)</td>
</tr>
<tr>
<td><strong>Age 50-64</strong></td>
<td>0.390*</td>
<td>−0.085</td>
<td>0.180</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td>(0.163)</td>
<td>(0.168)</td>
<td>(0.159)</td>
<td>(0.164)</td>
</tr>
<tr>
<td><strong>Age 65+</strong></td>
<td>0.324</td>
<td>−0.255</td>
<td>0.180</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>(0.168)</td>
<td>(0.173)</td>
<td>(0.167)</td>
<td>(0.170)</td>
</tr>
<tr>
<td><strong>College</strong></td>
<td>−0.163</td>
<td>−0.165</td>
<td>−0.009</td>
<td>−0.093</td>
</tr>
<tr>
<td></td>
<td>(0.168)</td>
<td>(0.171)</td>
<td>(0.160)</td>
<td>(0.165)</td>
</tr>
<tr>
<td><strong>University degree</strong></td>
<td>0.398**</td>
<td>0.315*</td>
<td>0.257</td>
<td>0.214</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.152)</td>
<td>(0.142)</td>
<td>(0.147)</td>
</tr>
<tr>
<td><strong>Francophone</strong></td>
<td>−0.327</td>
<td>−0.370</td>
<td>−0.255</td>
<td>−0.233</td>
</tr>
<tr>
<td></td>
<td>(0.184)</td>
<td>(0.190)</td>
<td>(0.191)</td>
<td>(0.198)</td>
</tr>
<tr>
<td><strong>Other language</strong></td>
<td>−0.284</td>
<td>−0.121</td>
<td>−0.419*</td>
<td>−0.393*</td>
</tr>
<tr>
<td></td>
<td>(0.173)</td>
<td>(0.177)</td>
<td>(0.166)</td>
<td>(0.172)</td>
</tr>
<tr>
<td><strong>Ontario</strong></td>
<td>−0.687*</td>
<td>−0.555</td>
<td>−0.529*</td>
<td>−0.205</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.291)</td>
<td>(0.221)</td>
<td>(0.229)</td>
</tr>
<tr>
<td><strong>Quebec</strong></td>
<td>−1.200***</td>
<td>−0.617</td>
<td>−0.913***</td>
<td>−0.491</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
<td>(0.317)</td>
<td>(0.258)</td>
<td>(0.267)</td>
</tr>
<tr>
<td><strong>West</strong></td>
<td>−0.727*</td>
<td>−0.470</td>
<td>−0.692**</td>
<td>−0.251</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.292)</td>
<td>(0.222)</td>
<td>(0.231)</td>
</tr>
<tr>
<td><strong>Political ideology</strong></td>
<td>−0.177***</td>
<td>−0.260***</td>
<td>−0.183***</td>
<td>−0.216***</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.025)</td>
<td>(0.024)</td>
<td>(0.025)</td>
</tr>
<tr>
<td><strong>κ1</strong></td>
<td>−5.769***</td>
<td>−6.563***</td>
<td>−6.265***</td>
<td>−6.869***</td>
</tr>
<tr>
<td></td>
<td>(0.378)</td>
<td>(0.386)</td>
<td>(0.336)</td>
<td>(0.356)</td>
</tr>
<tr>
<td><strong>κ2</strong></td>
<td>−4.675***</td>
<td>−5.453***</td>
<td>−5.132***</td>
<td>−5.550***</td>
</tr>
<tr>
<td></td>
<td>(0.363)</td>
<td>(0.370)</td>
<td>(0.320)</td>
<td>(0.336)</td>
</tr>
<tr>
<td><strong>κ3</strong></td>
<td>−4.115***</td>
<td>−4.871***</td>
<td>−4.178***</td>
<td>−4.450***</td>
</tr>
<tr>
<td></td>
<td>(0.357)</td>
<td>(0.362)</td>
<td>(0.311)</td>
<td>(0.325)</td>
</tr>
<tr>
<td><strong>κ4</strong></td>
<td>−2.953***</td>
<td>−3.728***</td>
<td>−2.467***</td>
<td>−2.587***</td>
</tr>
<tr>
<td></td>
<td>(0.348)</td>
<td>(0.351)</td>
<td>(0.299)</td>
<td>(0.311)</td>
</tr>
<tr>
<td><strong>κ5</strong></td>
<td>−1.097**</td>
<td>−1.832***</td>
<td>−0.630*</td>
<td>−0.588</td>
</tr>
<tr>
<td></td>
<td>(0.340)</td>
<td>(0.340)</td>
<td>(0.292)</td>
<td>(0.303)</td>
</tr>
</tbody>
</table>

| N                        | 1,661       | 1,661        | 1,795        | 1,797       |

Ordinal logistic regression models. $\kappa$ denote cut-point parameters.

*p < .05; **p < .01; ***p < .001
Figure C.5: Sensitivity analysis for homogeneous interaction assumption for cultural threat index

C.7 National media mentions of Paris attacks

In the main article, we demonstrate that the effects of the Paris terrorist attacks on public emotions and attitudes. In the conclusion, we suggest that the duration of these effects may be tied to the degree of coverage given to the attacks by the media. To provide evidence that is suggestive of this, we collected all articles from the two major national newspapers (The National Post and the Globe & Mail) and counted the frequency of articles mentioning the Paris attacks for each day following the attacks. We present these data in Figure C.6. Although the analysis of the relationship between media coverage and emotions and attitudes is a complex task, media coverage of the attacks appears to show a similar pattern to that of the duration of the effects of the attacks as shown graphically in the main article. Because this descriptive evidence is far from conclusive, and investigation of this relationship goes beyond the scope of the article, we leave further investigation to future research.

C.8 Survey question text

Below we present the wording for each question analyzed in the article.

Received by all respondents:
The Canadian government is currently considering whether to admit more refugees from Syria. Many of these refugees are [(blank), Muslims, Christians] fleeing from the civil war.

**Received by respondents in the threat branch of the survey (security, values/integration, and economic threat):**

Imagine that these [(blank), Muslim, Christian] refugees are permitted to settle in your own community. To what extent would you agree or disagree with the following:

- Their presence would be economically costly
- They would help grow the economy
- They would increase competition for jobs

**Answer categories:**

Strong disagree, Somewhat disagree, Slightly disagree, Slightly agree, Somewhat agree, Strongly agree

Imagine that these [(blank), Muslim, Christian] refugees are permitted to settle in your own community. To what extent would you agree or disagree with the following:

- They would fit well into Canadian society
- Their values would conflict with those of Canadians
• They would enrich our culture

**Answer categories:**

Strong disagree, Somewhat disagree, Slightly disagree, Slightly agree, Somewhat agree, Strongly agree

Imagine that these [(blank), Muslim, Christian] refugees are permitted to settle in your own community. To what extent would you agree or disagree with the following:

• Their presence would pose a threat to national security
• Their presence would lead me to fear for my safety
• Some would have links to terrorism

**Answer categories:**

Strong disagree, Somewhat disagree, Slightly disagree, Slightly agree, Somewhat agree, Strongly agree

**Received by respondents in the emotion branch of the survey (anxiety and sympathy):**

Imagine now that these [(blank), Muslim, Christian] refugees are permitted to settle in your own community. To what degree do you feel the following toward them:

• Sympathy
• Indifference
• Compassion
• Sadness
• Distress

**Answer categories:**

0 (None at all), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 (A great deal)

When you think about these [(blank), Muslim, Christian] refugees settling in your community, to what degree do you feel the following:
• Anxiety
• Pride
• Upset
• Worry
• Anger
• Hope
• Fear

Answer categories:
0 (None at all), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 (A great deal)

Received by all respondents:

If it were up to you, would you agree or disagree that these [(blank), Muslim, Christian] refugees should be permitted to settle in [Canada / your own community]?

Answer categories:
Strong disagree, Somewhat disagree, Slightly disagree, Slightly agree, Somewhat agree, Strongly agree

Would you consider contacting your Member of Parliament regarding this issue?

Answer categories:
No, Yes

If these [(blank), Muslim, Christian] refugees were permitted to settle in your own community, would you be willing to donate to a program that would help them integrate?

Answer categories:
No, Yes
Questions regarding national economy and personal financial circumstances

Over the past 12 months, do you think Canada’s economy has become worse, better, or stayed about the same?

Answer categories:
Worse, Stayed about the same, Better

Over the past 12 months, has your own economic situation and that of your family become better, worse, or stayed about the same?

Answer categories:
Worse, Stayed about the same, Better
Bibliography


Canadian Broadcasting Company. 2015a. “Justin Trudeau Justifies Refugee Delay, Says Liberals Want it Done Right.”.


Coyne, Andrew. 2015. “Trudeau cabinet should be based on merit, not gender.” The National Post, November 2.


Schlosser, Edward. 2015. “I'm a liberal professor, and my liberal students terrify me.” Vox, June 3.


Starr, Terrell Jermaine. 2015. “There's a good reason protests at the University of Missouri didn't want the media around.” The Washington Post, November 11.


Taub, Amanda. 2015. “I was a liberal adjunct professor. My liberal students didn't scare me at all.” Vox, June 5.


Werner, Erica. 2015. “Syrian Refugees on Congress’ Radar.”

