Deep Divides: Experiments in Public Opinion Toward and Among Minority Groups in the United States and Canada

Anja Kilibarda

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

© 2021

Anja Kilibarda

All Rights Reserved

Table of Contents

Acknowledgments . . . . **iii**

Introduction **1**

Chapter 1: Ideological Conflict and Attitudes Toward Immigrants in the United States and Canada **13**

1 Introduction 13
2 Economic and Cultural Threat Hypotheses 15
3 Political Ideology and Attitudes Toward Outgroups 17
4 Study Design 21
  4.1 Data 21
  4.2 Experimental Design 22
  4.3 Estimation and Expectations 25
5 Results 27
  5.1 Ideological Conflict in the United States 27
  5.2 Ideological Conflict in Canada 30
  5.3 The Structure of Preferences Over Ideology 34
6 Conclusion 41

Chapter 2: Lost in Translation: How Language Compounds Wording Effects in Multilingual Surveys **45**

1 Introduction 46
2 Language Theory in Social Psychology 48
  2.1 Studying Language Effects Beyond Bilinguals 50
  2.2 Wording, Language, and Racial Affect 51
  2.3 English and French in Canada 53
3 Experimental Design 55
4 Results 57

4.1 Wording Effects in English and French 57
4.2 Cultural-Linguistic Heterogeneity 59
4.3 What If? An Observational Correction Based on Experimental Results 63
5 Discussion 65

Chapter 3: Celebrity Messaging to Counter Vaccine Hesitancy: A Study Among Black Americans During the Covid-19 Pandemic 69
1 Introduction 70
2 Countering Vaccine Hesitancy 72
2.1 The Messenger (May) Matter 73
2.2 Celebrity Messaging in Low Trust Environments 76
3 Study Design 77
3.1 Survey and Experimental Manipulation 77
3.2 Bundling and Defining the Treatment 79
3.3 Outcomes and Estimation 80
4 Results 83
4.1 Black Public Opinion and Covid-19 83
4.2 Does Race-Concordant Celebrity Messaging Work? 85
4.3 Heterogeneity and Backfire Effects 89
5 Conclusion 90

Appendix: Chapter 1 108

Appendix: Chapter 2 118

Appendix: Chapter 3 121

Acknowledgments

This dissertation was a labour of love as much as it was a labour of the spectrum of human emotion, from joy at finding (some of) my intuitions are indeed borne out by evidence, to relief at having finally designed coherent experimental projects, to misery at the many points when I believed I didn't have the strength to finish it, no less in light of the multitude of challenges that seemed to rear their proverbial ugly heads at every turn. So it goes. But so it also triumphantly ends.

I would first like to thank my advisors Drs. Robert Y. Shapiro and Donald P. Green for their expert guidance through what is an exceptionally difficult and complicated endeavour. Along with my future co-author, Dr. Robert S. Erikson, they pushed me and challenged me in all the right ways to become an incomparably better scholar, with razor-sharp insights, attentiveness, and endless support over six years. I have been in academia long enough now to know that finding advisors who are as intelligent, capable, and available as they are kind, empathetic, and genuinely caring is no small feat. The three of you are the smartest people I have ever known, and I am in constant awe.

Next, I would like to thank my mother, Maja Kilibarda, and baby sister, Andrea Kilibarda, for being the complex and hilarious and endlessly supportive human beings that they are. I can say with certainty that I wouldn't be where I am without you. We're a bit of a mess but, as we say, we at least always keep it interesting! My mother taught me the grit and resilience necessary to keep going even when things seem utterly doomed and my sister taught me that pure and untainted love is absolutely possible. They are powerhouses in their own right. Thank you also to Robert Turzo, for always being there in our most joyful and most challenging times; you are incredibly special.

We move on now to the thing I am notorious for: having a lot of friends. Here are my most special ones and to whom I also dedicate this feat of 6 long years and roughly 100 pages. First, to my oldest friends: Olivia Gagnon, Barbora Racevičiūtė, and Sonja Klarić. Sonja's boundless support since I was literally 6 months old culminated in her chauffeuring me around the entire Balkans to scope out a dissertation passion project whose demise was sadly met by the arrival of the pandemic. But the memories made during that summer and the dozens of others we've spent together are my fondest. Throughout her most challenging years, Barbora never failed to show up for me exactly as she did in the half-dilapidated homes we shared during our undergraduate years. And Olivia Gagnon is my second sister, full stop. The fact that we are not blood related at this point is both baffling and irrelevant. I love you. To my Ph.D. mates, Oliver McClellan, Julia Rubio, Giovanna Invernizzi, and Anna Wilke, I cannot possibly thank you enough for both always challenging me to be intellectually sharper while simultaneously holding my hand when I cried and always, always keeping me laughing. To Henrike Dessaules, I'm so impossibly glad I found you through this program-adjacent. Your ability to carry me through the messiest and brightest of times throughout all these years has made me into the mostly put-together girl I am today. To my New York darlings, Ross and Rosalind and Quincy Graham, Trevor Moppin, Gayatri Kawlra, whom I met in so many strange and delightful ways, you are all so different and yet your strength and kindness made me certain that there was no better place for me than here with you in this big, jumbled city. Finally, to James Nathaniel Campbell, my sweetest love, who came in to usher me over the finish line with so much tenderness, and who showed me I am deserving of all the things I have achieved. I love you to the most profound depths.

Introduction

This dissertation examines three different subjects underpinned by one common approach—the survey experiment—and, broadly, one common aim: to better understand heterogeneity in public opinion in the United States and Canada. Specifically, it focuses heterogeneity as it relates to minorities and the cultural dynamics that emerge in multiracial and multiethnic countries. Contexts with diverse racial and ethnic compositions, diverse immigration and equity policies, and complex sociohistorical lineages are bound to be underpinned by deeply fragmented attitudinal dynamics. Yet only recently has research taken a deep dive into what the contours of this fragmentation might look like. As diversity increases in the West and cultural complexities deepen, understanding heterogeneity in public opinion toward and among different cultural, racial, and ethnic groups will become increasingly pressing. Luckily for the research community, the ability to study such heterogeneity is increasing as well. Fielding large-scale surveys has been facilitated by both the vast penetration of the Internet in the 21st century and the explosion of online marketplaces that allow researchers to buy survey respondents relatively cheaply and quickly.[1] This dissertation exploits these contextual developments to field three online survey experiments among a total of 40,000 respondents in Canada and the United States.

[1]Though, it must be noted, using data from a non-probability sample comes with its own challenges when trying to make inferences to the general population (see, for example: Dutwin and Buskirk, 2017).

Survey Experiments and Incremental Learning

With that understanding, I will begin by purposefully putting the cart ahead of the proverbial horse. Before discussing the motivation and theoretical underpinnings of studying opinion heterogeneity with regard to racial and immigrant minorities, it is instructive to lay out why survey experiments are particularly well-suited to conducting such research. That is, to explain what survey experiments do, on a psychological level, that allow researchers access to complex individual orientations and how they facilitate, on a practical level, meaningful conclusions about such orientations. First, how might we conceptualize treatments in survey experiments? This dissertation conceives of them less as interventions (though in conforming with current social scientific language the included papers use this term on occasion) and more as invitations to respondents to focus more on certain elements over others when constructing an opinion on an issue or object. Even the earliest student of political science would recognize this conceptualization as, essentially, "framing" (see: Chong and Druckman, 2007). Certain information is highlighted and thereby brought to "top of mind" for respondents, with the idea that they will then use this information in their subsequent evaluations (Gamson and Modigliani, 1989; Nelson and Kinder, 1996).

While the three experiments presented in this dissertation are not framing experiments by definition, all three, I suspect, operate in basically that way: by bringing certain considerations to bear over others when respondents think about the outcome question they are to answer. In the first paper, people are encouraged to consider how the characteristics of putative immigrants might bear on how they feel about said immigrants. In the second, English and French Canadians are encouraged to think about how "positively" or "negatively" they feel toward, or how much they "like" or "dislike" racial minorities. And in the third, Black Americans are encouraged to consider whether a race-concordant celebrity communicating vaccine information is a compelling messenger of such information.

Heterogeneity and the Contours of Public Opinion

Returning to the substantive focus of the dissertation—heterogeneity in public opinion when it comes to racial, ethnic, and immigrant minorities — it is, if nothing else, at least poetic that survey experimentation in the United States largely came about to better understand attitudes toward these groups. Studies like the earliest ones by Sniderman et al. (1993) and Sniderman et al. (2002) created the overarching infrastructure of what we know about intergroup attitudes. More recent work has delved deeper to complexify our understandings of attitudes both towards and of minorities. As population heterogeneity has increased, the need to study the implications of that heterogeneity has increased concomitantly. Foundational experiments illustrated what invocations, or treatments, matter or are effective. The aim of this dissertation, building on more recent and granularly-oriented work, is to help us understand *how* these treatments might matter, *under what conditions*, and to *whom*. The first two studies are direct and clear extensions of, on the one hand, existing work on a subject and, on the other, an existing matter of survey fact. The third emerged more from intuition derived from personal experience than from theoretical motivation. As Sniderman (2018) once put it: "I say this by way of acknowledgment, not by way of apology" (p.266). Nevertheless, extensive research, largely from the medical community, theoretically underpins the work and motivates its hypotheses. Ultimately, none of the three papers makes silver bullet contributions nor do they aim to. Instead, they intend to deepen our understanding of the subtle contours that can shape opinion about racial and immigrant minorities and that of those minorities. At a meta level, the dissertation conceives of itself as operating under a framework that understands scientific inquiry as a progressive updating process—where each study contributes to our posterior distribution of information, to put it in Bayesian terms, or casts a flashlight over an unexplored corner of a dark room, to put it in accessible terms; never conclusive, but always collaborative. Seizing on ever-expanding quantities of data and

ability to assess hyper-local differences is the specific collaborative endeavour to which I aim to contribute; the world is rife with endless sources of heterogeneity currently obfuscated, at least practically speaking, by insufficient data. Future work might continue exploring heterogeneity at localized geographic levels, across nationalities – Cuban Americans, versus Dominican Americans, versus Guatemalan Americans, for example – or across more choice-based divides, such as lifestyle. Without further ado, then, the following sections introduce the three papers that constitute the dissertation.

Ideological Conflict and Attitudes Toward Immigrants in the United States and Canada

One way in which survey experiments are particularly well-suited to probing heterogeneity is their flexibility in testing multiple hypotheses simultaneously. The first dissertation paper leverages a conjoint experimental framework to extend the findings of Hainmueller and Hopkins's (2014) work evaluating competing hypotheses about the underpinnings of Americans' attitudes toward immigrants. Large-scale immigration to Western countries in recent decades has led to significant debate about what factors are more or less likely to make native-born citizens welcoming of newcomers. To that end, Hainmueller and Hopkins (2014*a*), like Wright, Levy and Citrin (2016), evaluate how the characteristics of immigrants — those corresponding to socioeconomic and sociocultural theories— affect Americans' opinions toward them. These characteristics are all ascriptive, but may also signal immigrants' cultural values. And indeed, successful integration requires that majorities to some degree accept immigrants' cultural values. So do they? This paper probes the extent to which agreement or disagreement with indicators of putative immigrants' political values affect native-born Americans' and Canadians' attitudes toward immigrants. First, respondents to an online survey in both countries report their own attitudes on a series of salient political issues, as well as their party identification. Then, relying on a conjoint design, they see pro-

files of putative immigrants applying for permanent residency (or citizenship, randomized) comprised of some ascriptive characteristics, but also immigrants' (fictitious) 'responses' to the same political questions respondents had answered themselves. The analysis estimates the degree to which agreeing or disagreeing with an immigrant's political position affects a respondent's evaluation of that immigrant, and how that estimate compares to the effect of ascriptive characteristics. The results demonstrate that attitudinal conflict between citizens and immigrants predicts negativity toward immigrants to an equal or greater extent than the ascriptive characteristics typically found to generate anti-immigrant sentiment. Moreover, despite conservatives' reputation for intolerance, people across the political spectrum are equally willing to express negative attitudes towards immigrants who disagree with them about issues tapping political values. Instead of contributing to a more general theory of attitudes toward immigrants, the study relies on heterogeneous treatment effects to generate a more granular understanding of the multifaceted factors that affect such attitudes. This understanding is increasingly relevant as the foreign-born population in both Canada and the United States increases and as immigrants' political opinions become more prominent in political debates. Indeed, already by 2020, media coverage of the presidential election was consumed by pundits trying to understand the "unexpected" wave of support for Donald Trump on the part of Cuban-heritage Americans.[2] Methodologically, this type of survey experiment allows us to understand not only whether a certain hypothesis outperforms the null, but also how it stacks up against other credible contenders. That said, measuring the subtleties of attitudes toward immigrants, and any other group in society, requires not only reflecting on the multitude of factors that might shape such attitudes, but also how questions intended to measure such attitudes are actually constructed. The second paper takes up this issue in the Canadian context.

[2]Of course, not all Cuban-heritage Americans are foreign-born, but a little over half are (Noe-Bustamante, Flores and Shah, 2019).

Lost in Translation: How Language Compounds Wording Effects in Multilingual Surveys

Measuring public opinion toward minorities has a long history in Canada, like in the United States, because it has long been a diverse society. Since the 1960s, the Canadian Election Study (CES) has contained the primary measure of record — originally, just a simple survey question asking how positively or negatively, on a scale of 0 to 100, respondents feel toward racial minorities. Indeed, the CES has certainly been the survey source that has informed the greatest majority of research on Canadians attitudes since its inception. In 1997, however, it made a slight change to the way the racial minorities feeling thermometer question was posed. Instead of asking how positively or negatively people felt toward racial minorities, it asked how much they *liked* or *disliked* them. Incidentally, this change also corresponded to a precipitous and unprecedented drop in the positivity of Canadians' evaluations. To make matters more complicated, since the survey is administered in both French and English, and the change was made in both languages, there was a differential drop among francophone and anglophone respondents.

The second dissertation paper aims to understand whether this drop was the product of a genuine souring of attitudes toward racial minorities in that year, and a heterogeneous one across linguistic groups at that, or whether some portion of it can be attributed to the CES wording change. By way of evaluation, I conducted a large-scale experiment randomly assigning respondents to either the positive/negative wording of the question or the like/dislike wording in order to detect potential wording effects. The randomization was done among both anglophones and francophones. The results demonstrate that such a wording change induces a shift in evaluations of racial minorities, but that the effect is not the same for English and French survey takers. Moreover, the effect is heterogeneous by the specific racial group asked about and by respondents' pre-treatment levels of tolerance. These findings suggest

that changes in wording matter and that they are compounded by language in ways that cannot necessarily be accounted for by mere semantic differences between languages. While most studies of language effects focus on bilinguals, this paper offers some insight into the conditions under which language can matter to public opinion when the sample is not necessarily bilingual—a much broader proportion of most populations. In linguistically diverse societies, interviewing survey respondents in different languages is important to generating representative impressions of public opinion. Understanding how and the conditions under which language can compound other survey effects is critical to developing more valid and reliable survey instruments in rapidly diversifying contexts. Within these contexts, however, it is also imperative to study not only the opinions of majorities toward minorities, but also the unique dynamics that may guide the opinions of minorities as well. The dissertation pivots in the third paper to evaluate such dynamics.

Celebrity Messaging to Counter Vaccine Hesitancy: A Study Among Black Americans During the Covid-19 Pandemic

In what has now become an essentially comical platitude, 2020 brought with it the advent of "unprecedented times" by way of the covid-19 pandemic. Near the end of the year, however, a vaccine to inoculate against covid-19 had been developed and by the beginning of 2021, three vaccines had been approved by the FDA for emergency use in the United States. As an increasing number of people became eligible to take the vaccine, the fact that a meaningful proportion were *unwilling* to take it also came to light. As a White person living in a predominantly Black neighborhood, with a roughly equal number of White and Black friends, I came to notice that the reasons my White and Black friends gave for not being willing to a take a covid-19 vaccine tended to be different. Impressionistically, at least, my White friends seemed more concerned with perceived violations of individual rights; that is, they resented the notion that the government was trying tell them what to

do with their bodies (ie. get a covid-19 vaccine). By contrast, my Black friends seemed less concerned with infringements on their individual liberties and more with the extent to which the government — and the vaccines for which the government was advocating — were trustworthy.[3] Very soon, polls began widely reporting that vaccine hesitancy among Black Americans was indeed particularly high, and that the reasons Blacks across the country tended to give for not wanting a vaccine were indeed ones primarily related to distrust (see, for example: Langer Research Associates, 2020; Ndugga et al., 2021). To anyone with even a cursory sense of American history, and especially the history of medical experimentation in the country, this, of course, makes complete sense.

In response, the government and other organizations began devoting significant resources to trying to counter "vaccine hesitancy" in Black communities. One particularly visible strategy has been to use Black celebrities as vaccine spokespeople via television campaigns, with the implicit belief that race-concordant and high-profile, non-government messengers might be more effective than typical government or medical officials in persuading people to get a covid-19 shot. There is, to date, though, very little empirical evidence that celebrities are compelling messengers of medical information and that race-concordance between messengers and audience matters when it comes to encouraging vaccination. Relying on another online survey experiment, the third dissertation paper evaluates whether Black celebrities are indeed effective at persuading Black Americans to take a covid vaccine, both in general and compared to Black non-celebrities. Self-identified Black respondents were randomized to either see a video of television and film star Tyler Perry, or New York City nurse Sandra Lindsay (the first person to get vaccinated against covid-19 in the U.S.) take a covid-19 vaccine and discuss its benefits. A third group saw a control video in which Tyler Perry talks about his new film studio. The study does not find strong support for the effectiveness of

[3]I express gratitude to the long-time Black residents of my Bedford-Stuyvesant block who took me in as a friend when I appeared in their neighborhood seemingly out of nowhere at the beginning of the pandemic, and trusted me with their concerns even though I was by all accounts an outsider.

race-concordant celebrity messaging on intentions to take a covid-19 vaccine or on attitudes about such vaccines. While distrust in government and the medical establishment is indeed relatively high among Blacks compared to Whites and other racial groups, getting around any effect this might have on vaccine hesitancy by relying on celebrity spokespeople is not clearly effective. It may even backfire: having a strong sense of Black identity and belonging to the Black community was associated in the study with reporting a *lower* likelihood of taking a covid-19 vaccine and more negative attitudes toward vaccines. The work represents the first experimental test of which I am aware of the effectiveness of celebrity messaging on vaccination in the United States, and the first such among Black Americans specifically. Of course, ultimately, no work, and certainly no dissertation project, exists without shortcomings, flaws, things left to be desired — ones that will hopefully be taken up by others in the pursuit of shedding light into new corners of the dark room of what we don't know. The next section discusses the limitations of the three papers that constitute the dissertation and offers some reflections on future directions.

Limitations and Future Directions

The usual caveats that apply to survey experiments also of course apply here. One of these is naturally the question of external validity. It arises, in my view, primarily as concerns the first dissertation paper. Evaluating putative immigrants applying for permanent residency/citizenship in the context of an online survey is necessarily an artificial exercise; while people in the real world certainly pass judgement on immigrants on the basis of a multitude of characteristics every day, they are not typically positioned to have to choose which of two they prefer side by side on a predetermined set of factors. The deception in the experiment intends to alleviate some of the artifice of the task — telling respondents immigrants applying for such status also responded to the same survey to which they themselves are responding — but of course a level of artifice remains. We might also ask how visible immigrants' opinions

actually are to Americans and Canadians in the real world. I venture they are both visible and not. That is, I suspect native-born citizens have some assumptions about the opinions of different immigrant groups, or at least 'immigrants' as a whole. That said, I doubt these assumptions are, at the moment, particularly subtle. As immigrant voices become more prominent in public debate, though, it seems reasonable to suspect knowledge will increase. Indeed, it would be interesting for future work to probe what exactly native-born citizens think the political standings and cultural values of different immigrant groups in the country are, and to track this over time to see whether, how, and under what circumstances it evolves. In a more abstract sense, the study aims to evaluate the effect attitudinal difference might have on where Americans and Canadians draw the line with respect to tolerance. It is important for research to continue developing a subtler understanding of how much cultural difference native-born populations are willing to accept and what might be done, in liberal democracies, to push the mark at which people draw the tolerance line forward.

Limitations to the second dissertation paper come more in the form of questions about internal validity. First, while the CES asks an omnibus question about feelings toward "racial minorities," the experiment that forms the crux of the paper asked about a series of specific racial minority groups and uses as an outcome an index of responses to those questions. If the goal is to replicate the 1997 CES wording change, we might ask about the extent to which those two measures map onto each other. The experiment is also conducted using 2015 data, though the results do not meaningfully change when the data are reweighted to better resemble the 1997 data demographically. While the aim of the study is to understand how language compounds wording effects, it only examines this question in the context of English and French in Canada — some languages might have similarly compounding effects, but others might not at all. It is also important to iterate that while the study can make causal inferences about the effects of wording changes among different linguistic groups, it does not randomly assign the language of survey and therefore only speaks to

descriptive inferences about the effects of language. An obvious extension to this work is to also randomly assign language to a sample of bilinguals; what might be lost in terms of external validity due to the fact that bilinguals tend to be relatively unusual members of many populations would be gained in terms of generating an internally valid causal claim about language. Similarly, it would be instructive to replicate the experiment in different countries where different languages are spoken and with different sensitive questions that might be particularly prone to wording effects. On a larger scale, an expansive cross-country research collaborative engaged in a kind of A/B testing of such wording effects could represent a bigger project aimed at designing a broadly consistent baseline for commonly asked survey questions.

The third dissertation paper is a different beast than the first two in the sense that it concerns itself not with the opinions of a majority toward a minority, but the opinions of a minority and the dynamics that underpin them. As discussed in the paper, perhaps the primary limitation of the experiment is that it involves highly bundled treatments and is thus unable to identify exactly which of the many components that constitute the treatments actually undergird effects we observe (or do not observe). As with the other papers, though, the goal of this one was not to identify a silver bullet solution to encouraging people to take a vaccine, but rather to shed light on how experiences to which people may be exposed in the real world, in all their complexity, might affect vaccine orientations. It is one small piece of what is a much larger puzzle being built in particular by researchers in the medical community, as well as social scientists concerned with persuasion across domains. Like all research, but perhaps especially because it is the first experiment of its kind, it requires replication. Ideally, such replication would attempt to unbundle some of the factors inherent to the present treatments, though as I also note in the paper, this comes at the cost of sacrificing some external validity in that one would need to create treatments people would not necessarily be exposed to in their real lives. On the subject of real lives, it is important

to note that people receive many different kinds of messages on a day-to-day basis, and perhaps especially so when it comes to covid-19 vaccines over the last year. The limited effects reported in this study are plausibly in part due to the fact that the video treatments are but one message of the many people to which people are regularly exposed. There is also the question of durability — to the extent any effects of celebrity messaging exist, do they persist beyond the moment the respondent closes the survey window? When people go out into the world and hear others — perhaps even other celebrities — say otherwise? Future work would do well to study how such interventions fare over time. In addition to studying effects beyond the time of survey, it would also be instructive to understand what effects might look like beyond the sample at hand. For instance, what are the effects of a Black celebrity promoting vaccines on Latino or White respondents? What are the effects of Latino or White celebrities? And more generally, how much does race-concordance itself matter? How much does celebrity status itself matter? Might effects vary across health domains? When widespread distrust meets the need to persuade the distrustful to make a medically-sound choice, knowing the answers to these questions can contribute to a more diverse and flexible persuasive toolbox.

To a conception of the scientific process as one of incremental accumulation rather than silver bullet responses proving or discrediting a claim, survey experiments are suited perfectly. Sniderman makes this case eloquently in his 2018 review of advances in survey experimentation. "The research process," he writes, "is a process—a progression of trials. Survey experiments, because of their modesty, lend themselves to this conception of research. Modesty is the operative term: Survey experiments are radically imperfect" (p. 274). It is with this view of science and this appreciation for modesty and imperfection that I approached this dissertation. My hope is that as research marches on, its work is replicated and its limitations are explored, and that, ultimately, its findings can contribute to tracing a clearer pattern upon the deep complexities of public opinion.

Chapter 1: Ideological Conflict and Attitudes Toward Immigrants in the United States and Canada

Abstract: Large-scale immigration to Western countries in recent decades has spurred extensive research examining the factors that make native-born citizens more or less likely to accept newcomers. Building on the literature examining how immigrant characteristics shape natives' opinions about them, this paper evaluates whether immigrants' political attitudes matter to how they are perceived. Specifically, it asks: how would an immigrant agreeing or disagreeing with a native-born citizen about politics affect the citizen's attitude toward that immigrant? Results from survey-based conjoint experiments in the United States and Canada demonstrate that attitudinal conflict between citizens and immigrants predicts negativity toward immigrants to an equal or greater extent than other ascriptive characteristics typically found to generate anti-immigrant sentiment. Moreover, despite conservatives' reputation for intolerance, people across the political spectrum are equally willing to express negative attitudes towards immigrants who disagree with them about politics.

1 Introduction

Large-scale immigration to Western countries in recent decades has spurred an extensive literature examining the factors that make native-born citizens more or less likely to accept newcomers. This research is generally characterized by two main lines of reasoning. The first posits that labor market competition makes natives feel economically threatened and that feelings of economic threat turn into antipathy toward immigrants (e.g., Dancygier and

Donnelly, 2012; Mayda, 2006; Scheve and Slaughter, 2001). The second argues that antipathy toward immigrants emerges not from economic threat, but rather from feelings of symbolic, or cultural, threat (eg., Citrin et al., 1997; Kalkan, Layman and Uslaner, 2009; Sniderman and Hagendoorn, 2007). On balance, the latter view has received more empirical support (see: Hainmueller and Hopkins, 2014*b*). Work in this vein has tended to focus on how immigrants' language abilities, race, religion, and compliance with norms shape natives' opinions about them. This project examines how immigrants' *political attitudes* may matter. Specifically, it asks: how does an immigrant disagreeing with a native-born citizen about politics affect the citizen's attitude toward that immigrant? Answering this question entails, for the first time, integrating the political science literature on public opinion toward immigrants with the social psychology literature on ideological conflict.

In general, people prefer those who agree with them over those who do not. Attitudinal similarity has been shown to supersede considerations of similarity based on race, gender, and social status (Erwin, 1971; Crano, 1997; Fawcett and Markson, 2010). The "ideological conflict hypothesis" argues that similarity-liking applies to politics as well: despite conservatives' reputation for intolerance, people across the political spectrum have been shown to be equally willing to express negative attitudes towards, and even discriminate against, others if they disagree with them (Brandt et al., 2014). I seek to understand how ideological conflict and ideological similarity between natives and immigrants affects natives' attitudes toward immigrants in the two countries that receive among the largest number of immigrants each year: the United States and Canada. Relying on a conjoint design, an experiment embedded in a survey asks Americans and Canadians to choose between two putative immigrants whose profiles are composed of randomly assigned attributes. While six of these are ascriptive, tapping economic and symbolic threat hypotheses, five are attitudinal. The experiment randomly assigns the immigrants' putative party identification as well as their attitudes on four salient political issues: i) abortion rights, ii) gay marriage, iii) welfare spending, and

iv) environmental regulation. Respondents' own positions on the same issues, as well as their partisanship, are measured earlier in the survey in order to allow us to estimate how *distance* between native-immigrant political orientations affects natives' attitudes toward immigrants. The magnitude of these effects can then be compared to those of the six ascriptive attributes used to capture economic and symbolic threat. The results demonstrate that attitudinal conflict between citizens and immigrants predicts negativity toward immigrants to a greater extent than other ascriptive characteristics typically found to generate anti-immigrant sentiment. Moreover, people across the political spectrum are more or less equally willing to express negative attitudes towards immigrants who disagree with them about politics. For example, those who support a woman's rights to legal abortion tend to dislike immigrants who do not support that right as much as anti-abortion citizens dislike immigrants who are pro-choice. Interestingly, in the Canadian case, people on the political left are more punitive toward disagreeing immigrants than are people on the right. The paper begins with a review of the political science literature on immigration attitudes and the social psychology literature on ideological conflict. The study design is then described and expectations about the impact of political disagreement on attitudes toward immigrants are tested. Average Marginal Component Effects (AMCEs) show that political disagreement between citizens and immigrants is an important factor shaping attitudes toward immigrants across the political spectrum, though marginal means reveal that, in Canada, citizens on the left may be more more punitive than those on the right toward those who disagree with them. The paper concludes with a discussion of the implications of these results for understanding contemporary immigration dynamics in western countries.

2 Economic and Cultural Threat Hypotheses

First, a birds-eye view of immigration attitudes in the Unites States and Canada. In both countries, attitudes toward immigrants and immigration have generally become more positive

over time, though Canadians have tended to be more accepting of open immigration policies and immigrants than Americans and Americans paid scant attention to immigration politics until roughly the 1990s. Yet within both contexts tensions have abounded: in Canada, differences in attitudes have varied significantly across provinces due to unique local conditions (Bilodeau, 2010) and in the U.S., significant attitudinal inconsistencies exist across different aspects of immigration and immigrants (Levy and Wright, 2020). In terms of the sources of anti-immigrant attitudes, the literature long held that immigrants are disliked by native-born citizens because they constitute competitors for scarce economic resources (e.g., jobs, wages) (Bilodeau, Turgeon and Karakoc, 2012; Bobo and Hutchings, 1996; Semyonov, Raijman and Gorodzeisky, 2006; Mayda, 2006; Scheve and Slaughter, 2001; Sides and Citrin, 2007). However, as Hainmueller and Hopkins (2014*b*) point out, the evidence that attitudes toward immigrants primarily emerge from economic concerns is relatively weak (though see, for example: Iyengar et al., 2013; Schmuck and Matthes, 2017). Rather, attitudes toward immigrants seem to derive mostly from cultural factors. To the extent that native-born citizens dislike immigrants, such antipathy is thought to largely emerge from the perception that immigrants threaten national identity, culture, and norms (Citrin, 1990; Hainmueller and Hiscox, 2010; Sniderman and Hagendoorn, 2007).

Several recent studies have zeroed in on the characteristics of immigrants that may generate feelings of cultural threat, including not speaking the national language (Schildkraut, 2011), not respecting laws (Hainmueller and Hopkins, 2014*b*), and violating norms of propriety and culture (Figueroa, 2018) (for other examples, see: Clayton, Ferwerda and Horiuchi, 2019; Ford, 2011; Hainmueller and Hangartner, 2013; Hainmueller and Hopkins, 2014*a*; Harrell et al., 2012; Hellwig and Sinno, 2017; Wright, Levy and Citrin, 2016). In a review article, Hainmueller and Hopkins (2014*b*) conclude that "too frequently, culture operates as a residual category, describing any non-economic immigrant attribute. On its own, the claim that culture matters thus has less content than meets the eye" (p.242). Newman (2013) also note

that "despite its status as a prepotent source of support for anti-immigrant policies, cultural threat is presently an under-theorized and underanalyzed concept in the literature." (p.375). This paper aims to further clarify which cultural factors, specifically, affect attitudes toward immigrants. While culture is certainly about immigrant behavior, as other studies have shown—speaking national languages, following laws, etc.—this paper asks whether it may also be about immigrant beliefs. Perhaps the closest work to mine here is Levy and Wright's (2020) book that shows that values of civic fairness are the most important predictors of attitudes toward immigrants. Immigrants whose behaviors and outlooks comport with these values are viewed more favorably than those who shirk them, regardless of economics, race, and the other usual suspects. In a similar vein, I attempt to evaluate to what extent immigrants are punished for differing from citizens when it comes to their political beliefs. The work diverges from extant literature in that one's political beliefs are of course not a set of norms or values; however, they are likely to be strongly influenced by both culture, which also influences norms, and values.

3 Political Ideology and Attitudes Toward Outgroups

Research on the relationship between political beliefs and attitudes toward minority groups has focused less on the beliefs of minorities and more on the ideological positionings of majorities. This literature has tended to find that political conservatism among majority group members predicts more negative attitudes toward minorities than does liberalism (Adorno et al., 1950; Altemeyer, 1998; Esses, Haddock and Zanna, 1993; Feather, 1985; Gaertner and Dovidio, 1986; Katz and Hass, 1988; Kinder and Sears, 1981; Sniderman and Tetlock, 1986). Perhaps most notable for the purpose of this paper, though, is Rokeach's (1960) demonstration that it is not majority ideology per se that matters, but rather the *distance* between majority and minority ideology. A number of recent studies finds both liberals and conservatives tend to attribute negative characteristics to those whose values

are dissimilar from their own (Morgan, Mullen and Skitka, 2010; Skitka et al., 2002), distance themselves from those with different moral orientations (Skitka, Bauman and Sargis, 2005), and express intolerance toward those with whom they disagree (McClosky and Chong, 1985; Lambert and Chasteen, 1997; Yancey, 2010). As Sniderman and Hagendoorn (2007) put it:

> No one supposes there is no prejudice on the left. But the conventional wisdom that the danger lies to the political right tacitly assumes that being on the left to some degree immunizes one against the effects of intolerance. If one stops and gives some thought to the matter, however, the assumption is not obviously plausible. What reason is there to suppose that the psychology of intolerance changes depending on the political perspective from which a person views the world?

The idea that majority dislike of minorities is conditional on the distance between majority-minority beliefs has been subsumed under the umbrella of the "ideological conflict hypothesis." We know people often conflate racial and religious characteristics with target groups' or individuals' political orientations. Chambers, Schlenker and Collisson (2013) and Reyna et al. (2006) both show that when white conservatives are told a putative Black person shares their views, antipathy toward the person decreases, suggesting conservative prejudice toward Blacks may at least in part be grounded in the assumption that Blacks skew politically liberal. In a similar vein, British conservatives express more positive attitudes toward Muslims when informed that putative Muslims share some of their values (Helbling and Traunmüller, 2018). Building on these studies, this paper represents a first effort to measure the implications of majority-minority ideological conflict using concrete and multiple indicators of ideological orientations by testing how immigrants' opinions about abortion, gay marriage, welfare spending, and environmental regulation, as well as partisan orientation, affect how natives feel about them. The abortion and gay marriage issues were chosen to

capture a social ideological dimension, while the welfare spending and environmental regulation issues were chosen to capture an economic and government intervention dimension. All the issues are politically salient in both countries. Partisanship is manipulated to capture the partisan dimension of ideological divides in the two countries. By measuring the effects of ideological indicators alongside those associated with cultural and economic threat hypotheses (captured by ascriptive characteristics in the experiment), the study likewise aims to be the first to integrate the social psychology perspective on ideological conflict with the political science literature on attitudes toward immigrants.

Understanding how the characteristics of immigrants shape attitudes toward them can help us understand the scope of tolerance in a democratic society. We know that the cultural characteristics of immigrants matter, but beliefs are a part of culture and yet it is unclear what role they play. Indeed, it can be argued that value conflicts between immigrants and natives may pose bigger threats to toleration than even xenophobia or prejudice. The latter two are grounded in groups harboring misperceptions about one another. But when it comes to value conflicts, "each [group] believes that what the other believes is right, is wrong: and each is right about what the other believes is right." (Sniderman and Hagendoorn, 2007, p.29). While misperceptions can perhaps be corrected, the literature on ideological conflict suggests that "it is unlikely that group members who hold a strong conviction, be it cultural, religious, or political, will come to like and approve of beliefs and practices of out-group members who strongly subscribe to an alternative worldview (Verkuyten, Yogeeswaran and Adelman, 2019, p.6). When it comes to political elites, it is not clear whether the right-wing parties in both Canada and the United States take relatively restrictive immigration positions because they do not like immigrants or because they believe immigrants tend to politically sympathize with their opponents (e.g., Bilodeau and Kanji, 2010; Judis and Teixeira, 2002). Conversely, it is also not clear whether support for more lax immigration policies on the left would be lower if incoming immigrants were found to lean to the right. The next section

describes the procedure used to shed light on the relationship between ideological conflict and attitudes toward immigrants.

Both the United States and Canada receive among the highest number of immigrants per capita each year of any country in the world. They represent ideal contexts in which to study how immigrants' beliefs shape natives' attitudes toward them in part because by now both countries have politically consequential numbers of foreign-born residents. Moreover, in addition to being democratic societies with a commitment to toleration, they bear many structural and institutional similarities, including Anglo Saxon roots, first-past-the-post electoral systems, capitalist economies, and federal political structures. Yet they have taken different approaches to immigrant integration in ways that may contour how ideological conflict plays into immigrant-native relations. Whereas Canada has long taken a multicultural approach, the United States has billed itself as a "melting pot," with relatively more stringent expectations regarding immigrant conformity to norms. That said, there have been more high-profile debates pitting immigrant groups against groups of native-born citizens in Canada than there have been in the United States (see: Dufresne et al., 2019; Turgeon and Bilodeau, 2014; Harrell et al., 2012). We can think, for example, about the extremely contentious debate over religious accommodations in Quebec, spanning the Bouchard-Taylor Commission on the issue in 2008 to the 2019 law banning public sector employees from wearing religious symbols. Political conflict with immigrants over issues may thus be more salient in the minds of Canadians than Americans. Conversely, political polarization in the United States is generally thought to be greater in the United States than in Canada (see, for example: Kevins and Soroka, 2018); consequently, we should expect disagreement over political parties to have a greater impact on attitudes toward immigrants in the United States than in Canada.

4 Study Design

4.1 Data

In the United States, the study is conducted through Lucid's Fulcrum Exchange (n = 1,000) and in Canada through Vox Pop Labs' Internet panel (n = 2,000), both of which yield representative samples of native-born populations in the two countries.[1] After answering some demographic questions about themselves, respondents indicated the extent to which they agree or disagree with four different issue statements: marriage should only be between a man and a woman; women should have a legal right to access abortions regardless of the reason; environmental regulations should be stricter, even if it means consumer paying a higher cost; and government should increase spending on welfare. Although using measures of specific attitudes as indicators of values has a long pedigree in work on ideology and intergroup attitudes (Katz and Hass, 1988; Parsons, 1960; Rokeach, 1973), work connecting ideology to prejudice has tended to rely on single-item self-placement (for example: Chambers, Schlenker and Collisson, 2013; Crawford and Pilanski, 2014; Wetherell, Brandt and Reyna, 2013). To the extent that there is political conflict or comradery among immigrants and native-born citizens, it is ultimately over specific public policies. Moreover, attitudes about gay marriage and abortion are expected to be held with greater intensity than attitudes about environmental regulation and welfare spending because they tap moral mandates; while liberals and conservatives differ in their positions on moral issues (Graham, Haidt and Nosek, 2009), they tend to endorse their own positions with high and similar intensity (Skitka and Bauman, 2008). Thus, attitudinal difference between natives and putative immigrants on abortion and gay marriage are expected to more strongly predict negativity toward immigrants than are attitudes about the environment and welfare spending.

[1]The study received approval from the Institutional Review Board of Columbia University in the City of New York (#AAAS1209)

4.2 Experimental Design

After indicating their preferences on these four political issues, as well as their partisan affiliation using a standard party identification question, respondents are presented with the conjoint task. The prompt reads:

> This study considers immigration and who is permitted to live in the United States (Canada). There are different opinions about what sorts of people from other countries should be allowed to permanently live here. We'd like to know your opinion. Some immigrants applying for (permanent residency/citizenship) in the United States (Canada) answered the same questions you did about abortion, gay marriage, the environment, welfare, and party identification. You will see their responses and other basic information about them below. For each pair of people you see, please indicate which of the two you would prefer be given (permission to permanently live/citizenship) in the United States (Canada) or whether you would prefer that neither be given (permission/citizenship).[2]

Whether respondents were asked about permanent residency or citizenship was randomized in order to evaluate whether the kind of legal status a respondent was asked about would affect his or her response. The experiment asks about permanent residency and citizenship both because the putative immigrants have to have been in the country long enough to realistically have positions on issues and party identification and because receiving either permanent residency or citizenship are pivotal steps in being allowed to remain in the country and hold lasting political sway. That said, there is no statistically significant difference between the attribute effects of interest under the two treatment conditions and remaining analyses thus pool respondents across permanent residency/citizenship framings.[3] They are

[2]The French translation of the conjoint task available to Canadian respondents can be found in the Appendix.

[3]The results from these tests are presented in Tables 2 and 3 in Appendix.

then asked whether they would prefer Immigrant 1, Immigrant 2, or neither be given permanent residency/citizenship; this question forms the primary dependent variable of the study. Such a "forced choice" design, excluding a "both" option, allows us to observe the kinds of trade-offs people make when evaluating immigrants who vary along multiple dimensions. An example of the conjoint task from the U.S. study is presented in Figure 1. Respondents completed the task four times, evaluating eight profiles in total. While the order of profile attributes was randomized across respondents to preclude ordering effects, it was constant within respondents across profiles and tasks to reduce cognitive load. The order randomization was restricted such that issue attributes and ascriptive attributes always appeared one after the other, though the order of issue and ascriptive attributes within the blocks was randomized, and whether a respondent saw ascriptive or issue attributes at the top or bottom of the profiles was also randomized.[4] One limitation of this study design is that it may be prone to social desirability bias, given the politically charged nature of the question. However, extant work relying on conjoint experiments to assess the bases of attitudes toward immigrants has not found strong evidence of social desirability bias (see, for example: Hainmueller and Hopkins, 2014*a*), and respondents completed the survey entirely online. Another potential limitation is that it might not have meaningful external validity. However, this is a limitation that afflicts essentially all survey experimental work where people respond to hypothetical scenarios, and evidence shows that both responses to conjoint and vignette experiments tend to comport well with real-world behavior (Hainmueller, Hangartner and Yamamoto, 2015).

[4]The only cross-attribute constraint was that immigrants from Mexico could not also be Muslim, given the Muslim population of Mexico is less than 1% and therefore leads to a counterfactual that cannot be meaningfully evaluated.

Table 1: Immigrant Attributes with Levels to be Randomized (USA)

Attributes	Levels
Gender	Woman
	Man
Country of Origin	Syria
	Mexico
	Poland
	India
	China
	Somalia
	Iraq
Language	Speaks fluent English
	Speaks broken English
	Requires a translator in English
Profession	Nurse
	Fruit picker
	Child care provider
	Doctor
	Computer programmer
Job Experience	Mostly unemployed
	Mostly employed
	Continuously employed
Religion	Christian
	Muslim
	None
Abortion opinion	Agrees with legal abortion
	Disagrees with legal abortion
Environment opinion	Wants stricter environmental regulation
	Doesn't want stricter environmental regulation
Welfare opinion	Wants more welfare spending
	Doesn't want more welfare spending
Gay marriage opinion	In favor of gay marriage
	Against gay marriage
Party identification	Democratic Party
	Republican Party
	Independent

Note: In the Canadian sample, the political party attributes were: Liberal Party, Conservative Party, New Democratic Party, and Bloc Québécois. French respondents in Canada were presented with French language ability.

Figure 1: Conjoint Task Example (USA)

	Immigrant 1	Immigrant 2
Language	Requires a translator in English	Speaks broken English
Gender	Woman	Woman
Profession	Doctor	Nurse
Job experience	Mostly unemployed	Mostly unemployed
Religion	Christian	None
Country of Origin	Mexico	Somalia
Party affiliation	Independent	Democratic Party
Abortion opinion	Disagrees with legal abortion	Agrees with legal abortion
Environment opinion	Wants stricter environmental regulation	Doesn't want stricter environmental regulation
Welfare opinion	Wants more welfare spending	Wants more welfare spending
Gay marriage opinion	In favour of gay marriage	Against gay marriage

If you had to choose between them, which of these two immigrants would you prefer to live permanently in the United States?

- Immigrant 1
- Immigrant 2
- Neither

4.3 Estimation and Expectations

The study aims to capture the heterogeneous treatment effect of the immigrant issue and partisanship attributes conditional on respondents' own answers to the issue and party ID questions. An example would be the effect of an immigrant being pro-abortion on the preferences of respondents who are themselves anti-abortion. Average Marginal Component Effects (AMCEs) for each attribute level are estimated using a logistic regression where the dependent variable is whether an immigrant profile was chosen or not (0, 1) and the independent variables are immigrant attributes interacted with a variable capturing respondents' positions on the issues and party identification (for detail about AMCE definition and estimation, see: Hainmueller, Hopkins and Yamamoto, 2013). Because AMCEs are necessarily defined relative to a baseline attribute level (like categorical variables in standard regression analysis), it is also useful to evaluate subgroup marginal means in order to also get a sense of

the underlying *preferences* over attributes within subgroups (see: Leeper, Hobolt and Tilley, 2019). Although respondents were asked to indicate their positions on abortion, gay marriage, welfare, and the environment on 4-point Likert scales in the surveys, responses are binarized (pro- and anti- each issue position) to facilitate analysis.[5] The regression models take the form:

$$y = \beta_0 + \beta_1 x + \epsilon, \quad \forall z = 0$$
$$y = \beta_2 + \beta_3 x + \epsilon, \quad \forall z = 1$$

The AMCE of an attribute x for respondents scoring 0 on the pre-treatment covariate z is β_1, while the AMCE of the attribute for respondents scoring 1 is β_3 meaning the difference in AMCEs is given by $\beta_3 - \beta_1$. Though this quantity tells us about differences in the effect sizes across respondents who take different issue positions and have different party identifications, it does not tell us about the structure of preferences across the subgroups. AMCEs are necessarily sensitive to the baselines that are chosen, so estimating marginal means in addition to AMCEs is useful to illustrating the preferences over all attribute levels for subgroups. Marginal means using the equations above are defined by $\bar{y}_{z=1|x=1} - \bar{y}_{z=0|x=1}$ and estimated by $(\beta_2 + \beta_3) - (\beta_0 + \beta_1)$.

Following the ideological conflict hypothesis, I expect that disagreement between native born respondents and putative immigrants on the four political issues will lead to decreased probabilities of accepting immigrants. Similarly, I expect differences in party identification to lead to decreased probabilities of accepting immigrants. I also expect the effects of disagreement to be non-zero for respondents taking both the right- and left-wing positions on

[5]Presenting four different opinions as levels to each issue attribute would have been burdensome in terms of profile presentation; it is cognitively simpler for respondents to see that an immigrant "agrees with legal abortion" or "disagrees with legal abortion" than to see "strongly agrees," "somewhat agrees," etc. Respondents' answers to the four issue questions are thus binarized to measure agree/disagree and these binary variables are interacted with attributes in the AMCE models.

the issues and across partisans. Given that high-salience debates over issues dividing largely native-born and immigrant communities have been more prominent in Canada, I expect the magnitude of the effects of issue disagreement to be greater in Canada (see: Bilodeau, Turgeon and Karakoc, 2012; Dufresne et al., 2019; Turgeon and Bilodeau, 2014). Because, these debates have generally turned around issues where the Canadians involved stood on the political left and immigrant groups stood on the political right, I expect Canadians taking the leftist position on the issues to be more punitive toward those who disagree with them than Canadians taking the rightist position. Finally, due to greater partisan polarization in the United States than in Canada, I expect the magnitude of the effects of partisan disagreement to be greater in the United States. The next section provides tests of these hypotheses, first by estimating issue and partisan disagreement AMCEs and then by examining the structure of preferences over political ideology by marginal means [6].

5 Results

5.1 Ideological Conflict in the United States

Beginning with the United States, Figure 2 presents the probabilities that a given immigrant will be chosen if he or she *disagrees* with a native-born citizen on the four political issues under study, thereby testing the effects of political conflict.[7] Across all four issues, conflict between a putative immigrants' position on an issue and a native-born citizens' position tends to be associated with a decrease in the probability (ie. the AMCE) that the immigrant will be chosen to remain in the country, on the order of between roughly three and eight percentage points.[8] Respondents who are "supporters" on any given issue (orange coeffi-

[6]For a thorough discussion of the utility of estimating marginal means when examining heterogeneous treatment effects in conjoint experiments, see: Leeper, Hobolt and Tilley (2019).

[7]Analyses follow the pre-analysis plan registered with Evidence in Governance and Politics (EGAP) under ID #: 20190920AB.

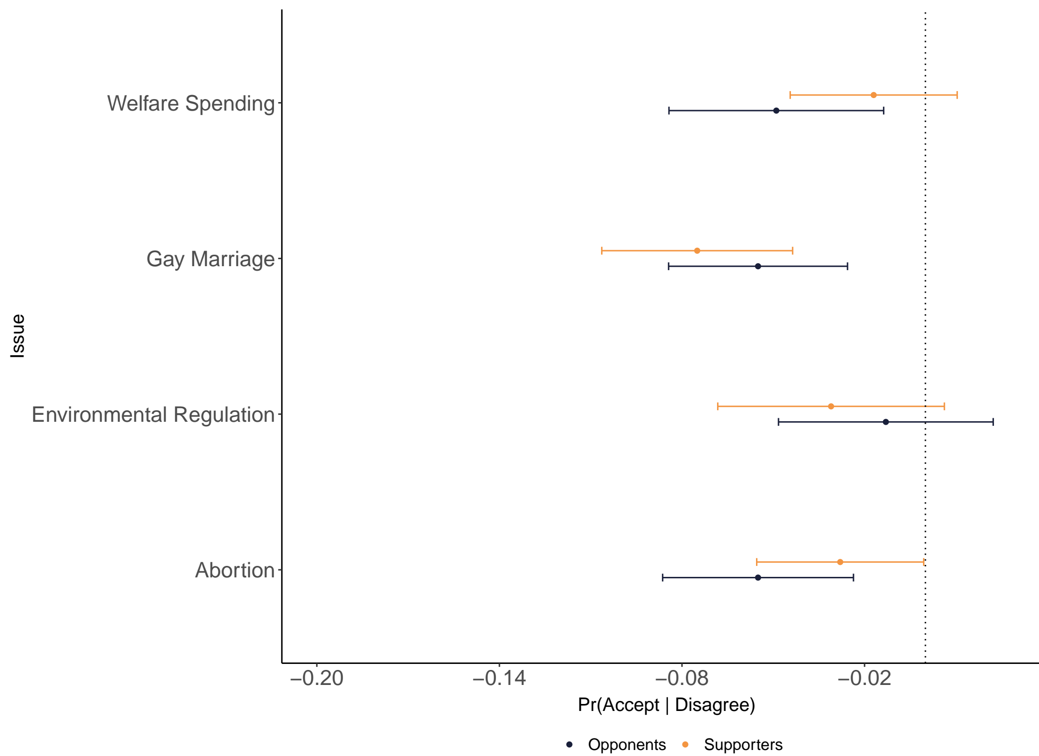
[8]Because the study focuses on ideological conflict, AMCEs are presented in terms of the probability of an immigrant being accepted when there is *disagreement* between respondents and immigrants, but AMCEs

cients in Figure 2) stand on the left-wing side of that issue (ie., are in favor of legal abortion, gay marriage, increased welfare spending, and greater environmental regulation) while "opponents" (navy coefficients) stand on the right-wing side of the issue (ie., are against legal abortion, gay marriage, increased welfare spending, and greater environmental regulation). We observe the largest negative effect of immigrant-native disagreement on the probability that an immigrant will be chosen to stay in the country when it comes to gay marriage: both supporters and opponents of gay marriage are less likely to choose immigrants who disagree with them on this question by about 7.5 and 5.5 percentage points, respectively. Similarly, supporters and opponents of legal abortion both penalize immigrants who disagree with them on abortion, though in this case, opponents are more punitive than supporters and the negative effects are slightly smaller in magnitude than when it comes to gay marriage. Results are more mixed when it comes to welfare spending and environmental regulation. Opponents of increased welfare spending are about 5 percentage points less likely to accept immigrants who disagree with them on the issue than immigrants who agree with them, but we observe no statistically significant effect of disagreement on supporters. The opposite is true when it comes to increased environmental regulation; respondents in favor of increased environmental regulation are about 3 percentage points less likely to accept immigrants who disagree with them than those who do not, but there is no statistically significant effect of disagreement among opponents.

Turning to partisanship, Table 2 presents AMCEs for party identification attributes across respondent partisanship. As expected, the effects of partisanship tend to be greater in magnitude than those of issue agreement or disagreement in the U.S. case, with partisan disagreement being associated with decreases in the probabilities of immigrants being chosen to stay of between roughly 5 and 10 percentage points. The effects are asymmetric among partisans, with Democrats being nearly 10 percentage points less likely to choose an immi-

for agreement are of course mathematically equivalent, simply flipped in terms of their signs.

Figure 2: Political Issue AMCEs Under Disagreement (USA)



Note: AMCEs of randomly assigned immigrant attributes by respondent subgroups from a logistic regression, where the dependent variable is whether a given immigrant profile is chosen (1) or not (0). AMCEs are estimated using respondent-clustered standard errors.

grant who is a Republican than a Democrat, while Republicans are only about 5 percentage points less likely to choose a Democrat than a Republican.

How do the magnitudes of these effects compare to those of the ascriptive attributes used to test economic and cultural threat hypotheses? The AMCEs for issue and partisan disagreement presented so far are all within respondent subgroups, meaning that in order to generate a full comparative picture of the effects of the remaining attributes, we would need to look at them by all issue- and party ID-based subgroups, ultimately generating an unwieldly $5 \times 2 = 10$ plots. Thus, for the purposes of efficient illustration, let us estimate the ascriptive AMCEs using subgroups defined by the gay marriage issue as an example. AMCE plots for all 11 attributes across all (2×7) respondent issue/partisanship subgroups are presented in

Table 2: Effect of Out-Partisanship on Attitudes Toward Immigrants (USA)

	Pr(Accept Immigrant | Party Disagreement)		
		Immigrant Party ID	
	Democratic	Republican	Independent
Democratic		−0.096*** (0.022)	−0.057* (0.023)
Republican	−0.053* (0.024)		−0.060 (0.024)
Independent	0.007 (0.024)	0.061* (0.024)	

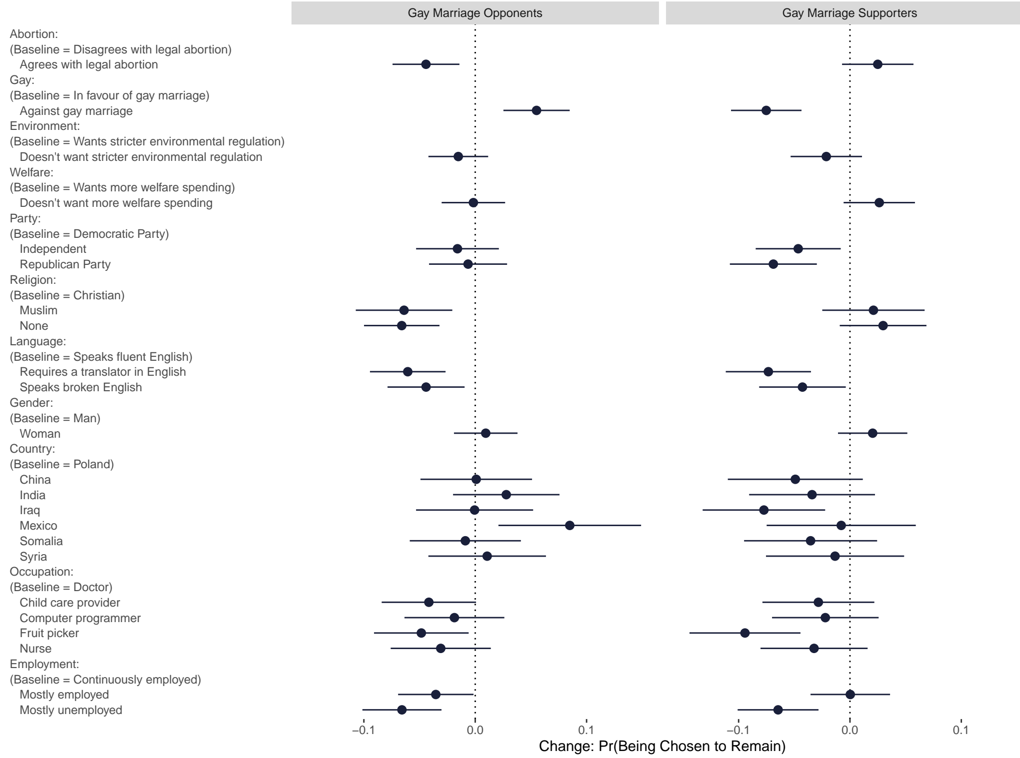
Note: Respondent-clustered std. errors in parentheses — *p<0.05; **p<0.01; ***p<0.001

the Appendix, but patterns are highly similar across subgroups. As Figure 3 shows, findings with respect to the ascriptive AMCEs cohere with extant work testing cultural and economic threat hypotheses: immigrants are penalized more for violating cultural norms, such as being unemployed, not speaking English fluently, or being non-Christian, than they are for their occupational choices. The effects of disagreeing over political issues tend to be at least as great, if not greater, in magnitude than these classical cultural factors. By way of example, among gay marriage supporters, immigrants are only 4.3 percentage points less likely to be chosen to remain if they speak broken English (compared to fluent English), but fully 5.5 percentage points less likely to be chosen if they do not support gay marriage.

5.2 Ideological Conflict in Canada

Similar patterns are observed in the Canadian case, though the AMCEs for both issue and partisan disagreement tend to be much larger in magnitude than in the U.S. case. What is perhaps most striking in Canada, as Figure 4 shows, is that *supporters* of the various issue positions—that is, people who hold the left-wing position on the issue—consistently tend to disfavor immigrants who disagree with them (compared to those who agree with them) to a much greater extent than do opponents (ie. those holding the right-wing position). Among

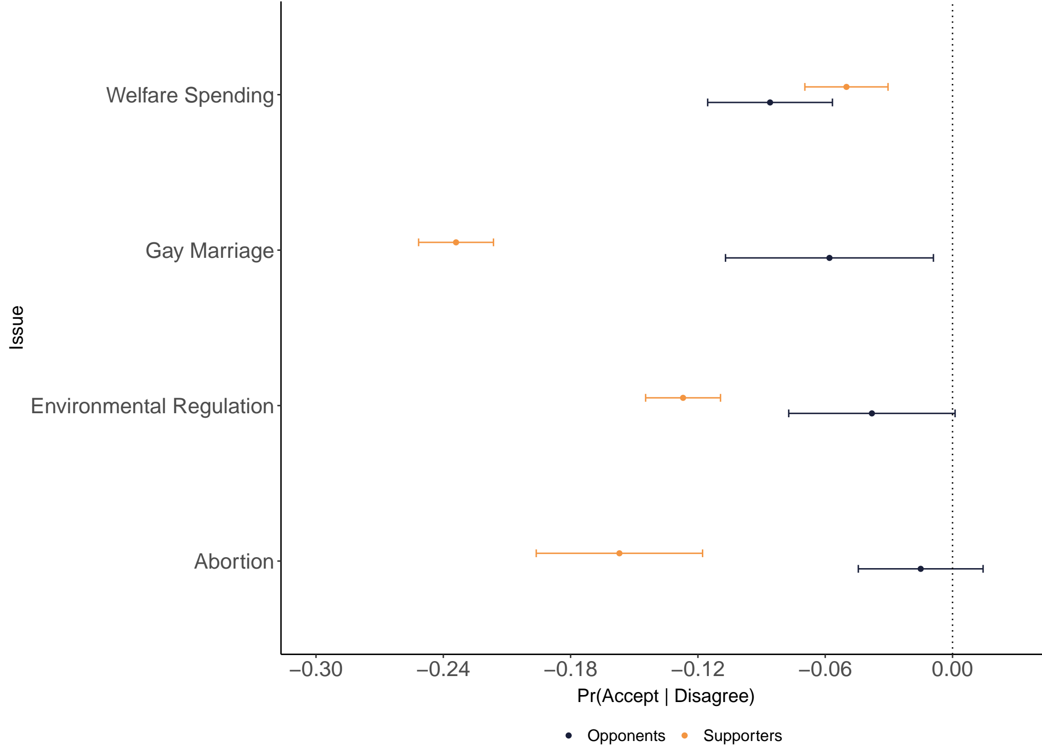
Figure 3: Gay Marriage AMCEs by Subgroup (USA)



respondents who support legal abortion, opposing legal abortion is associated with a nearly 16 percentage point decrease in the probability that the immigrant will be chosen to remain in the country (compared to an immigrant who supports legal abortion). And among gay marriage supporters, immigrants who are against gay marriage are fully 25 percentage points less likely to be chosen than immigrants who are pro-gay marriage. As in the American case, respondents in the Canadian case tend to be more reactive to disagreement on the two moral issues, abortion and gay marriage, than on the economic/government intervention issues, welfare spending and environmental regulation. But the magnitudes of the negative effects of issue disagreement are overall quite striking and, as expected, larger than in the U.S. case.

Estimates for partisanship AMCEs are presented in Table 3. In general, an immigrant identifying with an opposing party is associated with between a 6 and 16 percentage point

Figure 4: Political Issue AMCEs Under Disagreement (Canada)



Note: AMCEs of randomly assigned immigrant attributes by respondent subgroups from a logistic regression, where the dependent variable is whether a given immigrant profile is chosen (1) or not (0). AMCEs are estimated using respondent-clustered standard errors.

decrease in the probability of the immigrant being chosen to stay. The AMCEs of out-partisanship are generally greatest among respondents who identify with the Bloc Québécois, arguably reflecting tensions in the country between Quebec and the rest of Canada. In line with expectations, the effects of issue disagreement are greater than partisan disagreement in Canada, which is the opposite of what we observe in the American case. This may reflect the fact that partisan polarization tends to be lower in Canada (Kevins and Soroka, 2018), on the one hand, and on the other, the fact that the country has seen numerous high-profile debates over issues, including abortion and gay marriage, where certain immigrant groups vocally took stands against the status quo.

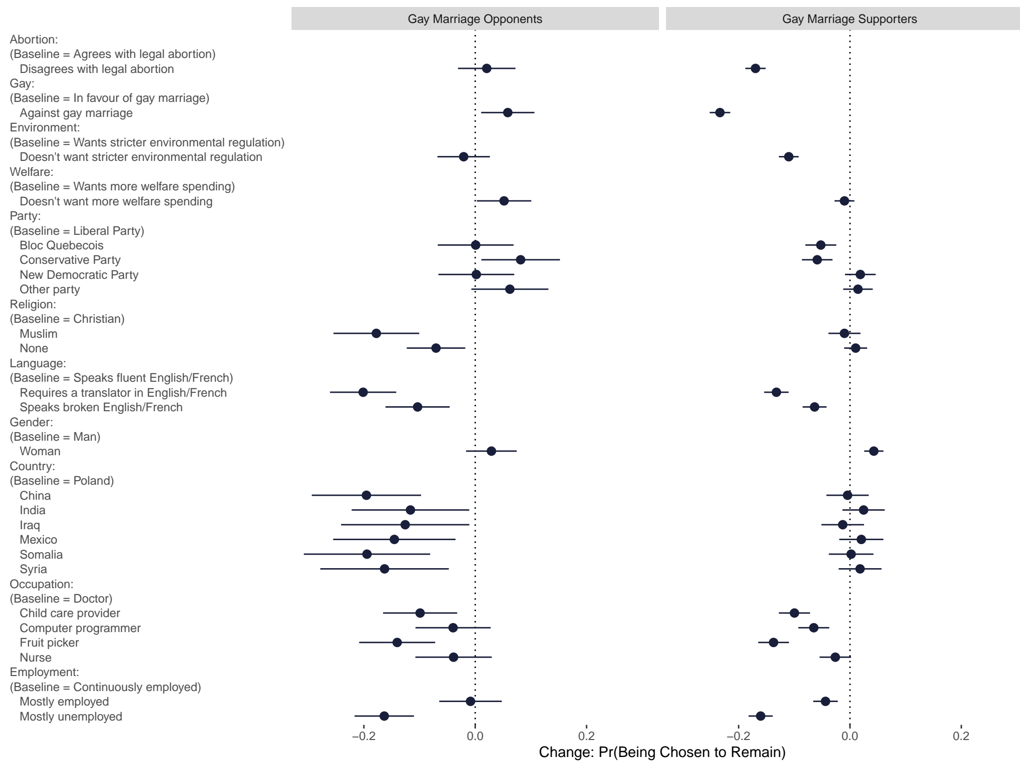
To compare the AMCEs on political issues to those of the other ascriptive characteristics

Table 3: Effect of Out-Partisanship on Attitudes Toward Immigrants (Canada)

	Pr(Accept Immigrant | Party Disagreement)			
	Immigrant Party ID			
	NDP	Liberal	Conservative	Bloc Quebecois
Liberal	−0.000		−0.052	−0.105***
	(0.023)		(0.023)	(0.022)
Conservative	−0.030	−0.036		−0.090**
	(0.030)	(0.027)		(0.031)
NDP		−0.102***	−0.058*	−0.132***
		(0.025)	(0.025)	(0.025)
Bloc Quebecois	−0.11*	−0.12*	−0.157**	
	(0.053)	(0.052)	(0.059)	

Note: Std. errors in parentheses *p<0.05; **p<0.01; ***p<0.001

Figure 5: Gay Marriage AMCEs by Subgroup (Canada)



included in immigrant profiles, we can once again focus on subgroups formed around the gay marriage issue. As in the U.S., the AMCEs for the ascriptive attributes in the Canadian case, illustrated in Figure 5, in general support the findings of extant work on cultural and economic threat: Canadians tend to prefer immigrants who are employed to those who are unemployed, immigrants who speak fluent English (or French) to those who speak broken English (French), etc. Violations of cultural norms also matter more to Canadians than do economic factors. But the effect of the gay marriage issue among gay marriage opponents is as large as these ascriptive characteristics, and among gay marriage supporters, is larger than all other characteristics. Put differently, Canadians who are pro-gay marriage care more that a putative immigrant disagrees with them on this issue than how well the immigrant speaks English (or French), what country they came from, whether they are employed or not, etc. Full AMCE figures for subgroups formed by the remaining issues can be found in the Appendix, but in general tell the same story: the effects of immigrant-native issue and partisanship disagreement tend to be as large, if not larger in magnitude, than the effects of characteristics associated with economic and cultural threat hypotheses. At least when revealed, immigrants' political orientations matter significantly to natives' attitudes toward them.

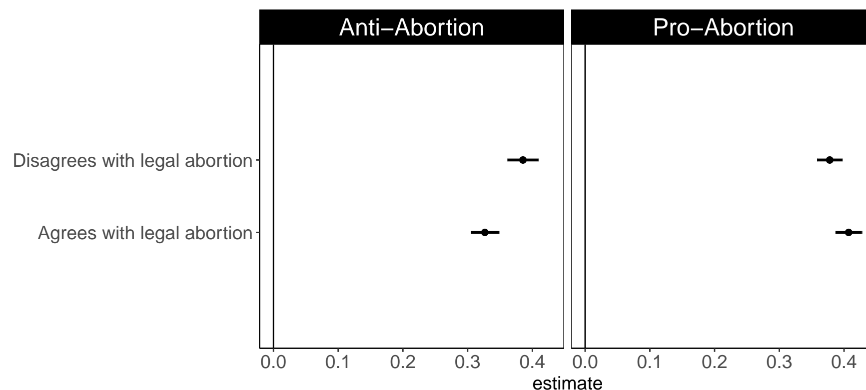
5.3 The Structure of Preferences Over Ideology

The United States

While understanding the causal effects of attributes on the probability that different subgroups will choose a given immigrant, it is also useful to assess the *structure* of respondent preferences over immigrants with different attributes. Greater causal effects of a given attribute in one subgroup as opposed to the other does not necessarily imply that the subgroup is on average more supportive of immigrants across different levels of the attribute. To address the question of preferences, Figures 6 to 10 present the conditional marginal means of

each of the issue and partisanship attribute levels in each of the attitudinal subgroups, as well as differences in marginal means between the subgroups in the United States.[9] Generally speaking, these mean evaluations further support the expectation that native-born immigrants prefer immigrants that are ideologically similar to themselves on salient issues. In Figure 6, differences in attitudes toward immigrants who are pro- and anti-abortion are evident only for anti-abortion Americans (left pane), who significantly prefer immigrants who agree with them over those that do not. The story is different when it comes to gay marriage, however (Figure 7). Here we see perhaps the greatest evidence for ideological conflict between natives and immigrants: both anti- and pro-legal abortion respondents significantly privilege immigrants in their own camp and punish those in the opposing one. Given both abortion and gay marriage are salient moral issues, it is not immediately clear why anti-abortion Americans do not seem to particularly mind immigrants who disagree with them on this issue, but anti-gay marriage Americans do mind when immigrants disagree with them on that issue; future work might further explore this pattern.

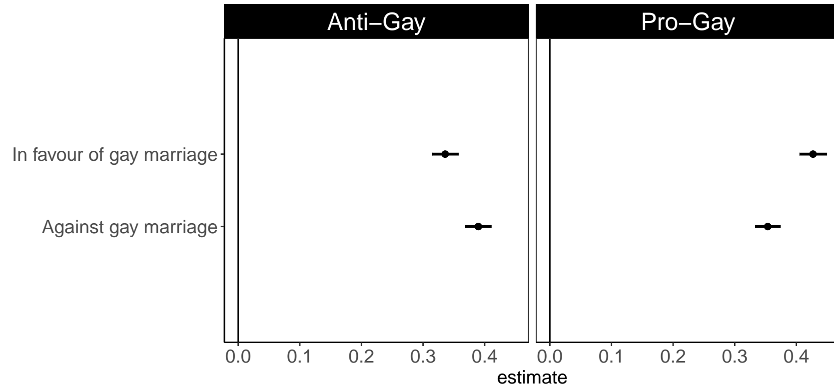
Figure 6: Abortion Marginal Means by Subgroup (USA)



When it comes to welfare spending (Figure 8), significant differences in preferences over pro- and anti-welfare spending immigrants only emerge for anti-welfare spending Americans (left pane). On the question of tightening environmental regulation in Figure 9, anti-

[9]It is important to note evaluating marginal means by respondent subgroups was not in the registered pre-analysis plan.

Figure 7: Gay Marriage Marginal Means by Subgroup (USA)



regulation respondents have similar preferences over immigrants with different positions on the issue (left pane), but pro-regulation respondents reward pro-regulation immigrants and punish anti-regulation ones (right pane). In summary, then, ideological conflict is most clearly manifest when it comes to gay marriage and, secondarily, abortion, which coheres with what we see in AMCEs. The punishing of immigrants whose opinions are different from one's own is less consistent when it comes to welfare spending and environmental regulation, the two economic and government intervention questions.

Finally, Figure 10 illustrates marginal means for the party ID attribute across respondents who identify as Republicans or Democrats. Partisans of each party clearly prefer immigrants who identify with their own party to immigrants identifying with the opposing party, and differences in the marginal means are significant across the board, in the expected directions. To conclude, then, while the pattern concerning the degree to which Americans on the political left and right disfavor immigrants who do not share their views is somewhat inconsistent, the overall picture illustrates that citizens do tend prefer those on their side. How do divides over these issues and partisanship play out in the Canadian context?

Figure 8: Welfare Marginal Means by Subgroup (USA)

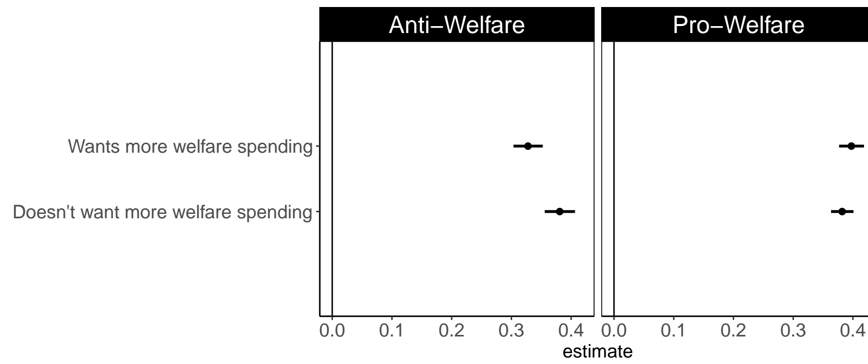


Figure 9: Environment Marginal Means by Subgroup (USA)

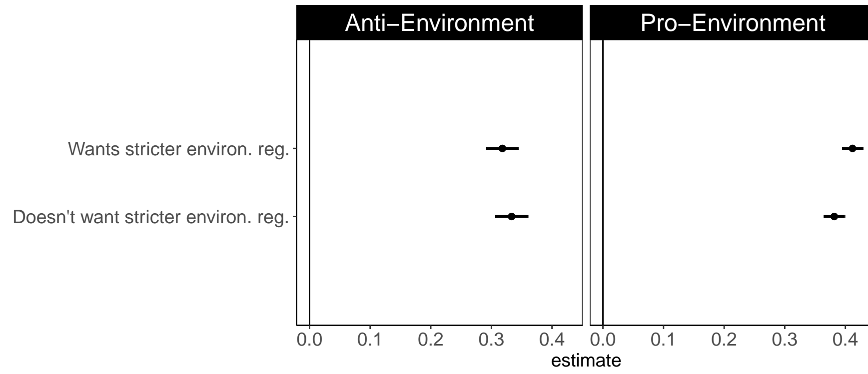
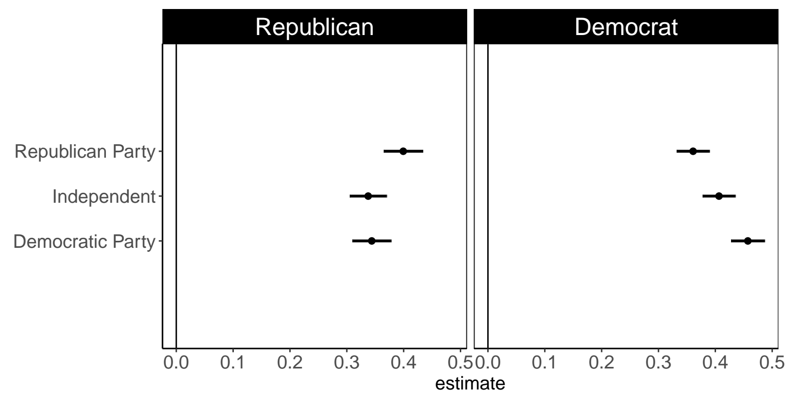


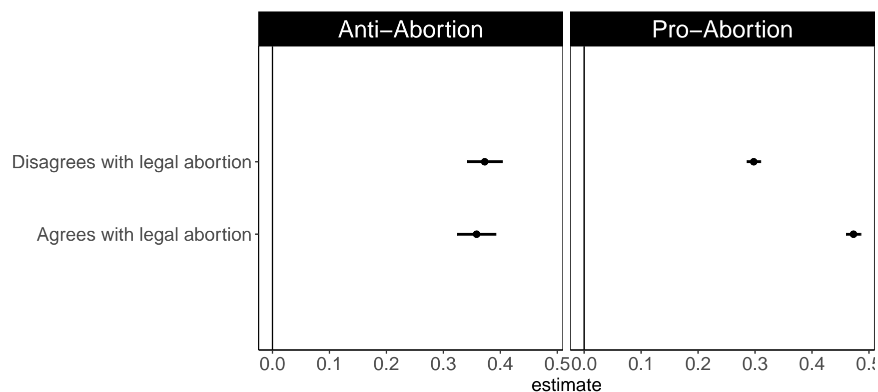
Figure 10: Partisanship Marginal Means by Subgroup (USA)



Canada

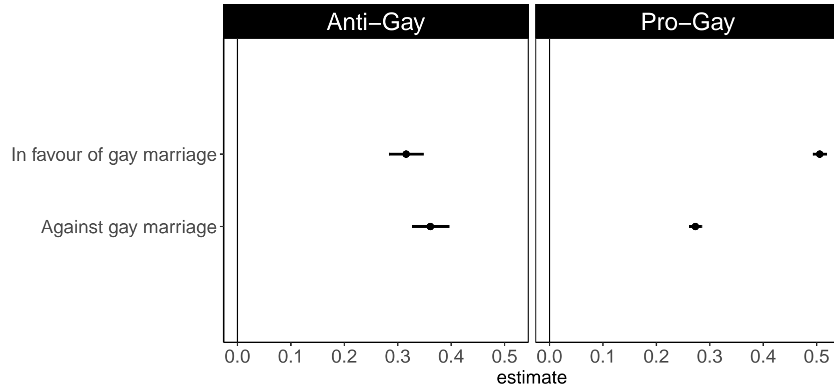
In the Canadian case, ideological conflict is more apparent among Canadians who hold the left-wing position on the four issues. For instance, when it comes to legal abortion, evidence for ideological conflict emerges only when it comes to pro-abortion Canadians, who significantly prefer immigrants who agree with them to those who do not (Figure 11). There is little difference in anti-abortion Canadians' preferences over immigrants who agree or disagree with them on the issue. A similar pattern is found for the gay marriage issue (Figure 12): anti-gay marriage Canadians do not make enormous distinctions between anti- and pro-gay marriage immigrants, but pro-gay marriage Canadians do. The gap between pro- and anti-gay marriage Canadians' views of pro-gay marriage immigrants is also striking, with pro-gay marriage Canadians preferring these immigrants 20 percentage points more than do anti-gay marriage Canadians. These preference illustrations cohere well with observed AMCE patterns.

Figure 11: Abortion Marginal Means by Subgroup (Canada)



Unlike in the United States, Canadians on both sides of the welfare spending issue privilege immigrants on their own side and punish those on the opposing side (Figure 13). But only pro-environment Canadians make meaningful distinctions between immigrants who agree and disagree with them on environmental regulation (Figure 14). When it comes to

Figure 12: Gay Marriage Marginal Means by Subgroup (Canada)



issues, then, ideological conflict is apparent in the preferences of Canadians over immigrants, but is more clearly expressed by those respondents holding the left-wing position on issues. One caveat must be added to this analysis, however: there are many more respondents holding the left- than right-wing position, meaning the analysis of the right-wing subgroups may be relatively underpowered.

The story is somewhat different when it comes to partisanship. Both Conservative and Liberal partisans significantly prefer immigrants of their own party to those who identify with the Bloc Québécois, but immigrants of other party affiliations are not necessarily punished (Figure 15). New Democrat Canadians significantly prefer co-partisans to immigrants who identify with either the Bloc or the Conservative Party, but they do not visibly punish immigrants who identify as Liberal or other (Figure 15). To summarize, then, like Americans, Canadians of all political stripes prefer immigrants who see the world on their terms and disfavor those who do not, just as the ideological conflict hypothesis would expect. However, left-leaning Canadians are more sensitive to ideological conflict than their rightist compatriots, which is not the case in the United States, where both sides are roughly equally reactive.

Figure 13: Welfare Marginal Means by Subgroup (Canada)

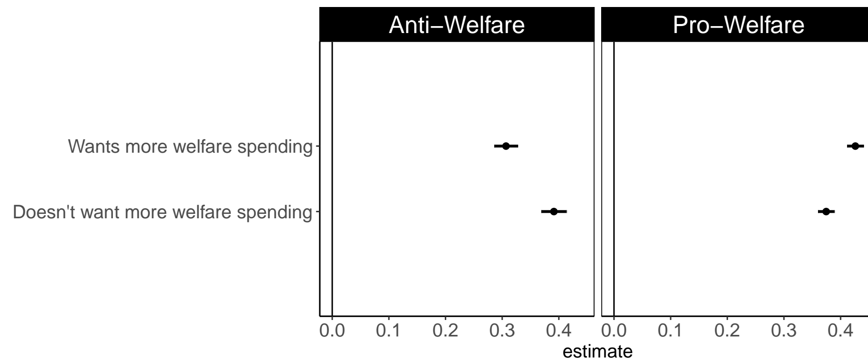


Figure 14: Environment Marginal Means by Subgroup (Canada)

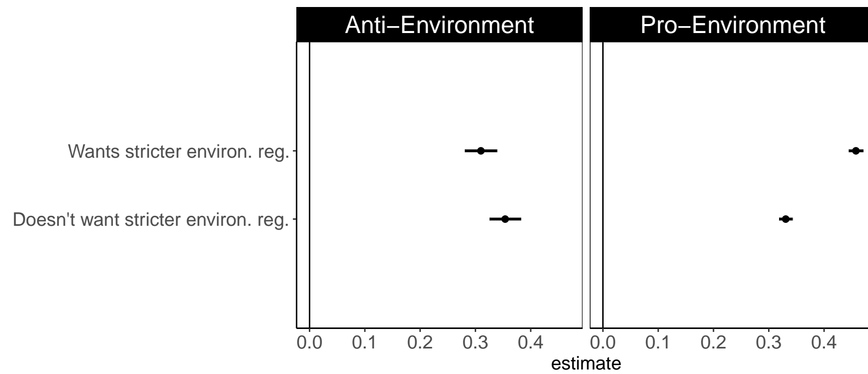
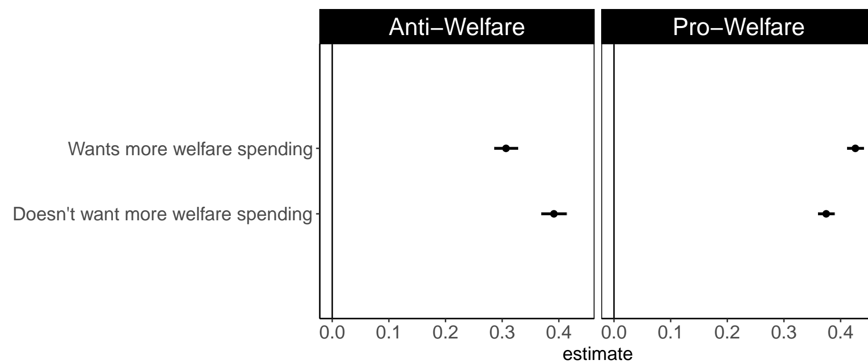


Figure 15: Partisan Marginal Means by Party Subgroups (Canada)



6 Conclusion

Research evaluating how the characteristics of immigrants affect native-born citizens' attitudes toward them have tended to find that cultural, or symbolic, factors matter more than economic factors. However, such work has tended to focus on cultural factors associated with immigrant behavior—the extent to which an immigrant speaks English, for example, or follows the rule of law. This paper has aimed to broaden the scope of this literature by asking whether immigrants' beliefs may also matter to natives' perceptions. It posed the question: how does an immigrant agreeing or disagreeing with a native-born citizen about politics affect the citizen's attitude toward that immigrant? Theoretically, it looks to the social psychology literature on ideological conflict for answers. The "ideological conflict hypothesis" (Brandt et al., 2014) suggests that what generates intergroup antipathy is not necessarily a given ideological stance that an outgroup or outgroup member may have, but rather the distance between the ingroup and outgroup's stances. In online surveys, native-born Americans and Canadians were asked their opinions on a series of four salient political issues: i) abortion rights, ii) gay marriage, iii) welfare spending, and iv) environmental regulation, as well as to indicate their political partisanship. Later, a conjoint experiment assigned them to evaluate the profiles of putative immigrants, but ones whose positions on the same issues and political parties were randomly varied. Such a design allowed me to conduct subgroup estimates in order to describe the extent to which an immigrant disagreeing with a native-born citizen–and thus being in a measure of ideological conflict with them—affects a citizen's willingness to have the immigrant remain in the country permanently. In order to benchmark the effects of ideological conflict against the effects of other often-studied cultural and economic characteristics, each immigrant profile also included six ascriptive attributes drawn from the literature, such as occupation and language ability.

The results demonstrate that in both the Canadian and American contexts, political dis-

agreement over salient issues and partisanship negatively impacts natives' attitudes toward immigrants at least as much, if not more, than traits typically associated with anti-immigrant sentiment. Conflict over moral issues, namely legal abortion and gay marriage, is associated with a decrease in willingness to allow an immigrant to remain in the country tantamount to the effects of not speaking English or not being employed, for instance. Native-born citizens who hold both the left- and right-wing positions on the four issues under study also tend to have *similar* propensities to punish immigrants who disagree with them, as the ideological conflict hypothesis would expect. In the Canadian case, left-wing citizens tend to be somewhat more punitive toward immigrants who disagree with them than are right-wing citizens; this may be due to religious accommodation debates that pitted relatively conservative religious practices of some immigrant groups against certain leftist principles.

One theoretical implication of these findings is that the left's generally positive orientation toward immigrants might not necessarily extend to immigrants who do not share its political views, especially on moral issues. Practically, this may mean that everyday tolerance toward immigrants and support for more open immigration policies among leftists may be limited if significant numbers of immigrants are either shown or thought to express relatively right-of-center political preferences. For instance, we might wonder whether Canadians with left-wing attitudes on gay marriage became less favorable toward immigrants after seeing overwhelmingly immigrant groups protesting the inclusion of gay relationships into public sex education curricula. By contrast, tolerance of immigrants and support for more open immigration policies may increase among right-wing citizens if such citizens are to find out that critical groups of immigrants support their views on non-immigration issues. In the same vein, right-wing political parties may reorientate toward capturing the vote of newly enfranchised immigrants if they know these immigrants share their views on pivotal issues. We might wonder, to give another example, about the dynamics underlying Hispanic support for Donald Trump at levels (e.g., Yglesias, 202) that seem at odds with his rhetoric, and

what such support may do to both Democratic and Republican support for immigration from Latin American countries.[10] Of course, the opposite of these scenarios may also be true: that leftists are rendered further sympathetic to immigrants by believing immigrants view the political world in the same way they do, and rightists further repulsed.

The point ultimately, though, is that immigrants' political opinions *matter*: they can meaningfully move opinions about immigrants among citizens across the political spectrum. And being a self-proclaimed leftist indeed does not inoculate one against anti-immigrant sentiment. Attitudes toward immigrants are malleable and conditional not only on the extent to which immigrants behave in ways consistent with host-country norms, but also on what they believe. Many immigrants currently in the United States and Canada cannot yet vote – but many will be able to very soon, offering an institutionalized way to convert their opinions into policies. That said, non-institutionalized forms of participation, such as boycotting, petitioning, and protesting are already available to essentially all, and can likewise affect meaningful political change. One might ask about the real-world import of the study, given that political beliefs and some ascriptive characteristics—religion, or education, for example—are likely to correlate with some of the issue positions included in the experiments. I argue that while this may indeed be the case, understanding the specific sources of attitudes is critical to adjudicating between theories of public opinion and thus has implications for real world politics. Perhaps the most obvious questions this paper brings up are i) what is the actual distribution of immigrant opinion on these political issues under study and ii) how might native-born citizens come to know immigrants' opinions on these issues? Answering the first question requires surveys that comparably, if not identically, measure opinions on the issues of interest across a diversity of sufficiently large immigrant groups in the United States and Canada. The second is often answered when different im-

[10]Of course, while many Hispanic people in the United States are native-born citizens, a substantial proportion are immigrants, given the intense growth of immigration from Latin American countries to the U.S. in recent decades (Budiman, 2020).

migrant groups gain in number and their views become more frequently reflected in news media, if not through other directly visible political actions. It would be noteworthy, however, for future work to empirically evaluate where native-born citizens do indeed believe different immigrant groups stand on salient issues. Finally, one might ask how the effects of immigrants' political opinions stack up with the effects' of their more deeply-held values, such as civic-mindedness or patriotism. While research does not yet exist to speak to this comparison empirically, given the strength of antipathy generated by the perception that some immigrants cannot or will not adopt national values documented in Levy and Wright (2020) and Figueroa (2018), and the more general strength with which people tend to cling to their values, it is doubtful that antipathy toward immigrants who are perceived as unable to integrate values-wise would be tempered by a concordance in more micro-level political beliefs. Indeed, the "funnel of causality" proposed by Lewis-Beck et al. (2008), derived from decades of research on public opinion, suggests that values causally precede, and are more fundamental and unchangeable, than more granular attitudes. That said, the values-issues comparison is empirically-testable, and attitudes might nevertheless soften due to belief-concordance toward immigrants among native-born citizens who do not believe immigrants disavow national values. Ultimately, both value- and attitude-based lines of research suggest a more complete understanding of attitudes toward immigrants should include not only behavioral, but also attitudinal, characteristics of those immigrants.

Chapter 2: Lost in Translation: How Language Compounds Wording Effects in Multilingual Surveys

Abstract: In linguistically diverse societies, interviewing survey respondents in different languages is important to generating representative impressions of public opinion. But it is not until recently that the ways in which language might affect survey response has attracted sustained attention among social scientists. A growing body of research suggests public opinion differs meaningfully by survey language. However, we still do not know much about exactly *how*, *when*, and *among whom* language influences survey responses. Using observational longitudinal data from the Canadian Election Study (CES) and a large-scale wording experiment among French and English-speakers, this paper assesses how wording effects may be compounded by language in a multilingual context. In 1997, the CES changed the wording of its feeling thermometer question about racial minorities; at the same time, evaluations of racial minorities dropped precipitously. Experimentally replicating this change, the paper shows that switching wordings induces a shift in evaluations of racial minorities, but that the effect is not the same for English and French speakers. Moreover, the effect is heterogeneous by the specific racial group asked about and by respondents' pre-treatment levels of tolerance within each linguistic group. These findings suggest wording effects can be compounded by language in ways that cannot necessarily be accounted for by the semantic differences between languages. While most studies of language effects focus on bilinguals, this paper offers some insight into the conditions under which language can matter to public opinion when the sample is not necessarily bilingual—a much broader proportion of most populations.

1 Introduction

Multilingual surveys are critical to studying comparative public opinion. Indeed, large-scale studies of public opinion, like the World Values Survey and the Comparative Study of Electoral Systems, have been developing questionnaires in tens of languages for decades. As individual countries have diversified over time, the importance of developing multilingual surveys within single national contexts has increased. In linguistically diverse societies, interviewing respondents in different languages is important to generating representative impressions of public opinion (de la Garza et al., 1992; Fraga et al., 2010; Inglehart and Norris, 2002). But it not until recently that the ways in which language might affect survey response has attracted sustained attention among social scientists (Davidov et al., 2014; Lee, 1991; Pérez, 2015; Pérez and Tavits, 2015; Pérez, 2016). Of course, researchers developing surveys are aware of potential translation issues.[1] But if people respond differently to the "same" question in different languages, is it merely an artifact of insufficient translation? Recent research suggests it is not (Boroditsky, 2001; Hong et al., 2000; Lee, 1991; Pérez, 2015; Pérez, 2016; Ross, Xun and Wilson, 2002). For example, answering questions in a genderless language leads to more positive attitudes toward women in government than answering the same question in a gendered language (Pérez and Tavits, 2015). Yet as Pérez (2015) puts it, we still do not know much about exactly *how*, *when*, and *among whom* language influences survey responses. This paper aims to contribute to answering these questions.

Specifically, it focuses on better understanding the conditions under which language matters and for whom. While most work on language effects randomizes the language of survey among bilinguals, this study moves beyond this often small and unique population to understand language effects among the (more typical) majority of people who are not

[1]McGorry (2000) outlines four suggested procedures for translating survey questions: 1) one-way translation; 2) double translation; 3) translation by committee; and 4) decentering. One-way translation—most often from English—seems to be the norm in most comparative political surveys. The World Values Survey, for example, indicates that "in many cases", but not all, double translation is used.

necessarily bilingual. It is motivated by the observation that in 1997, the Canadian National Election Study (CES), by far the most important barometer of public opinion in Canada, changed the wording of its question about feelings toward racial minorities from asking how "positively" or "negatively" a person felt toward racial minorities to asking how much they "like" or "dislike" racial minorities. Importantly, this wording change was made in both the English and French versions of the survey. Canada represents a good test case to understand language effects because it is comprised of two clear linguistic communities, the vast majority of whom (82.5%) are not bilingual (Canada, 2012). At the same time the wording change was implemented, we observe a large drop in evaluations of racial minorities, and one that is different in magnitude among anglophones and francophones. To evaluate whether this wording change produced a "wording effect" that caused evaluations to drop, and to understand whether any such wording effects may operate disparately across speakers of different languages, the study relies on a large-scale survey experiment ($n = 35000$) administered to both French and English speakers in the country. To replicate the CES change, respondents were randomly assigned to receive either the "like /dislike" version of the questions or the "positive/negative" versions. The results show that switching from the more cognitive "positive/negative" wording to the more affective "like/dislike" wording indeed generates more negative responses to feeling thermometer questions about racial groups.

Re-estimating 1997 CES attitudes once this effect is taken into consideration, we observe a fluctuation in racial attitudes more in line with previous years. Moreover, the experimental effect of the wordings is heterogeneous across English and French respondents. If this difference were mainly a matter of poor translation across the two languages or of differences in semantics, then, within each linguistic group, the effect should be constant across the different racial groups and across respondents' pre-treatment characteristics. However, I find that treatment effects vary not only by linguistic group, but also within each linguistic

group by the specific racial group respondents are asked about, as well as by respondents' pre-treatment racial tolerance levels. This variation suggests that languages can produce complex framing effects that go beyond semantics and require deeper assessments of translation quality (Pérez, 2016). Ultimately, survey research in multilingual contexts must take into consideration that language may alter respondents' answers in a way that cannot necessarily be eliminated by improving extant translation methods and may require a shift to a more extensive A/B type of testing model. The study begins by reviewing theories of language in political psychology before presenting the longitudinal observation that motivates the experimental portion of the work. The next two sections describe the experiment and present both the average treatment effects (ATEs) of the wording change and conditional average treatment effects (CATEs) for French and English-speaking groups. Using the experimental results, the study finally demonstrates what public opinion about racial minorities among anglophones and francophones would have looked like to researchers had the 1997 CES wording change not been implemented. A discussion of the implications for multilingual survey design follows.

2 Language Theory in Social Psychology

Research in cognitive psychology has extensively documented the effects of language on the way people think about the world and themselves within it (Boroditsky, 2001; Boroditsky and Gaby, 2010; Boroditsky, Schmidt and Phillips, 2003; Fuhrman et al., 2011). As such, we can reasonably expect language to also affect the way people view sociopolitical matters and answer survey questions about them. The literature on the link between language and social psychology can be roughly divided into two camps. On the one hand, some scholars conceptualize the effects of language as primarily operating through its semantic structure. This perspective has been most notably advanced by Slobin's concept of "thinking for speaking." According to Slobin (1996), the varying grammatical structures of languages

actively filter people's experiences of reality by forcing them to differentially focus on some elements of the world or situations and not others. A recent demonstration of this effect can be found in Pérez (2015), where speakers of genderless languages are shown to be more supportive of efforts to redress gender inequality because gender distinctions are less salient to them and they thus are less likely to perceive a natural gender hierarchy. Similarly, Boroditsky (2001) and Li and Gleitman (2002) have shown that the way space and time are encoded in languages affects how people respond to survey questions that involve temporal or spatial reasoning.

Yet other scholars understand the main effects of language as having much more to do with the way language invokes cultural considerations (Hong et al., 2000; Ralston, Cunniff and Gustafson, 1995). Lee (1991), for instance, writes that language effects "cannot be reduced to a technical matter about omitted variable bias, measurement error, or status deference [between the speakers of two different languages]" (p.20). Studies like Ross, Xun and Wilson (2002) give linguistic credence to Markus and Kitayama's (1991) famous essay on collectivistic versus individualistic self-construals across cultures; the authors randomly assign bilingual Chinese Canadian subjects to complete a study in either Chinese or English and show that, consistent with notions about the relatively collectivistic nature of Chinese culture compared to Anglo-American culture, those who completed the study in Chinese gave more collectivistic answers to survey questions. More directly relevant to the present paper are studies that demonstrate that orientations toward other ethnic groups vary according to the language in which one is asked about these groups and in a way that goes beyond what can be accounted for by semantic structures. Danziger and Ward (2010) show that when bilingual Arab Israelis completed a survey reporting attitudes toward Arabs and Jews, they evaluated Arabs less favorably when completing the survey in Hebrew than in Arabic. Similarly, Ogunnaike, Dunham and Banaji (2010) show that bilingual Arabic and French speakers evaluate Arabic names more negatively in French than in Arabic and that bilingual

English and Spanish speakers evaluate Spanish names more negatively in English than in Spanish. Indeed, the authors write, "languages, besides their first function of communication, are also repositories of rich cultural values that help to constitute national, ethnic, and individual identities" (p.999).

2.1 Studying Language Effects Beyond Bilinguals

Informed by these studies, I hypothesize that language conditions how different linguistic groups interpret even wording changes that may seem innocuous on their face. Moreover, I expect that differential effects of wording changes among linguistic groups are not exclusively the product of semantic differences between languages, but that they emerge by virtue of language evoking culturally-relevant considerations. The main goal of this paper is to contribute to answering the question Pérez (2015) raises about the need to better understand specifically *how*, *when* and *among whom* language affects survey responses. But it goes about the task differently than the studies thus far reviewed. Most of these examine how language affects bilinguals' responses to survey questions. This study examine how a real-world question-wording change in a widely-used national survey of political opinion affected respondents depending on which of two—English or French—linguistic communities they belonged. While studies that randomly assign language of survey to bilinguals have the advantage of being able to causally identify the effects of language itself, they are importantly limited with respect to external validity in that their inferences only apply to bilinguals (see: Gerber and Green, 2012; McDermott, 2011). Most people in most countries are not bilingual and understanding the role language can play in shaping the responses of the non-bilingual majority is an important task, especially if we expect linguistic diversity to increase within countries at a greater rate than second or third language acquisition. It is important to also note that in most countries, bilinguals are likely very systematically different from monolinguals, whether by virtue of being better educated, having higher incomes, being immigrants,

or living in more linguistically diverse areas. Thus, instead of randomly assigning language, this study randomly assigns the wording of a question of perennial political importance—feelings toward racial groups—to people who self-select into taking a survey in either English or French in Canada. The first goal is to evaluate whether there indeed is a wording effect; subsequently, I evaluate whether this effect is conditional on respondents' language. If it is, a third goal is to shed light on whether this conditionality may be better attributed to semantic differences between the two languages or whether cultural encodings are likely to be at play.

2.2 Wording, Language, and Racial Affect

This study is motivated by the observation that in 1997, the Canadian National Election Study (CES)—the primary source of information about the public opinion of Canadians—changed the wording of its feeling thermometer question about racial minorities in both the French and English versions of its survey.[2] From 1968 to 1997, the CES asked "Using the 0 to 100 scale, where 0 means *very negative* and 100 means *very positive*, I would like you to tell me how you feel about racial minorities." In French, the same wording applied, substituting the words "positif" and "negatif" for "positive" and "negative." Then, from 1997 onward, the question read "Use any number from 0 to 100. 0 means you *really dislike* racial minorities, and 100 means you *really like* racial minorities." In French, the valenced terms were changed to "aimez vraiment beaucoup" (really like) and "aimez vraiment pas du tout" (really dislike). Given how widespread CES usage is among Canadian academics, this question certainly informed a great number of studies about prejudice in the country since the 1960s. Whether a switch from a positive/negative wording to a like/dislike wording had an effect on observed racial attitudes in the country is an empirically-testable question.

[2] It is not immediately clear why this change was made. The reasoning behind it is not documented in the materials accompanying the surveys. The Principal Investigator for the survey at the time also does not recall why the change was made, according to personal communications.

A quick review of the use of feeling thermometers toward social groups in public opinion research shows a great deal of wording heterogeneity. While the wording in the American National Election Study (ANES) has remained relatively consistent since 1964[3], at least in English, this is often not the case when it comes to comparative surveys. And different surveys use different wordings. The Comparative Study of Electoral Systems (CSES), for instance, uses feeling thermometers where the lowest point is labeled *strongly dislike* and the highest *strongly like*, compared to the ANES' *warm/cold* labeling. Given that the terms 'warm' and 'cold' cannot be applied to feelings about social groups in many languages, it makes sense to not apply the original ANES labeling in comparative or multilingual contexts. Indeed, the ANES does not use the warm/cold formulation in the Spanish version of its survey and the CES does not use it in its French version.[4] In addition to being inconsistent over time, feeling thermometer wordings in comparative surveys are also often inconsistent across the countries they survey within a given language, given that different teams typically work on different countries. The CSES, for instance, is inconsistent across francophone countries, with the high points in Switzerland, for instance, being labeled "sympathie [que] vous éprouvez" (feel sympathy), while in France they are labelled, "appréciez," (appreciate) and in Canada, "aimez" (like, or love).

Returning to the motivating case of the 1997 wording change in the CES, Figure 1 plots Canadians' average scores on the racial feeling thermometer question in each year the question was asked. Between 1968 and 1984, attitudes toward racial minorities got colder by about four points on average but tended to hover around the 69- to 74-degree mark. The attitudes of the anglophone majority and the francophone minority in the country were not perceptibly distinguishable from the population average for the majority of the period, with

[3]Minor changes are noted on pages 63-64 of the ANES Time Series Cumulative Data File (1948-2012) codebook: www.electionstudies.org.

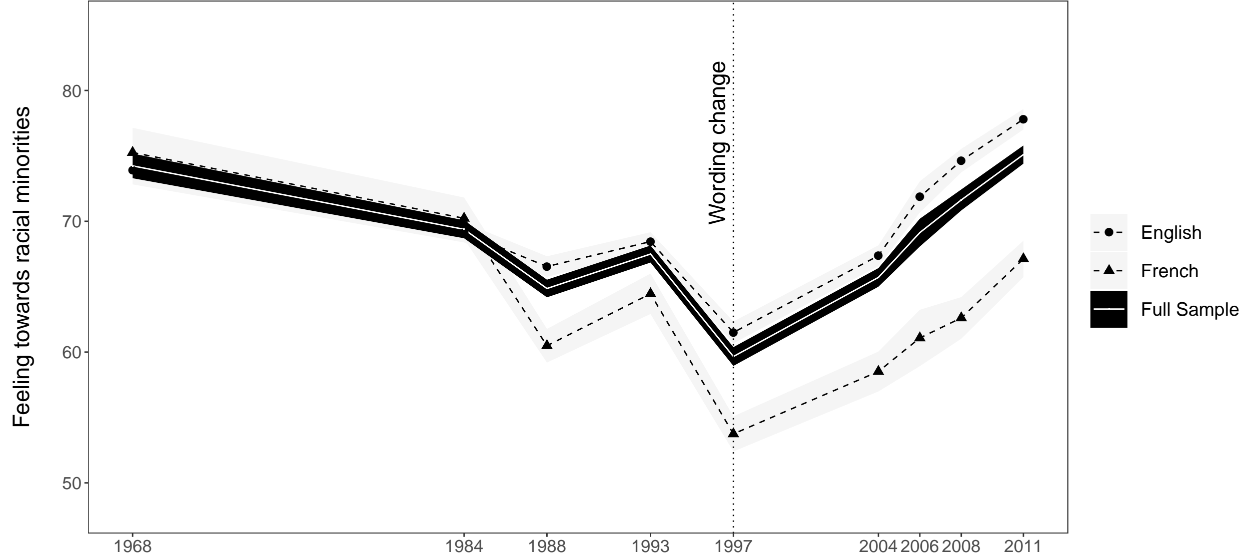
[4]All CES campaign-wave studies surveyed a representative sample of Canadian residents, in earlier decades through face-to-face interviews and in recent decades through random-digit-dialing.

a slight divergence among francophones only being driven by the 1988 time point. We see, however, that after 1997, racial attitudes in the country became precipitously more negative (ie. colder) overall (by about 8 points) and that this decline was slightly greater among francophones than anglophones. The question is whether this change reflected real shifts in attitudes or simply respondents' interpretations of the question. There is reason to believe that attitudes toward racial minorities really did become more negative during this period among both francophones and anglophones. The 1997 wording change came immediately after the 1995 Quebec secession referendum, where the then-premier of the province famously attributed the failure of Quebec sovereignty to "money and the ethnic vote." This rhetoric may have increased hostility toward those perceived as being "ethnic" among the largely francophone Quebec sovereignists. At the same time, English-speaking western Canada saw the rise of right-wing populist parties who similarly espoused anti-racial minority views. We can isolate what portion of this observed decline in warmth toward racial minorities is attributable to the 1997 wording change in the survey versus what portion is the product of genuine attitudinal change through an experimental intervention.

2.3 English and French in Canada

Research has long shown that people tend to be susceptible to different wordings of the same question because different wordings evoke different considerations, or "frames" in people's heads, leading them to construe issues and respond to questions differently (Schuman and Duncan, 1974; Schuman and Presser, 1977; Lau, Smith and Fiske, 1991; Druckman, 2001). For example, in a well-known study, Sniderman and Theriault (2004) demonstrate that people are 40% more likely to support a hate group being able to hold a rally when the question is prefaced with the statement "Given the importance of free speech" than when it is prefaced with "Given the risk for violence." The wording change in our case is more subtle given the basic meanings of the two wordings are on their face the same. We might characterize the

Figure 1: Racial Attitudes Over Time in the Canadian Election Study



Source: Canadian Election Study (1968-2011).
Note: White and patterned inner regions represent average opinions and shaded regions represent 95% confidence intervals around these means.

difference in terms of one of the wordings being more cognitive – evaluating how positively or negatively one feels toward a group – versus more affective – how much one likes or dislikes a group. In line with what we observe in the CES data, I expect the affective wording to lead to more negative evaluations of racial minorities, though at this stage remain agnostic as to why. Speculatively, it might be the case that cognitive formulations make people more likely to consider social desirability in answering the question in a way they would not when simply asked to tap into their feelings – what they like or dislike. To the extent that thinking in a language evokes the cultural considerations associated with that language (Pérez, 2015), I also expect wording effects to operate heterogeneously between English and French respondents. The fact that English and French Canada have two different cultures is widely recognized (Quebec is considered both informally and legally a "distinct society"), and the difference and separation between the two has led both scholars and laypeople to characterize the Canadian cultural landscape as culturally consisting of "two solitudes." One might

counter that we may observe heterogeneous treatment effects across anglophones and francophones not because different languages evoke different cultural considerations, but simply because the wording change constituted a semantically larger change in French than in English. That is to say, the phrases "aimez vraiment beaucoup" and "aimez vraiment pas du tout" in French have stronger connotations than the English "like" and "dislike;" in fact they might be translated as "really like/dislike a lot," or even "really love/really do not love a lot." In order to test whether heterogeneity in wording effects across the two languages is simply a matter of such semantic differences, I evaluate whether effects are constant across different racial groups and across respondents' pre-treatment tolerance levels. If the differences are indeed only a matter of semantics, and not language evoking cultural considerations, then we should see constant effects across racial groups and respondents' pre-treatment tolerance levels. If effects in fact vary, then there is reason to believe language indeed evokes cultural considerations–or acts as a 'frame' by bringing cultural considerations to the top of the head. Of course, this is not a silver bullet test, no less because it only produces observational insights – but it can provide meaningful suggestive evidence in a context where language itself is not randomly assigned.

3 Experimental Design

To assess the causal effect of the CES wording change of the feeling thermometer question toward racial minorities, an experiment replicating the change was embedded in an online survey administered in 2015 by Vox Pop Labs ($n = 35000$).[5] Survey respondents chose whether to take the survey in English or French, and then were randomly assigned to one

[5]VPL maintains an online panel of respondents who are recruited through the Vote Compass voting engagement application sponsored by the Canadian Broadcasting Corporations (CBC) before each provincial and federal election (since 2011). Raking weights are applied to the data using 2016 Census values on age, income, and education. Distributions of both unweighted and weighted data on key demographics look similar to those of the 2011 CES, and weighting does not ultimately produce any significant changes in estimates.

of two experimental conditions. Half (within each linguistic group) were assigned a series of feeling thermometer questions about nine racial minority groups with the affective wording (in English) "How do you feel about [the group]? Use any number from 0 to 10. 0 means *very negative* and 10 means *very positive*." The other half were assigned the same feeling thermometer questions, except with the question worded as "How do you feel about [the group]? Use any number from 0 to 10. 0 means you *really dislike* the group and 10 means you *really like* the group."[6] As in English, the French wordings of the questions mimic the pre- and post-1997 CES wordings of the questions and are randomly assigned within French respondents. The design thus effectively block-randomizes by language of survey. Those who received the *positive/negative* wording are considered the control group, given it is the baseline wording from which the CES deviated in 1997, and those who received the *like/dislike* wording, the treatment group. Instead of simply asking about "racial minorities," the survey asks about specific racial groups in order to gain leverage over disentangling whether language-based effect heterogeneity is a matter of semantics or cultural frames languages evoke. French-speaking Quebec and the rest of (predominantly) English-speaking Canada are home to different racial minority populations largely due to different patterns of immigration. They have also approached immigration and integration differently, and ultimately different immigrant groups have factored into public debates in quite different ways (see, for example: Bilodeau, Turgeon and Karakoc, 2012; Dufresne, Jeram and Pelletier, 2014; Dufresne et al., 2019; Turgeon and Bilodeau, 2014).

Relying on the potential outcomes framework (Rubin, 1974), we can denote by $Y_i(0)$ the outcome a subject would express if he or she received the positive/negative wording and $Y_i(1)$ the outcome a subject would express if he or she received the like/dislike wording. Recovering the Average Treatment Effect (ATE) involves estimating $\tau = E[Y_i(1)] -$

[6]In analyses, all 10-point measures are multiplied by 10 to mimic CES response categories that are scaled from 0 to 100; there is no evidence to suggest that feeling thermometer responses are not scale invariant in this sense.

$E[Y_i(0)]$, given that the probability of assignment to treatment is identical within French and English-speaking subject groups. Recovering the Conditional Average Treatment Effect (CATE) within language groups involves estimating $\gamma = (E[Y_i(1)] - E[Y_i(0)]|X_i)$, where X = $\{French, English\}$. The main advantage of this experimental design is that, although it does not randomly assign and thus causally identify the effect of language, it is not constrained to only samples of bilinguals — a limitation of extant work on language effects. It does, however, causally identify the effect of question framing within linguistic groups and thus offers insight into the conditions under which language may matter among people who may be monolingual *or* bilingual — a sample that is more representative of the populations of most countries. By collecting responses from over 35,000 people, I also overcome the issue of low statistical power present in the small-scale experiments ($n < 50$) that dominate the study of language effects (see, for example: Boroditsky, 2001; Fuhrman et al., 2011) and the often associated "college sophomore" issue (Sears, 1986). The large sample also powers the experiment to reliably detect heterogeneous treatment effects.

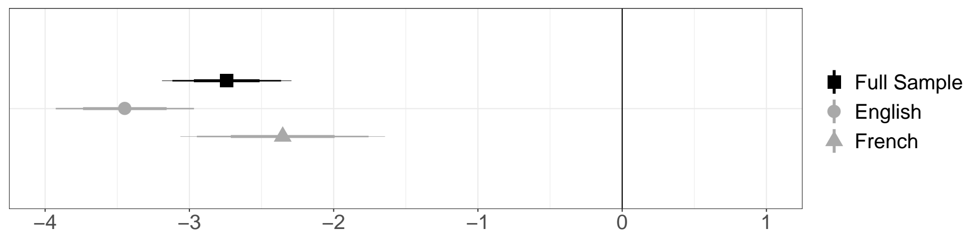
4 Results

4.1 Wording Effects in English and French

Inconsistency of question wording and response labeling in different linguistic iterations of the same survey indicates that researchers might not think such inconsistencies matter. Unfortunately, the results presented here suggest they do, at least to some extent. Changing the wording of questions asking people to evaluate racial minorities from the more cognitive *positive/negative* to the more affective *really like/really dislike* causes a decline in positivity toward racial minority groups comparable to many socio-demographic determinants of these attitudes. The square black point estimate in Figure 2 represents the ATE and corresponding confidence intervals from a t-test comparing attitudes toward racial minorities among

those who received positive/negative wording versus the like/dislike wording over the pooled sample. Those exposed to the *like* wording of the feeling thermometer questions tended to report feelings about 2.7 percentage points less positive than those who received the *positive* wording in the sample as a whole. The outermost bars of the point estimate represent 95% confidence intervals, and treatment differences between anglophones and francophones are very nearly statistically significant at this level.[7] This is about one third of the overall decline observed after the 1997 CES wording change. Given that treatment assignment was randomized within anglophone and francophone samples, we can also causally infer a statistically significant treatment effect within each linguistic group.[8] The overall takeaway though is that being asked whether you like or dislike racial minorities instead of being asked how positively or negatively you feel towards racial minorities causes a small, but significant, decrease in the warmth of feelings toward racial minorities. This effect is significant among both anglophones and francophones.

Figure 2: Average Wording Treatment Effect (All Racial Group Questions Combined)



Source: Vox Pop Labs, 2015.

The differences *between* treatment effects in anglophone and francophone samples are significant at the 90% confidence level and very nearly significant at the 95% confidence

[7]Relaxing linearity assumption and re-estimating the confidence intervals via randomization inference yields effectively the same results.

[8]The chapter Appendix also contains estimates of treatment effects on feelings toward groups that do not represent racial minorities in the Canadian context for comparison – Whites, Christians, and atheists. Therein, we observe statistically significant treatment effects within anglophone and francophone samples (and between these samples) only when it comes to whites.

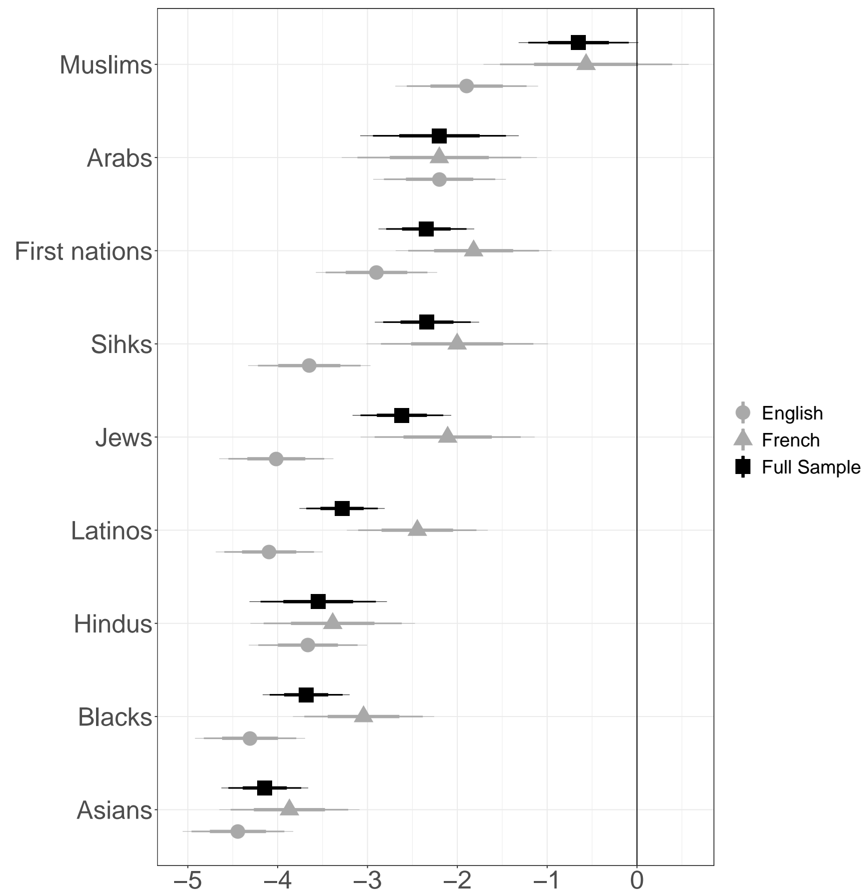
level (Figure 2). This result suggests that wording changes likely affect different linguistic groups differently.[9] Yet it also runs counter to what is observed in the CES data, where it is francophones, not anglophones like in Figure 2, that become more negative toward racial minorities when exposed to the "like/dislike" wording. Given that both the CES and between-linguistic-group estimates from our experiment are observational in nature, we do not have a silver-bullet explanation of this discrepancy. Once again, this is a limitation of research aiming to understand language effects among those who are not necessarily bilingual. But it ultimately does give us a clue that question wordings about intergroup attitudes are sensitive to framing effects and that these effects are not constant across languages.

4.2 Cultural-Linguistic Heterogeneity

There is thus some evidence of heterogeneity of wording effects by language, as the growing body of theory on language effects would predict. In terms of answering *how*, *when*, and *among whom* language might influence survey responses, the results thus far suggest that even when researchers try to render functionally equivalent translations, these translations might produce different wording effects among anglophones and francophones. To offer some insight into how this might function, through the semantic structures of languages or through the cultural frames they evoke, I disaggregate effects among the nine different racial minority groups about which the survey asked. As Figure 3 shows, when it comes to feelings toward three of the nine groups—Latinos, Jews, and Blacks—treatment effects among anglophones and francophones are statistically different from one another. The pattern of differences among anglophones and francophones appears relatively sensible given the literature on intergroup attitudes in Canada, though again, we can only speculate as to causes (Bilodeau, Turgeon and Karakoc, 2012; Dufresne, Jeram and Pelletier, 2014; Turgeon and Bilodeau,

[9]In Table 2 in the chapter Appendix, I re-estimate the ATE and CATEs by language in a linear regression, adjusting for individual-level sociodemographic covariates that may differ between treatment and control or linguistic groups. The results remain substantively unchanged.

Figure 3: Average Wording Treatment Effect by Individual Racial Group Question

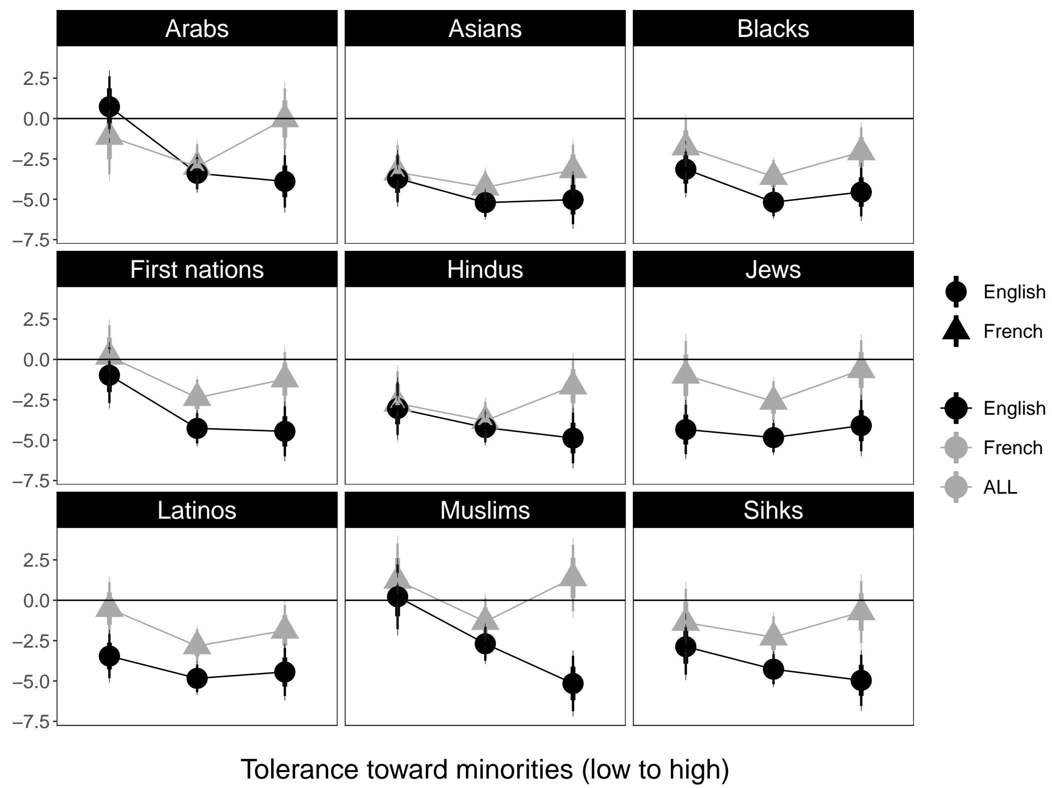


Source: Vox Pop Labs, 2015.

2014). When it comes to racial minorities who have long been members of Canada's ethnic makeup (Latinos, Jews, Blacks) and about whom social norms are likely stronger, there are differences in treatment effects among anglophones and francophones, with the treatment being associated with more negative affect in anglophones. Yet when it comes to groups who form more recent sources of immigration to the country (e.g., Sikhs, Hindus, Muslims), differences among francophones and anglophones are less clear. As discussed in the theoretical section of the paper, there may be differential norms related to different groups encoded in the different cultures to which French and English correspond in Canada and the different debates involving these groups that have emerged across provinces. Ultimately, though, the fact that there is heterogeneity in treatment effects between francophones and anglophones across specific racial groups suggests that the effect of the treatment overall does not differ in the two languages purely as a result of differences in semantics.

A second clue that differences in treatment effects between anglophones and francophones may occur because language induces cultural considerations comes from the fact that treatment effects also vary between francophones and anglophones by pre-treatment attitudes on interethnic tolerance. Questions about the immigration and integration of mostly racial minorities have represented perhaps the most heated lines of division among the two linguistic groups in recent years (Dufresne et al., 2019; Giasson, Brin and Sauvageau, 2010; Turgeon and Bilodeau, 2014). Could the more affective "like/dislike" wording of the feeling thermometer questions be associated with a different effect among those with strong views about these issues? I expect having strongly positive prior views on immigration and integration questions, and thus perhaps being more in tune with more liberal social norms with respect to tolerance, to meaningfully distinguish treatment effects among anglophones and francophones when it comes to the groups who form recent sources of immigration and who have been at the center of integration debates — that is, Hindus, Sikhs, Jews, and Muslims. Differences of means among anglophones and francophones are plotted in Figure 4 by social

Figure 4: Average Wording Treatment Effect by Racial Group and Pre-Treatment Tolerance



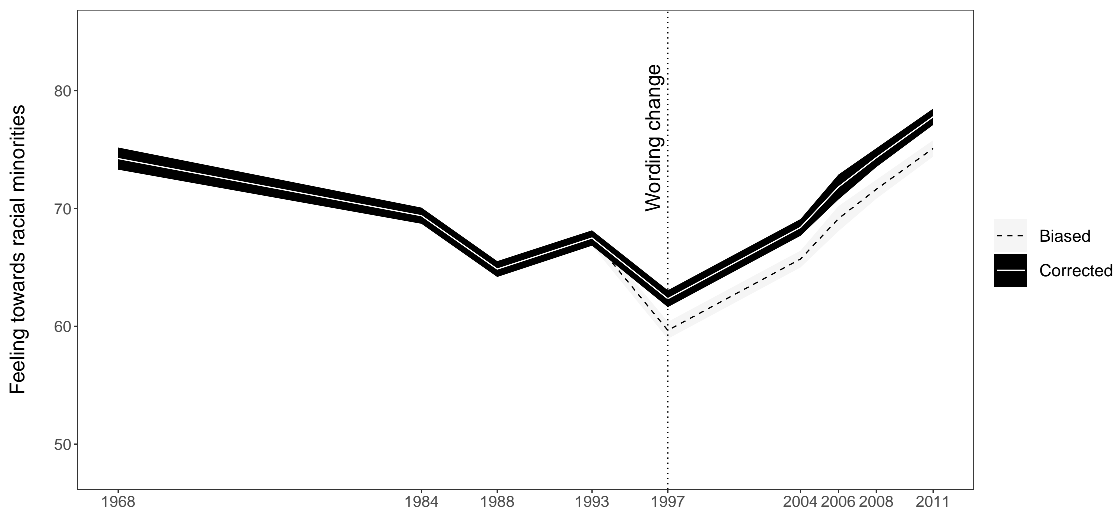
Source: Vox Pop Labs, 2015.

group and by (on the x-axis) a combined index of support for increased immigration and support for more religious accommodation (labeled 'tolerance'). The expectation is born out and is generally in support of the hypothesis that social norms about minorities are more strongly associated in English, given we observe the greatest decrease in positivity toward Hindus, Sikhs, Jews, and Muslims among anglophones who are most tolerant in general, and one that is statistically different from that of francophones (there is no difference between anglophones and francophones otherwise). Once again, though, we lack the data here to directly test hypotheses about social norms and discussion to that end remains only suggestive. The main takeaway from this set of findings, though, is that not only do wording effects differ across French and English speakers when we disaggregate to the nine different racial group questions, but also when conditioning on English and French speakers' pre-treatment attitudes toward interethnic tolerance. This pattern offers further evidence in support of the notion that language effects are more likely to be the product of a language-culture connection between languages than differences in the technical or grammatical structures of languages.

4.3 What If? An Observational Correction Based on Experimental Results

At its base, this study harnesses a large-scale experiment to estimate the effect of wording changes on peoples' feeling thermometer evaluations of racial minorities. Based on these results, we can go back and ask: what if the survey that motivated this study had never changed its question wording? What longitudinal trend in attitudes would we have observed? To answer this question, we can re-construct 1997 attitudes in the CES after taking into consideration what we now experimentally know about feeling thermometer wording effects. That is to say, we can 'add' the overall combined treatment effect of about 2.7 that we observe in the experiment to the racial feeling thermometer trend line in the longitudinal CES. Based on estimates from this study, if the CES had not changed its feeling thermometer

Figure 5: Corrected Racial Minority Attitudes Over Time



Source: CES and Vox Pop Labs.

wording in 1997, the trend in citizen attitudes from 1993 to 1997 would not have wavered significantly more dramatically than it had in other years (Figure 5). Since the 1960s, we would have generally seen an increase in the positivity of citizens' attitudes toward racial minorities.[10]

This outcome is surprising given that positivity declined more pronouncedly among francophone respondents of the CES between 1993 and 1997 than among anglophone respondents. It may thus well be the case that during that time period francophone attitudes toward racial minorities did indeed become more negative, perhaps as a result of contextual factors, like anti-immigrant rhetoric around the Quebec referendum. The fact that positivity increased in subsequent survey years retaining the same wording speaks to the presence of context effects. But ultimately, the effects of wording changes are at least to some meaningful

[10]Given I am imputing a 1997 value based on 2015 data, one might wonder whether demographic differences between samples taken at two different time frames may make the imputation inappropriate. A priori, there is no obvious reason to suspect that the effect of wording would change within language groups over time. And demographic change tends to come slowly. That said, I applied raking weights to the 2015 data based on 1997 CES language, age, education, and income values. The results are not meaningfully different and are entirely in the same direction.

extent compounded by linguistic factors. When taken together with the observed longitudinal differences among francophones and anglophones in the CES, this finding offers support for the hypothesis that wording changes differently affect French- and English-speakers' reported feelings toward racial groups.

5 Discussion

Comparative surveys, as well as surveys in multilingual contexts, often must be designed in multiple languages to ensure broad sampling. While researchers designing such surveys are aware that care must be taken when it comes to translation to maintain question equivalence, only recently have scholars begun to take empirical stock of how, exactly, language might affect survey responding. If we see that people respond differently to the same question posed in different languages, how might we account for this discrepancy? Poor translation, of course, does occur, but even the best translation efforts may be belied by the cognitive or sociopsychological dynamics underpinning how subjects respond. To the extent that even 'equivalent' wordings in different languages lead speakers of those languages to respond in systematically biased ways, what might the source of this differential be? Two main lines of reasoning have shaped our thinking about how language might factor into the way people understand and respond to survey questions: on the one hand, equivalent wordings may elicit different responses in speakers of different languages because there is something about the semantic structure of the language that forces speakers to focus on some elements of the wording and not others, bringing different considerations to top of mind; on the other hand, the effects of such wordings across languages might be less about how the languages are structured and more about the cultural considerations they evoke. Because research on language in survey responding is relatively nascent, we still do not know much about exactly *how*, *when*, and *among whom* language influences survey responses (Pérez, 2015). This paper aims to contribute to answering these questions in the Canadian case, motivated

by a real-world change in the national public opinion survey of record in the country, the Canadian Election Study (CES).

In 1997, the CES changed the wording of its feeling thermometer question toward racial minorities from a formulation that asked people how "positively" or "negatively" they felt toward racial minorities to one that asked how much they "liked" or "disliked" racial minorities. Importantly, this wording change was also made in the French version of the survey. Immediately after, we observe an unprecedented drop in the warmth of people's racial attitudes. Moreover, the magnitude of the drop is different for anglophones and francophones. To analyze whether this wording effect had i) a causal effect on racial attitudes and ii) whether observed differences by linguistic group can be better attributed to linguistic semantics, or the way the languages culturally encode ideas, I conduct a survey experiment replicating and extending the change. The results show that switching from a "positive/negative" wording to a "like/dislike" wording causes a substantial drop in feelings toward racial minorities among both anglophones and francophones. If this effect were mainly a matter semantics, or even poor translation, across the two languages, then treatment effects within francophone and anglophone samples would be constant across the nine separate racial groups about which the survey asks. They would also be constant across respondents' pre-treatment intergroup tolerance levels. But they are not. For both anglophones and francophones, wording effects are heterogeneous across specific racial minority groups, offering greater evidence in favor of the theory that wording effects operate through the way languages encode relevant sociocultural considerations. And indeed, patterns of heterogeneity comport with patterns of immigration across English- and French-speaking Canada and cultural dynamics we may expect to be at play. While these are not definitive answers to the question of whether language effects are predominantly semantic or cultural, they suggest that wording effects in surveys may vary by language not only because of semantics, which might be more easily reduced through translation efforts or at least accounted for in estimation, but because of the

different ways cultural concepts are encoded in language, which might be slipperier. When we re-estimate 1997 CES attitudes once average treatment effects from the experiment are taken into consideration, the decline in warmth of feeling toward racial minorities is not as precipitous and more consistent with the magnitude of fluctuations in previous years. To return to the questions of *how? when?* and *among whom?* language effects might matter, the present study offers this: language effects are more likely to matter by virtue of bringing certain cultural considerations to top of mind when respondents answer survey questions; this seems to be the case when 'equivalent' changes are made in two languages–in this case, French and English–and among people who are not necessarily bilingual, a much broader proportion of the populations of most countries.

Ultimately, scholars must take into consideration that languages themselves may produce 'framing' effects that complicate survey question design. Without A/B testing wording changes, or even just different articulations of wordings that on their face look the same in different languages, studies risk reporting results that are biased either simply by the framing effects of language, as other researchers have noted, but also results whose linguistic bias is compounded by differential wording effects across languages. In the Canadian case, counterintuitively to what the 1997 CES data descriptively illustrate, the 'like' wording effect actually leads to a greater drop in anglophones' positivity toward racial minorities, although francophones baseline positivity is lower. We likely observe a greater drop for francophones in the descriptive graph, despite a smaller effect of the wording change, because francophones' attitudes really were more negative in that year – again, not implausible in light of anti-immigrant rhetoric following the failed 1995 secession referendum. When it comes to anglophones, research relying on pre-1997 data would have overstated their positivity, while post-1997 work would have understated it. Taken together, research using the 1997 data onward paints a grimmer picture of Canadians' racial attitudes than was truly the case. Of course, A/B testing every question to be translated into multiple languages

is an immensely resource-intensive, if not impossible endeavor. But language clearly does matter, and more work needs to be done to narrow the conditions under which it does. The contribution of this paper here is principally to point out that language can compound the bias that wording effects can already introduce when there is insufficient evidence about how wordings compare.

If language effects are indeed the product of the fact that different languages are differently associated with different cultural concepts, then rigorous qualitative translation attempts may still fall short. Scholars must instead actively integrate language as a covariate in research design in order to render more accurate representations of latent public opinion. Similarly, translation efforts might also begin incorporating the kind of A/B testing, especially when it comes to sensitive subjects or topics about which respondents are expected to know little, that are common in industry today. Such an endeavor is increasingly pressing as societies diversify linguistically and producing surveys in multiple languages within a given context or set of contexts is necessary to generating more representative images of a population's attitudes. While strides have been made in understanding language effects among bilinguals, this group tends to be small and systematically different from the much larger population of monolinguals. By actively incorporating monolinguals, research on survey design and the role of language in political psychology can make further strides in our understanding of the conditions under which, and among whom, language matters.

Chapter 3: Celebrity Messaging to Counter Vaccine Hesitancy: A Study Among Black Americans During the Covid-19 Pandemic

Abstract: Like other historical disease pandemics, the solution to the covid-19 pandemic has been inoculation through vaccination. As soon as a covid-19 vaccine was developed, and eventually approved, government and medical officials in the United States have encouraged Americans to get vaccinated as soon as they are able. A barrier to full vaccination, however, has been vaccine hesitancy, or a delay in the acceptance or refusal of vaccines despite the availability of vaccination services. Such hesitancy has been found to be higher among Black Americans than others, presenting a potential barrier to immunity among this group (Funk and Tyson, 2021; Ndugga et al., 2021). Significant resources have thus been devoted to countering hesitancy, and one particularly visible strategy has been to use Black celebrities as vaccine spokespeople. Relying on a survey experiment, this study evaluates whether televised instances of Black celebrities are indeed effective at combating vaccine hesitancy among Black Americans, in general and as compared to non-celebrities. Respondents to an online survey are randomly assigned to watch a video in which either film and TV star Tyler Perry, or New York City nurse Sandra Lindsay (the first person to get vaccinated against covid-19 in the U.S.) take a covid-19 vaccine and discuss its benefits. A third are assigned to a control video condition, where Tyler Perry talks about his new movie studio. The findings demonstrate that while distrust in government and the medical establishment among Black Americans is indeed relatively high, getting around any effect this might have

on vaccine hesitancy by relying on alternative spokespeople, like celebrities, is not clearly effective. Analyses do not detect treatment effects of the Perry video, or the Lindsay video, on respondents' intentions to get a covid-19 vaccine, nor on their attitudes about the safety and efficacy of such vaccines. This study represents the first experimental test of which I am aware of the effectiveness of celebrity messaging on vaccination in the United States, and the first such among Black Americans specifically.

1 Introduction

Since covid-19 vaccines became widely available in the United States, Black communities in the have persistently significantly lower covid-19 vaccination rates than non-Black communities (Funk and Tyson, 2021; Ndugga et al., 2021). While differential rates of access to vaccine sites likely explains a portion of the discrepancy, early research has shown individual hesitancy to take the vaccine also matters a lot (CDC, 2021; Langer Research Associates, 2020). Early work has suggested this hesitancy may be driven by the belief that the government and medical establishment cannot be trusted to look out for the interests of Black people (Bogart et al., 2021; Ellis, 2020; Gramlich and Funk, 2020; Langer Research Associates, 2020; Savoia et al., 2021). Reasonably so, given the United States' history of endangering Black lives in the pursuit of medical scientific discovery and the neglect of Black communities by health care systems. To encourage vaccine uptake, governments, organizations, and community leaders have made extensive efforts to combat hesitancy among Black Americans specifically. At least implicitly, if not explicitly, a lot of these interventions have relied on the notion that racial in-group members may have greater persuasive power than out-group members — the campaign spokespeople were, in the vast majority of cases, also Black. Much of this messaging was televised: from talk shows interviewing Black nurses and doctors about the vaccine to news outlets sharing videos of national and local celebrities like Samuel L. Jackson receiving one, there has been no shortage of media efforts relying on Black leaders

to convince Black Americans to get vaccinated.

Despite how widespread, visible, and certainly, costly, these campaigns have been, there has yet to be any empirical evidence demonstrating their effectiveness. This study aims present a first step in this direction. Primarily, it aims to evaluate whether messaging by a Black celebrity might impact Black Americans' intentions to vaccinate and opinions about vaccines, and how any such impacts might compare to those when messaging is from a Black non-celebrity. An experiment embedded in an online survey taken by a nationally representative sample of Black Americans randomly assigned respondents who had yet to receive a covid-19 vaccine by April 2021 to watch short videos of either film and TV personality Tyler Perry, or Sandra Lindsay, the New York nurse who was first to receive a covid-19 vaccine, receive a vaccine and promote its benefits. A control group watched a placebo video wherein Tyler Perry talks about his new film studio. Both treatment videos had aired on national television and had also received hundreds of thousands of views on YouTube. Prior to the experimental intervention, respondents' opinions about medicine in general, Black identity, and trust in government were measured to get a sense of respondents' attitudes toward these potentially moderating factors. The results fail to find any meaningful effect of watching either treatment video on the probability a respondent would take a covid-19 vaccine or on respondents' evaluations of the safety and efficacy of such vaccines. Nor is any effect of the treatments found on the probability an individual would recommend to others to get vaccinated.

Evaluating whether the messenger indeed matters when it comes to reducing covid-19 vaccine hesitancy can help shed light on the extent to which different figures might be useful in promoting other medical, and even more broadly socially positive, efforts that stand to benefit the Black community in the United States, as well as other subgroups in general. The paper first offers an overview of extant research on encouraging vaccine uptake, as well as the role of celebrity and co-ethnicity in persuasion. The experimental design and

estimation strategy is introduced and descriptive and causal findings from the survey are presented. The conclusion discusses the implications of the findings, namely the potential role of celebrity messengers in future health care campaigns under circumstances of low public trust in medical and governmental institutions.

2 Countering Vaccine Hesitancy

By the timing of this study, eight vaccines had been approved around the world to protect against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that produces the disease known as covid-19. In the United States, three were in use: those made by Pfizer, Moderna, and Johnson & Johnson. Yet over 30% of Americans were still indicating they were not willing to take a covid vaccine, as were 40% of Black Americans (Funk and Tyson, 2021). Refusal to get vaccinated under conditions where the barriers to accessing a vaccine are relatively low is often attributable to hesitancy on the part of the individual, for a diversity of psychological, cognitive, emotional, or spiritual reasons. The WHO's Strategic Advisory Group of Experts on Immunization (SAGE) defines vaccine hesitancy as a delay in the acceptance or refusal of vaccines despite the availability of vaccination services — a definition applied in this paper.

There has, of course, been extensive research evaluating interventions to boost vaccination rates among the vaccine hesitant prior to the covid-19 pandemic. This research tends to fall along four general lines, as identified in Romaniuc et al.'s (2021) recent review: (1) interventions experimentally manipulating beliefs about social norms surrounding vaccination (e.g., Alatas et al., 2019; Bokemper et al., 2021; Frank, 2020; Karing, 2019; Milkman, Patel, Gandhi, Graci, Gromet, Ho, Kay, Lee, Bogard, Brody, Chabris, Chang, Chapman, Dannals, Goldstein, Goren, Hershfield, Hirsch, Hmurovic, Horn, Karlan, Kristal, Lamberton, Meyer, Oakes, Schweitzer, Shermohammed, Talloen, Warren, Whillans, Yadav, Zlatev, Berman, Evans, Snider, Tsukayama, Van den Bulte, Volpp and Duckworth, 2021), (2) in-

terventions that make vaccines easier or perceived to be easier to obtain (e.g., Chapman et al., 2010; Giubilini et al., 2019), (3) interventions involving sending reminders to individuals to get vaccinated (e.g., Currat, Lazor-Blanchet and Zanetti, 2020; Milkman, Patel, Gandhi, Graci, Gromet, Ho, Kay, Lee, Bogard, Brody, Chabris, Chang, Chapman, Dannals, Goldstein, Goren, Hershfield, Hirsch, Hmurovic, Horn, Karlan, Kristal, Lamberton, Meyer, Oakes, Schweitzer, Shermohammed, Talloen, Warren, Whillans, Yadav, Zlatev, Berman, Evans, Snider, Tsukayama, Van den Bulte, Volpp and Duckworth, 2021; Milkman, Patel, Gandhi, Graci, Gromet, Ho, Kay, Lee, Akinola, Beshears et al., 2021; Yokum et al., 2018), and finally, (4) interventions communicating information about, as well as either the risks or benefits of, being vaccinated (e.g., Cucciniello et al., 2020; Lorini et al., 2020; Mowbray et al., 2016; Nyhan and Reifler, 2015; Olson, Berry and Kumar, 2020). Of these, interventions that convey a social norm of vaccination and those that require people to opt out of vaccination appointments rather than opt in tend to have been found to the greatest impact on vaccine uptake.

2.1 The Messenger (May) Matter

Studies that involve some sort of public messaging about vaccination have varied *whom* the message to get vaccinated is coming from. Generally speaking, people are more likely to seek out information from those with whom they share physical or cultural attributes (Ibarra, 1993, 1995; Spence, Lachlan and Griffin, 2007). This might be the case because ethnic or racial in-group members are likely to be evaluated as more trustworthy and caring—key components of source credibility (McCroskey and Teven, 1999)—than out-group members. A line often attributed to Teddy Roosevelt famously reads: “people don’t care about what you know, unless they know that you care.” When it comes to health concerns in specific, a 2013 experiment showed that Black Americans were more likely to believe Black sources than White ones when it came to updating their beliefs about medical information online (Spence

et al., 2013). An observational study about vaccine messaging found Black parents were more likely to evaluate other Blacks as credible sources of vaccine information in specific (Fu, Haimowitz and Thompson, 2019). Most recently, a survey experiment found that a race-concordant messengers were particularly effective at inducing covid-19 information seeking behaviors in Black participants (Alsan et al., 2021), although a follow-up study by the same authors that used a more complex design failed to replicate such effects (Torres et al., 2021). Looking at flu vaccines, Alsan and Eichmeyer (2021) find Black messengers are more effective at convincing Black American men to take a flu vaccine, though this result is observational in nature. These studies all involve one-way communications – when it comes to two-way communications, there is considerable evidence that race-concordance matters to Black Americans' health orientations (see, for example: Cooper-Patrick et al., 1999; Saha et al., 1999; Street et al., 2008).

Indeed, evidence that race-concordance matters is consistent with Dawson's (1994) seminal work demonstrating that race is a critical decision-making heuristic for Black Americans because individual outcomes are closely tied to the outcomes of the racial group. Ultimately, an endorsement from an individual who is also Black may signal that the action — in this case, taking a vaccine — will serve group members' interests well. As a celebrity, Tyler Perry, upon whom we rely in this study, might be particularly well-positioned to persuade Black Americans to get vaccinated not only because he himself is Black, but also because his career has been almost exclusively focused on producing Black-oriented content, potentially signaling particular investment in the Black community. Racial cues may also particularly salient in the context of the covid-19 pandemic, where uncertainty is high, due to the novelty of the situation, and information low, due to the complex scientific nature of vaccines (Popkin, 1994).

The vaccine encouragement literature also suggests that when it comes to vaccines, people are more likely to listen to health care providers, such as doctors and nurses than people

outside of the medical profession (Argote et al., 2021; Larson, 2020; Olson, Berry and Kumar, 2020). This makes sense, considering source credibility is particularly important when it comes to complex and sensitive information (Atkin, 2001; Eagly and Chaiken, 1975). But in the context of covid-19, where mistrust in health care institutions, official sources, and the government has been rampant, people might be relatively more receptive to messaging from leaders outside of these domains (Savoia et al., 2021). And given relatively higher mistrust of health and government institutions among Black Americans (Langer Research Associates, 2020), we might expect this to especially be the case among that subgroup (see also: Bogart et al., 2021). One such category of leaders might be celebrities — indeed, a large literature in cultural evolution shows that people tend to be more receptive to information from prestigious sources, including celebrities (see: Brewer et al., 2017; Mesoudi, 2011). And celebrities are widely used to promote all kinds of other pro-social behaviors, like turning out to vote and recycling. That said, little work has empirically examined the impact of celebrity endorsement on vaccination in specific. A recent New York Times interview with Drs. Heidi Larson and Simon Piatek of the Vaccine Confidence Project reports: "'We know that very scientific messages alone do not work and do not build trust,' Dr. Larson said. And it matters who delivers it: think Kim Kardashian (in the United States) or a Bollywood star (in India), not a white-coat doctor from the WHO or your federal government, Dr. Piatek said" (Anderson, 2020). That said, the only experimental assessment of the effect of celebrity endorsement of vaccination of which I am aware was conducted by Alatas et al. (2019), who randomly manipulated whether tweets encouraging people to get vaccinated in Indonesia came from a celebrity or a non-celebrity, and evaluated subsequent vaccination attitudes and behaviors. The results found celebrity encouragement to have a positive, though modest, effect on opinions about vaccines and vaccination behaviors.

2.2 Celebrity Messaging in Low Trust Environments

Why might celebrities be particularly persuasive communicators of vaccine information when it comes to Black Americans, though? Indeed, why would someone like Tyler Perry, who presumably has a lot on his Hollywood plate, create an hour-long special titled "Covid-19 Vaccine and the Black Community" and broadcast it specifically on BET?[1] Since vaccines started being widely touted as the solution to the covid-19 pandemic, polls and academic studies showed that Black Americans were significantly less likely than other racial groups to say they would get a vaccine (e.g., Gramlich and Funk, 2020; Malik et al., 2020; Funk and Tyson, 2021; Walsh, 2020), and, as vaccines became widely available, indeed were vaccinated at lower rates (Ndugga et al., 2021). Correlational studies evaluating the antecedents of the racial gap in intention to vaccinate found higher levels of mistrust in government and medical institutions among Black Americans, as well as experiences of racial discrimination and strong Black identity, to be significant predictors of unwillingness to get a covid-19 vaccine (Langer Research Associates, 2020; Ndugga et al., 2021). Such mistrust is likely to have arisen as a rational and self-protective response to an extensive history of general and health-related mistreatment of Black Americans at the hands of the U.S. government and medical establishment. Long before the covid-19 pandemic, qualitative work showed that the Tuskegee Syphilis Study almost always came up when Black Americans expressed vaccine hesitancy—the Study being one of the quintessential historical instances in which the government and medical establishment put Black lives in danger in pursuit of scientific discovery (Quinn et al., 2016). Moreover, Whetten et al. (2006) show that Black Americans are less likely than Whites to ask questions during medical examinations and are more likely to believe that physicians do not prioritize their well-being. As such, they show that Blacks are also more likely than Whites to seek medical information from sources other than doctors.

[1]BET stands for Black Entertainment Television and is an American cable television channel aimed at Black audiences.

If people are unlikely to trust government and medical sources, it makes sense for campaigners to turn to vaccine spokespeople outside of those communities. A focus-group-based study conducted during the covid-19 pandemic reported that Black participants indicated that they "trusted nonclinical organizations and influential formal and informal leaders to promote the vaccine." and that "some participants suggested that partnerships with Black celebrities (e.g., hip-hop artists) would encourage vaccination" (Bogart et al., 2021, p.8). Because of their visibility and social clout, celebrities might also have significant power over conveying a social norm of vaccination and inducing vaccine uptake in that way. Research indeed shows that perceiving vaccination to be a social norm in one's community is a significant predictor of the likelihood of getting vaccinated (Romaniuc et al., 2021). In an environment where trust in government and the medical community is low, outsiders like celebrities might be more effective in conveying pro-vaccination messages. Since much of the actual Black-targeted celebrity-based campaigning for covid-19 vaccines in the U.S. has used Black celebrities, we should also ask whether it matters that the celebrities are in fact Black. Taken together, the combined literature suggests we should expect a positive, though quite small, effect of celebrity and co-ethnic interventions on vaccination intentions and attitudes among Black Americans in the context of the covid-19 pandemic.

3 Study Design

3.1 Survey and Experimental Manipulation

Informed by these literatures, this study aims to provide a first step to assessing the extent to which celebrity and co-ethnicity and the combination of the two can affect Black Americans' vaccination orientations. This section lays out how the evaluation is conducted. A sample of 2000 Black-identifying Americans were recruited through the Lucid marketplace to participate in an online survey in April of 2021 — when covid-19 vaccines had already

become widely available in the United States (though varyingly across states) — and about 64% of Americans had received one dose of the vaccine.[2] Prior to the experimental intervention, respondents to the survey were asked a series of questions measuring several known determinants of covid-19 vaccine hesitancy, such as mistrust in government, experiences of racial discrimination, news consumption, and orientations toward alternative medicine. Subsequently, they were asked if they had already received a covid-19 vaccine and if so, routed past the experiment on to the outcome questions. About 50% of the sample had already received one dose of the vaccine—higher than the national average among Black Americans.

In the first condition, participants saw a 6-minute-long video in which Tyler Perry speaks to Gayle King (also a Black woman) on the nationally broadcast CBS This Morning about his experience taking the covid-19 vaccine while also encouraging others, especially Black Americans, to do the same. At one point in the video, Perry is also seen receiving the vaccine — an excerpt from his hour-long BET special in which he discusses with doctors and promotes covid-19 vaccination. A link to watch the video can be found in the chapter Appendix. Tyler Perry is both a highly visible celebrity (indeed, among TIME Magazine's 100 Most Influential People in 2020) and one whose career largely consists of producing film and television content targeted toward Black audiences. To test whether a co-ethnic celebrity might have a greater effect on respondents' orientations toward covid-19 vaccines than a co-ethnic non-celebrity, a second experimental condition randomly assigns individuals to watch a slightly shorter (4 minute) video of Sandra Lindsay, the Black New York City nurse who was the first person in the United States to receive a covid-19 vaccine, on December 14th, 2020. Like in the Perry condition, the video is of a broadcast on daily national television where Lindsay talks to an interviewer about receiving the covid-19 vaccine, discusses its benefits, and encourages others to take it too. The third prong of source credibility is competence (McCroskey and Teven, 1999). Since people generally tend to trust healthcare professionals

[2]This study received ethical approval through the IRB of Columbia University (Protocol #AAAT6467)

and we can expect a nurse to be considered a competent source, the fact that Lindsay is a nurse may circumvent Black Americans' reluctance to trust medical institutions about vaccines. Finally, those assigned to the control condition see an unrelated 5-minute video where Tyler Perry talks to Gayle King about his new movie studio. In addition to having been broadcast on national television, all of these videos have also been freely available to the public YouTube. After Alsan et al.'s (2021) work, this study is to my knowledge only the second experimental study evaluating the persuasive effects of Black messengers on Black Americans' vaccine orientations, and the first on covid-19 vaccine orientations.

3.2 Bundling and Defining the Treatment

The treatment videos obviously constitute "bundled" or compound treatments. Trying to evaluate in a controlled experiment the effects of real-world events is oftentimes complicated by bundling and thus introduces a trade-off between external and internal validity. In this case, it is hard to imagine unbundling the treatments, even in principle. In the context of the present study, we would need a version of the Perry treatment video in which he is a celebrity, and another in which he is not. Arguably, we could independently randomly assign the video to samples who are aware of Perry's celebrity status and to those who are not, but finding such groups in sufficiently large sizes would be an incredibly difficult endeavor, and no causal conclusions could be drawn from comparisons between groups anyways. We would perhaps get closer to unbundling by designing a video that is identical to the treatment video except in that the person speaking is not Perry, but rather a layperson with otherwise similar features, but this would require extensive video production and editing, and would then land us in the realm of questioning external validity; such a control video would have never existed "in the wild" and we could question the plausibility of Gayle King interviewing a completely unknown layperson about covid-19 vaccines.

Here, I am primarily interested in the effects of appeals that were actually made and

distributed on a mass basis during the covid-19 pandemic. To account for the bundling that often occurs when real appeals are used treatments, we need to be careful about how we define the treatments. Here, treatments are simply defined as 'having seen the Tyler Perry covid-19 vaccine clip' or 'having seen the Sandra Lindsay covid-19 vaccine clip.' The choice of these videos was motivated by a broad theoretical literature, but the aforementioned design constraints and validity trade-offs mean we cannot test specific hypotheses related to celebrity (or co-ethnicity/race-concordance) specifically. That said, it seems more theoretically plausible that any effects of the videos are attributable to celebrity or co-ethnicity rather than other elements such as Perry or Lindsay's appearance, the interviewers, or circumstantial factors present in the videos. But it is important to keep in mind that we cannot be sure.

A final note should also be made to address the fact that the videos are of slightly different lengths. One might wonder whether longer or shorter exposures to the media stimuli might constitute yet another stick in the bundled treatment, and in some sense, they probably do, but length differences are not vast. And again, cutting the longer Perry video to be the length of the shorter Lindsay video would also mean presenting it in a version that was not actually the version seen by people in the real world. Given the nature of the study and the motivation to retain as much external validity as possible, clips were kept at their original length.[3]

3.3 Outcomes and Estimation

The different questions that form the outcome measures of this study all tap different dimensions of attitudes toward covid-19 vaccination in order to get a sense of what aspects of vaccination orientations might be affected by treatment.[4] The first and most basic one asks

[3]The Perry treatment video does include further discussion between Perry, King, and a third interviewer after the portion presented in the experiment, but it is entirely unrelated to covid-19 and focuses on Perry's film career.

[4]The full questionnaire can be found in the chapter Appendix.

how likely respondents would be to get a covid-19 vaccine; those who indicated anything but 'certainly would not get it' were also asked how long they would wait to get the vaccine when it became possible for them to get it. Subsequent questions ask how likely people thought they would be to experience side effects, how likely they were to recommend to others to get a vaccine, how much they trusted that covid-19 vaccines were safe, how effective they thought vaccines were, and how confident people were that the vaccines were tested for safety in Black people specifically.

The primary estimands of interest are the Average Treatment Effects (ATEs) of having been in one of the two treatment conditions (vaccine videos) compared to the control condition (the unrelated Tyler Perry video) on the subsequent outcome questions, as well as the difference in ATEs between having been in the Tyler Perry condition versus the Sandra Lindsay condition. Relying on a Neyman-Rubin (Rubin, 1974) causal model, the estimation procedure can more precisely be defined as follows. Each respondent i has the potential outcome $Y_i(Z = 0)$ when untreated, $Y_i(Z = 1)$ when assigned to the Tyler Perry vaccine video and $Y_i(Z = 2)$ when assigned to the Sandra Lindsay video. The effect of each treatment on a given respondents' outcome is defined as the difference in potential outcomes between the treatment condition $\mathbf{Z}$ and the control condition: $Y_i(\mathbf{Z}) - Y_i(Z = 0)$.[5] I am interested in the pairwise comparisons between $E[Y(\mathbf{Z}) - Y(0)]$. To generate estimates of the Average Treatment Effect (ATE), we can regress the outcomes on the treatment indicators for each video on using OLS regression since outcomes are continuous. The general model can be written as follows:

$$Y_i = \beta 0 + \beta Z1 + \beta Z2 + e_i,$$

Where $Z1$ corresponds to the treatment indicator for the Perry vaccine video and $Z2$ corre-

[5] Since respondents are not able to skip the video page to which they are assigned nor navigate away from the page until the video ends, treatment assignment is assumed to be equal to treatment received.

sponds to the treatment indicator for the Lindsay treatment indicator.

To the extent that non-zero treatment effects are observed, we may, in an exploratory manner, assess whether these treatment effects are heterogeneous conditional on pre-treatment covariates of interest, such as trust in government. In this case, $Z1$ and $Z2$ indicators, respectively, would be interacted with X_k covariates, where k indexes the covariate. All analyses were pre-registered through the EGAP Registry while study data was still being collected. Given the difficulty of changing medical attitudes, and the mixed findings of the literature on vaccine persuasion, I expect that both the Perry and Lindsay videos will have weakly positive effects of respondents' covid vaccination intentions and vaccine attitudes, but that the effect of the Perry video will be greater in magnitude. Similarly, when it comes to heterogeneous treatment effects, I expect those who are least trustful of the government and most inclined to rely on alternative medicines to show the least movement on outcomes compared to the relatively trustful and those who do not use alternative medicines. I am agnostic about the expected direction of effects when it comes to indicators of Black identity. On the one hand, we might expect those with a strong racial identity to be more receptive to race-concordant messengers and especially those like Perry who exist outside of the medical establishment but bear cultural cache. On the other hand, we might expect the same people to be more deeply distrustful of vaccine messaging in general, and to in fact be resentful of the use of race-concordance to persuade. Before discussing the results from the estimating equations outlined here, I first offer a descriptive picture of the data along sociodemographic and relevant attitudinal dimensions.

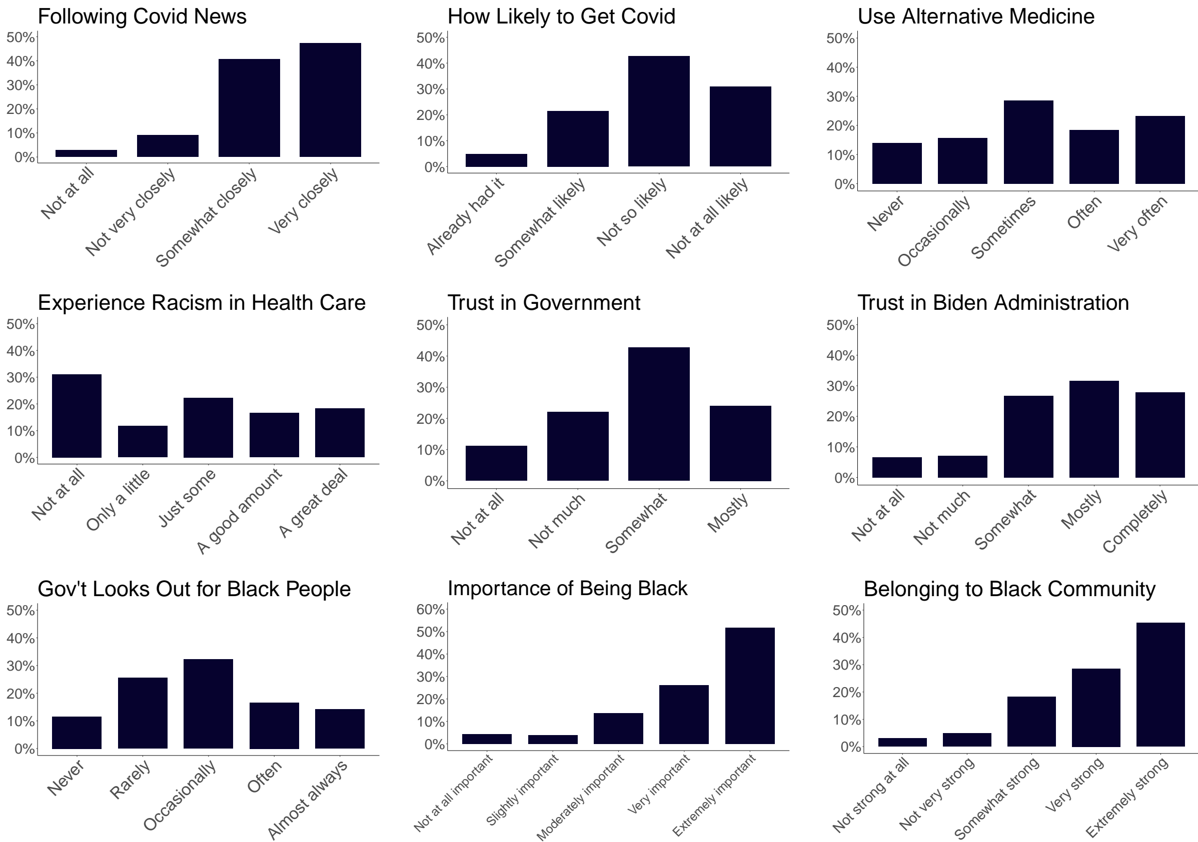
4 Results

4.1 Black Public Opinion and Covid-19

Given the recency of the pandemic and the fact that evidence about the nature of Black Americans' attitudes toward vaccines is thus still scarce, it is useful to get a sense of opinions among the sample about covid-19 in general, as well as opinions about factors that the literature suggests might be related to vaccination, such as trust in government, attitudes toward medicine, and experiences of racism and Black identity. Figure 1 presents graphs illustrating frequencies across responses to questions capturing these attitudes. First, and unsurprisingly, considering the nature of the pandemic and expansive press coverage, the vast majority of people were following news about the covid-19 pandemic closely or very closely. That said, about 70% of respondents did not think it was likely they would actually get covid-19, and only 10% had already had it (roughly in line with the American population proportion). In terms of managing illnesses such as covid-19, nearly half the sample reported relying on alternative medicines, instead of conventional interventions like vaccines, often or very often.This comports with reports showing the prevalence of the belief especially among the Black community that covid can be prevented through the ingestion of certain supplements; indeed, several prominent Black dietitians and alternative medicine practitioners have publicly made this claim (for a summary, see: Collins-Dexter, 2020). Reliance on alternatives to conventional medicine might be especially high among Blacks because of historical mistreatment by mainstream medical establishments and a continued lack of diversity in medical practice. When asked how much racism has impeded their ability to receive good quality health care, however, respondents were divided: while a little over a third said it did a great deal or a good amount, a third also said it did not at all, leaving the rest in the middle.

When it comes to trusting non-medical institutions, namely in this case, the government,

Figure 1: Descriptive Statistics



respondents are similarly divided. Only slightly over a third trust say the federal government can be trusted completely or at least mostly. Evaluations of the Biden administration are significantly rosier, though: about 60% are completely or mostly trusting. That said, that figure stands in contrast with the observation that only 30% of respondents believe the government can be trusted to look out for the interests of Blacks specifically. Nevertheless, we can deduce a meaningful amount of distrust in what the government is up to in general. If such distrust has a racial component to it, then it is important to understand the extent to which respondents indeed feel attached to a racial identity and community. Here, 78% of respondents report feeling that being Black is very or extremely important to their identity and a similar proportion report very or extremely strong feelings of belonging to the Black community. In summary, then, distrust of the institutions managing the pandemic is fairly high, though it depends what we are considering in specific. Respondents also report strong racial identities and strong feelings of belonging to the racial community. With this background in mind, we can move into assessing how appeals from a Black celebrity — an elite member of the racial in-group divorced from governmental or medical establishments — might affect the probability of getting vaccinated against covid-19 and attitudes toward the vaccine. Overall, these descriptive results comport with those found in other studies, the most expansive of which thus far has been that conducted by Langer Research Associates (2020).

4.2 Does Race-Concordant Celebrity Messaging Work?

The graphs under Figure 2 present the effects of the experimental interventions on each of the seven different outcome measures. In each, the navy-blue (dark) points correspond to the estimated ATE of having seen the Sandra Lindsay video versus the control video; the turquoise (lightest) points to the ATE of having seen the Tyler Perry video versus the control video; and the red (medium colored) points to the ATE of having seen the Tyler Perry video

versus the Sandra Lindsay video. The points are the OLS coefficients on the treatment (relative to the comparison category) from the regression of the relevant outcome on the binary treatment indicator, and bars represent 95% confidence intervals. Each outcome variable is recoded to range between 0 and 1. We begin with the most basic question of interest: does encouragement from a celebrity or other Black community leader to get vaccinated against covid-19 actually affect Black Americans intentions to take a vaccine?

The top left panel of Figure 2 presents the ATEs corresponding to this question. Evidently, none of the estimated potential group comparisons reveal treatment effects statistically distinguishable from zero; that is to say, exposure to the Perry video has no effect on vaccination intentions compared to control nor compared to exposure to the Lindsay video, and neither does exposure to the Lindsay video compared to control.[6] Nor is there any observable effect for either the Perry or Lindsay treatments on how long a person would wait to take a vaccine if they were at all open to it (second panel), or on their perceptions of whether vaccines have harmful side effects (third panel). If either the Perry or Lindsay treatments proved to have any effect on the likelihood that a person would encourage *others* to take the vaccine, then we could at least say there might be some "downstream" effect of such messaging, one that creates the perception of a vaccination norm; this might be the case insofar as those who viewed the Perry video were nearly statistically significantly more likely to encourage others to take the vaccine than those who saw the Lindsay video. While celebrity messaging might not make people more likely to get the vaccine themselves, they

[6]The survey measured post-treatment whether respondents had seen, prior to the survey, the video to which they were exposed — 28% of respondents in the Perry treatment group reported having already seen the Perry video, whereas 50% of respondents in the Lindsay group reported already having seen that video. Though the latter proportion is high, and suggests the information may be less impactful because it is less novel, it is hard to avoid given how publicized Lindsay's vaccination was. People may also be answering the question based on having seen the still image from the video, which was even more widely publicized. It might also of course be the case that people assigned to one video condition may have at some point before the study also encountered the other video, raising a question about compliance. Unfortunately, because the question about whether respondents had seen the video to which they were assigned is asked post-treatment, treatment effect estimates cannot be conditioned on prior exposure.

might spread a positive word to others about getting the vaccine, potentially nevertheless contributing to a perceived "norm" of vaccination. Again, though, this result is only significant at the $p = 0.06$ level and thus must be taken with the requisite grain of salt. The final three outcomes evaluated are those related to the safety and efficacy of covid-19 vaccines and treatment effects are displayed in the last three panels of Figure 2. Evidently, neither the Perry nor the Lindsay treatments had a statistically distinguishable effect on respondents' perceptions of how safe the vaccine is, how effective it is in actually preventing covid-19, and how safe it is specifically for Black Americans. It is important to keep in mind, however, that muted effects may in part be the product of prior exposure – either to one's own treatment condition or to the other condition or to both. Working with real-world messages is important for purposes of external validity, but in a non-field experimental setting can face this constraint. Moreover, although both treatment videos contain pro-vaccine information, respondents are likely to have heard counterarguments to this information in their regular lives prior to the study, further limiting any potential effects.

That said, when all of the seven outcome questions are scaled together to create an additive index measuring vaccine orientations more generally, exposure to the Perry treatment (compared to the Lindsay treatment) leads to a nearly statistically significant 5 percentage point increase in how positively disposed respondents are to covid vaccines ($\beta = 0.05$; $p = 0.06$). With a larger sample of people who have yet to receive a vaccine, we might expect this effect to indeed be significant at $p = 0.05$. Even if race-concordant celebrities are not necessarily effective at changing people's minds about specific elements of vaccination, they may have a slight effect on generating more positive dispositions toward it in a broader sense. The availability of the series of sociodemographic and attitudinal pre-treatment questions allow us to explore how any such effects may vary among subgroups of individuals. In total, respondents were asked 13 such questions and an indicator for each question can be interacted with each of the two treatments. An F-test reveals that, compared to the basic

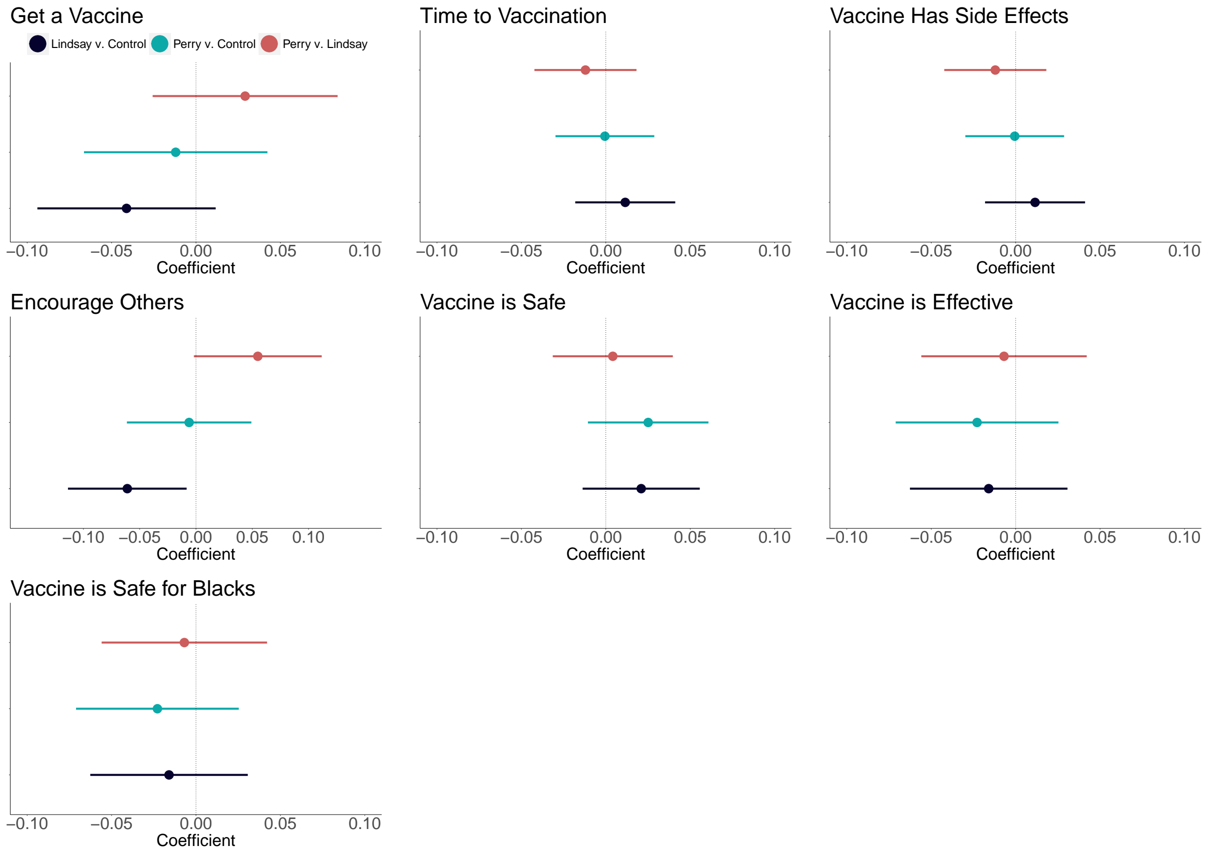


Figure 2: Average Treatment Effects by Outcome

model regressing outcome on treatment, the coefficients in the fully interacted model are not jointly zero (Perry vs. control: $F = 1.71$, $p < 0.004$; Lindsay vs. control: $F = 1.73$, $p = 0.004$).

4.3 Heterogeneity and Backfire Effects

To explore potential treatment effect heterogeneity, I estimate 11 models of the scaled outcome on each of the two treatment indicators (Perry vs. control, Lindsay vs. control) where each model interacts the treatment indicator with a pre-treatment covariate (the three separate measures of trust in government are scaled together for efficiency, hence 11 and not 13). Though there is no meaningful heterogeneity in effects among respondents of different ages, party identifications, genders, education levels, regions, some interesting effects emerge when it comes to Black identity and sense of belonging to the Black community. While the coefficients in the regression interacting identity and belonging with the Perry treatment (vs. control) are generally not statistically significant, and variance is limited by very small sample sizes, they nevertheless all trend in the same direction: exposure to the Perry treatment seems to be associated with a *decrease* in the likelihood of taking a vaccine for those who report Black identity as being at least slightly important to them or belonging to the Black community as being at least slightly important (Table 1).[7] With larger samples, one might expect to uncover significant effects here given the consistency of the trend. These tendencies imply there may be a kind of racial identity-based backfiring to Black celebrity messaging. Indeed, if we look at the comments under the version of the video posted on YouTube, the idea of Perry being a 'traitor' to the racial group and a puppet of a White establishment appears extremely frequently. On the whole, then, while there may be a weakly positive effect of race-concordant celebrity messaging on vaccine orientations overall, there is a subset of Black Americans for whom this messaging may elicit an undesired response.

[7]And similar effects are found if we compare those who received the Perry treatment to those who received the Lindsay treatment, instead of to control.

The idea of backfire effects of different kinds of vaccine messaging has long been documented (see, for example: Nyhan et al., 2014) and is thus not entirely surprising. The next section summarizes the general findings and discusses their implications for the role of celebrities and media in mass vaccination campaigns.

5 Conclusion

Governments and medical establishments in the United States have long relied on celebrities to encourage people to get vaccinated against diseases. In the 1950s, Elvis Presley famously took the polio vaccine backstage on the Ed Sullivan Show, and in the 1980s, author Roald Dahl publicly campaigned for the measles vaccine. Since vaccines against covid-19 became available in the winter of 2020, dozens of celebrities took the vaccine publicly on TV or social media to promote the mass vaccination required to end the pandemic; New York Magazine even began keeping an updated list of every celebrity that had been vaccinated. Likewise, many events featuring celebrities were put on to promote vaccination; hip hop artist Ciara, for instance, played host to a star-studded NBC vaccination special featuring guests like Barack and Michelle Obama to encourage people to take a covid-19 vaccine. Though a lot of time, money, and certainly effort has been invested in celebrity-driven campaigns, it is unclear whether they achieve their intended goals. Extant literature suggests people might turn to celebrities or other community leaders, instead of medical professionals or government figures, when trust in the latter two is low. This study evaluates whether televised celebrity encouragements might be effective in the particular case of vaccinating Black Americans against covid-19. In the U.S., Black Americans tend to be more distrustful of the government and medical establishment than Whites (Spence et al., 2013), and research conducted in the last year has demonstrated this to also be the case with respect to covid-19 in specific (Ndugga et al., 2021; Langer Research Associates, 2020). Although celebrities may lack the credibility of government or the expertise of doctors, might they nevertheless

Table 1: CATEs by Racial Identity and Belonging

	Positivity Toward Covid Vaccines	
Perry Video	0.219	0.132
	(0.146)	(0.135)
Being Black is extremely important	0.073	
	(0.083)	
Being Black is very important	0.058	
	(0.085)	
Being Black is moderately important	0.071	
	(0.090)	
Being Black is slightly important	0.061	
	(0.100)	
Perry video × extremely important	−0.211	
	(0.150)	
Perry video × very important	−0.144	
	(0.152)	
Perry video × moderately important	−0.247	
	(0.157)	
Perry video × slightly important	−0.354*	
	(0.184)	
Belonging is extremely strong		0.243***
		(0.089)
Belonging is not very strong		0.122
		(0.106)
Belonging is somewhat strong		0.200**
		(0.091)
Belonging is very strong		0.184**
		(0.091)
Perry video × extremely strong		−0.180
		(0.140)
Perry video × very strong		−0.042
		(0.142)
Perry video × not very strong		−0.016
		(0.191)
Perry video × somewhat strong		−0.121
		(0.143)
Constant	0.410***	0.274***
	(0.080)	(0.085)
Observations	321	322
R^2	0.029	0.048
Adjusted R^2	0.0004	0.020

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

act as effective messengers when the former are not trusted?

To take a step in the direction of answering this question, this study gauges whether a televised instance of a Black celebrity taking and discussing the covid-19 vaccine generates a shift in the probability that an individual Black person is willing to take the vaccine, as well as changes in evaluations of the safety and effectiveness of such vaccines. Respondents to an online survey were randomly assigned to one of three experimental conditions. In the first treatment condition, and the principle one of interest, respondents were exposed to a video in which film and television personality Tyler Perry receives and discusses the vaccine on broadcast television. The 5-minute video is in fact a "summary" clip of an hour-long special Perry hosted on BET earlier in the year to encourage Black Americans to take a covid-19 vaccine. In the second condition, respondents were assigned to watch a video of New York City nurse Sandra Lindsay, who is also Black, and was the first person to receive a covid-19 vaccine in the U.S., taking it and discussing it. The third condition was a control video condition. Contrary, perhaps, to the intuitions of the celebrities and producers of these mediatized vaccination campaigns, the results show that co-ethnic celebrity endorsement does little to move the attitudes of Black Americans when it comes to intentions to get vaccinated, evaluations of covid-19 vaccines, or probability of recommending vaccination to others. Nor is there any effect of exposure to a non-celebrity co-ethnic endorsement (ie. the Sandra Lindsay condition). A weakly positive effect is detected for race-concordant celebrity messaging when all vaccine-orientation outcomes are taken together, though, suggesting that although such messaging might not be effective for specific attitudes toward vaccines, they may nevertheless change people's more general or abstract orientations toward vaccination. These findings are in line with work that suggests that celebrities are unlikely to have much impact on other public causes like voter turnout (see, for example: Green and Gerber, 2015), but support literature that suggests that in cases of low trust, non-establishment actors might be more persuasive than the usual institutional suspects. They also contrast with

Alatas et al.'s (2019) work demonstrating the effectiveness of celebrity Twitter messaging on vaccination in Indonesia – the only experimental work of which I am aware that tests the effects of celebrity campaigns on vaccination orientations. Although celebrities, and especially co-ethnic ones, might be seen as more trustworthy or caring than the government or members of the medical establishment, they might not be seen as more credible, and in a health context, credibility is what might matter most. Extant literature also suggests that when it comes to vaccines, people are more likely to trust friends, family, and members of their immediate community over more distal figures (Romaniuc et al., 2021) – this may also explain the null and weak findings reported here. The notion in extent literature that vaccine messaging can engender backfire effects (Nyhan et al., 2014) receives some support from this study as well, given that strong feelings of Black identity seem to actually be associated with more negative orientations toward covid-19 vaccines after exposure to celebrity messaging. In any case, these findings provide an initial sense that while public vaccination campaign resources may not entirely be wasted on celebrity advertising, they might be better allocated to other avenues. Given that government bodies are increasingly, as of August 2021, pushing social media influencers and other celebrities to advocate for covid-19 vaccines, further work randomizing other features of messengers is certainly warranted.

As always, there are several limitations to the study. First, and perhaps most importantly, the video treatments used here are necessarily "bundled" and we cannot attribute causal effects to the effect of "celebrity" or co-ethnicity per se. Given the incredible paucity of research on the effects of celebrity promotion on public health in general, and vaccine uptake in specific, though, assessing the effects of these particular campaigns nevertheless represents a step toward understanding the extent to which such campaigns matter. Ideally, these results will one day exist among others and together paint a more conclusive picture. The study also does not compare how people might respond to a White, or other non-Black, celebrity. It might be the case that while race concordance matters for trust, histories of

White supremacy in the United States still make it so that everyone, regardless of race, perceives White people as more credible than non-White people. Here I aimed to focus, however, on the effects of Black celebrity among Black Americans – a still relatively understudied population in most survey efforts. It may also be the case that there the treatments presented here had some statistically significant effects that were too small for the study to be sufficiently powered to detect. Given the dearth of relevant extant evidence, power analyses conducted prior to study implementation assumed effects around which there was great uncertainty. Finally, it is important to keep in mind that reported intentions to vaccinate, and reported attitudes toward vaccines, do not always comport with actual vaccination behaviors. People might be responding expressively to survey questions, for whatever reason, and getting vaccinated when they say they will not (or vice versa). It is also necessary, from a normative standpoint, to remind the reader that vaccine hesitancy is not the only driver of vaccination rates and that access to vaccines (e.g., ability to take time off work, reach sites, make appointments etc.) likely matters a great deal as well. Ultimately, future work should take up these issues and work toward constructing a larger body of evidence of how celebrity messaging might affect public health interventions. Such a research project is especially pressing in an era of social media and influencers, where even what we might consider "micro-celebrities" have outsized platforms from which to speak, and especially in contexts where trust in government is low. Evidently, government and medical institutions believe celebrity interventions to be important vectors in advancing public health — as such, it is imperative to understand whether they indeed are or whether resources might better be allocated elsewhere.

References

Adorno, Theodor W., Else Frenkel-Brunswick, Daniel J. Levinson and R. Nevitt Sanford. 1950. *The Authoritarian Personality.* New York: Harper and Row.

Alatas, Vivi, Arun G. Chandrasekhar, Markus Mobius, Benjamin A. Olken and Cindy Paladines. 2019. "When Celebrities Speak: A Nationwide Twitter Experiment Promoting Vaccination In Indonesia." *NBER Working Paper No. 25589* .

Alsan, Marcella, Fatima Cody Stanford, Abhijit Banerjee, Emily Breza, Arun G. Chandrasekhar, Sarah Eichmeyer, Paul Goldsmith-Pinkham, Lucy Ogbu-Nwobodo, Benjamin A. Olken, Carlos Torres, Anirudh Sankar, Pierre-Luc Vautrey and Esther Duflo. 2021. "Comparison of Knowledge an Information-Seeking Behavior After General Covid-19 Public Health Messages and Messages Tailored for Black and Latinx Communities." *Annals of Internal Medicine* 4:484–492.

Alsan, Marcella and Sarah Eichmeyer. 2021. "Experimental Evidence on the Effectiveness of Non-Experts for Improving Vaccine Demand." *NBER Working Paper No. 28593* .

Altemeyer, Bob. 1998. The Other "Authoritarian Personality". In *Advances in Experimental Social Psychology*, ed. M.P. Zanna. Vol. 30 San Diego: Academic Press pp. 47–92.

Anderson, Jenny. 2020. "She Hunts Viral Rumors About Real Viruses." *The New York Times, Accessed July 13 2021* .

Argote, Pablo, Elena Barham, Sarah Daly, Julian E Gerez, John Marshall and Oscar Pocasangre. 2021. The Shot, the Message, and the Messenger: COVID-19 Vaccine Acceptance in Latin America. Technical report Working Paper.

Atkin, Charles. 2001. "Impact of Public Service Advertising: Research Evidence and Effective Strategies." *Kaiser Family Foundation* .

Bilodeau, Antoine. 2010. "The Development of Dual Loyalties: Immigrants' Integration to Canadian Regional Dynamics." *Canadian Journal of Behavioural Science* 43(3):515–544.

Bilodeau, Antoine, Luc Turgeon and Ekrem Karakoc. 2012. "Small Worlds of Diversity: Views toward Immigration and Racial Minorities in Canadian Provinces." *Canadian Journal of Political Science* 45(3):579–805.

Bilodeau, Antoine and Mebs Kanji. 2010. "The New Immigrant Voter, 1965-2004: The Emergence of a New Liberal Partisan?" *Voting Behaviour in Canada* pp. 65–85.

Bobo, Lawrence and Vincent L. Hutchings. 1996. "Perceptions of Racial Group Competition: Extending Blumer's Theory of Group Position to a Multiracial Social Context." *American Sociological Review* 61(6):951–972.

Bogart, Laura M., Lu Dong, Priya Gandhi, Samantha Ryan, Terry L. Smith, David J. Klein, Luckie Alexander Fuller and Bisola O. Ojikutu. 2021. "What Contriutes to Covid-19 Vaccine Hesitancy in Black Communities, and How Can it Be Addressed?" *RAND Corporation Report* .

Bokemper, Scott E, Gregory A Huber, Alan S Gerber, Erin K James and Saad B Omer. 2021. "Timing of COVID-19 Vaccine Approval and Endorsement by Public Figures." *Vaccine* 39(5):825–829.

Boroditsky, Lera. 2001. "Does Language Shape Thought?: Mandarin and English Speakers' Conceptions of Time." *Cognitive Psychology* 43(1):1–22.

Boroditsky, Lera and Alice Gaby. 2010. "Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community." *Psychological Science* 21(11):1635–1639.

Boroditsky, Lera, Lauren A Schmidt and Webb Phillips. 2003. Sex, Syntax, and Semantics. In *Language in Mind: Advances in the Study of Language and Cognition*, ed. Dedre Gentner and Susan Goldin-Meadow. Boston: MIT Press pp. 61–79.

Brandt, Mark J, Christine Reyna, John R Chambers, Jarret T Crawford and Geoffrey Wetherell. 2014. "The ideological-Conflict Hypothesis: Intolerance Among Both Liberals and Conservatives." *Current Directions in Psychological Science* 23(1):27–34.

Brewer, Noel T, Gretchen B Chapman, Alexander J Rothman, Julie Leask and Allison Kempe. 2017. "Increasing Vaccination: Putting Psychological Science into Action." *Psychological Science in the Public Interest* 18(3):149–207.

Budiman, Abby. 2020. "Key Findings About U.S. Immigrants." *Pew Research Center* .

Canada. 2012. *Linguistic Characteristics of Canadians.* Ottawa: Statistics Canada.

CDC. 2021. "Increasing COVID-19 Vaccine Uptake among Members of Racial and Ethnic Minority Communities: A Guide for Developing, Implementing, and Monitoring Community-Driven Strategies." *Report* .

Chambers, John R, Barry R Schlenker and Brian Collisson. 2013. "Ideology and Prejudice: The Role of Value Conflicts." *Psychological science* 24(2):140–149.

Chapman, Gretchen B, Meng Li, Helen Colby and Haewon Yoon. 2010. "Opting in vs Opting Out of Influenza Vaccination." *JAMA* 304(1):43–44.

Chong, Dennis and James N. Druckman. 2007. "Framing Theory." *Annual Review of Political Science* 10:103–126.

Citrin, Jack. 1990. "Language Politics and American Identity." *Public Interest* 99:96–109.

Citrin, Jack, Donald P. Green, Christopher Muste and Cara Wong. 1997. "Public Opinion Toward Immigration Reform: The Role of Economic Motivations." *The Journal of Politics* 59(03):858–881.

Clayton, Katherine, Jeremy Ferwerda and Yusaku Horiuchi. 2019. "Exposure to Immigration and Admission Preferences: Evidence from France." *Political Behavior* pp. 1–26.

Collins-Dexter, Brandi. 2020. "Canaries in The Coal Mine: Covid-19 Misinformation and Black Communities (Report)." *Shoreinstein Center on Media, Politics and Public Policy, Harvard Kennedy School* pp. 1–21.

Cooper-Patrick, Lisa, Joseph J Gallo, Junius J Gonzales, Hong Thi Vu, Neil R Powe, Christine Nelson and Daniel E Ford. 1999. "Race, Gender, and Partnership in the Patient-Physician Relationship." *Jama* 282(6):583–589.

Crano, William D. 1997. "Vested Interest, Symbolic politics, and Attitude-Behavior Consistency." *Journal of Personality and Social Psychology* 72(3):485–491.

Crawford, Jarret T and Jane M Pilanski. 2014. "Political Intolerance, Right and Left." *Political Psychology* 35(6):841–851.

Cucciniello, Maria, Paolo Pin, Blanka Imre, Greg Porumbescu and Alessia Melegaro. 2020. "Altruism and Vaccination Intentions: Evidence from Behavioral Experiments." *medRxiv* .

Currat, Michael, Catherine Lazor-Blanchet and Giorgio Zanetti. 2020. "Promotion of the Influenza Vaccination to Hospital Staff During Pre-Employment Health Check: a Prospective, Randomised, Controlled Trial." *Journal of Occupational Medicine and Toxicology* 15(1):1–7.

Dancygier, Rafaela M and Michael J Donnelly. 2012. "Sectoral Economies, Economic Contexts, and Attitudes Toward Immigration." *The journal of politics* 75(1):17–35.

Danziger, Shai and Robert Ward. 2010. "Language Changes Implicit Associations Between Ethnic Groups and Evaluation in Bilinguals." *Psychological Science* 21(6):1–2.

Davidov, Eldad, Bart Meuleman, Jan Cieciuch, Peter Schmidt and Jaak Billiet. 2014. "Measurement Equivalence in Cross-National Research." *Annual Review of Sociology* 40:55–75.

Dawson, Michael C. 1994. *Behind the Mule: Race and Class in African-American Politics.* Princeton: Princeton University Press.

de la Garza, Rodolfo O., Louis DeSipio, F. Chris Garcia, John Garcia and Angelo Falcon. 1992. *Latino Voices: Mexican, Puerto Rican, and Cuban Perspectives on American Politics.* Boulder: Westview.

Druckman, James N. 2001. “The Implications of Framing Effects for Citizen Competence.” *Political Behavior* 23(3):225–256.

Dufresne, Yannick, Anja Kilibarda, André Blais and Alexis Bibeau-Gagnon. 2019. “Religiosity or Racism? The Bases of Opposition to Religious Accommodation in Quebec.” *Nations and Nationalism* 25(2):673–696.

Dufresne, Yannick, Sanjay Jeram and Alexandre Pelletier. 2014. “The True North Strong and Free Healthcare? Nationalism and Attitudes Towards Private Healthcare Options in Canada.” *Canadian Journal of Political Science* 47(03):569–595.

Dutwin, David and Trent D Buskirk. 2017. “Apples to Oranges or Gala Versus Golden Delicious? Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples.” *Public Opinion Quarterly* 81(S1):213–239.

Eagly, Alice H and Shelly Chaiken. 1975. “An Attribution Analysis of the Effect of Communicator Characteristics on Opinion Change: The Case of Communicator Attractiveness.” *Journal of personality and social psychology* 32(1):136–144.

Ellis, Nicquel Terry. 2020. “Nearly One Third of Black Americans Remain Hesitant to Get Covid-19 Vaccine, Study Finds.” *CNN* .

Erwin, Donn. 1971. *The Attraction Paradigm*. Academic Press.

Esses, Victoria M., Geoffrey Haddock and Mark P. Zanna. 1993. Values, Stereotypes, and Emotions as Determinants of Intergroup Attitudes. In *Affect, Cognition and Stereotyping*, ed. Diane M. Mackie and David L. Hamilton. New York: Academic Press pp. 137–166.

Fawcett, Christine A and Lori Markson. 2010. “Similarity Predicts Liking in 3-Year-Old Children.” *Journal of Experimental Child Psychology* 105(4):345–358.

Feather, Norman T. 1985. “Attitudes, Values, and Attributions: Explanations of Unemployment.” *Journal of Personality and Social Psychology* 48(4):876.

Figueroa, Lucila. 2018. “Cultural Norms and Immigrants in the United States: The Green Card Experiment.” *Journal of Race, Ethnicity and Politics* 3(2):293–330.

Ford, Robert. 2011. “Acceptable and Unacceptable Immigrants: How Opposition to Immigration in Britain is Affected by Migrants’ Region of Origin.” *Journal of Ethnic and Migration Studies* 37(7):1017–1037.

Fraga, Luis, John A. Garcia, Michael Jones-Correa, Gary M. Segura, Rodney Hero and Valerie Martinez-Ebers. 2010. *Latino Lives in America: Making it Home*. Philadelphia: Temple University Press.

Frank, Robert H. 2020. *Under the Influence: Putting Peer Pressure to Work*. Princeton: Princeton University Press.

Fu, Linda Y., Rachel Haimowitz and Danielle Thompson. 2019. "Community Members Trusted by African American Parents for Vaccine Advice." *Human Vaccines & Immunotherapeutics* 15(7-8):1715–1722.

Fuhrman, Orly, Kelly McCormick, Eva Chen, Heidi Jiang, Dingfang Shu, Shuaimei Mao and Lera Boroditsky. 2011. "How Linguistic and Cultural Forces Shape Conceptions of Time: English and Mandarin Time in 3D." *Cognitive Science* 35(7):1305–1328.

Funk, Cary and Alec Tyson. 2021. "Growing Share of Americans Say They Plan To Get a COVID-19 Vaccine – or Already Have." *Pew Research Center* .

Gaertner, Samuel L. and John F. Dovidio. 1986. The Aversive Form of Racism. In *Prejudice, Discrimination and Racism*, ed. John F. Dovidio and Samuel L. Gaertner. Orlando, FL: Academic Press pp. 61–89.

Gamson, William A. and Andre Modigliani. 1989. "Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach." *American Journal of Sociology* 95(1):1–37.

Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* New York: W.W. Norton.

Giasson, Thierry, Colette Brin and Marie-Michèle Sauvageau. 2010. "La Bon, la Brute et le Raciste: Analyse de la couverture médiatique de l'opinion publique pendant la 'crise' des accomodements raisonnables au Québec." *Canadian Journal of Political Science* 43(2):379–406.

Giubilini, Alberto, Lucius Caviola, Hannah Maslen, Thomas Douglas, Anne-Marie Nussberger, Nadira Faber, Samantha Vanderslott, Sarah Loving, Mark Harrison and Julian Savulescu. 2019. "Nudging Immunity: The Case for Vaccinating Children in School and Day Care by Default." *HEC Forum* 31(4):325–344.

Graham, Jesse, Jonathan Haidt and Brian A Nosek. 2009. "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology* 96(5):1029.

Gramlich, John and Cary Funk. 2020. "Black Americans Face Higher Covid-19 Risks, Are More Hesitant to Trust Medical Scientists, Get Vaccinated." *Pew Research Center* .

Green, Donald P. and Alan S. Gerber. 2015. *Get Out the Vote: How to Increase Voter Turnout.* 3rd ed. Washington, D.C.: Brookings Institution Press.

Hainmueller, Jens and Daniel J. Hopkins. 2014*a*. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes Toward Immigrants." *American Journal of Political Science* 59(3):529–548.

Hainmueller, Jens and Daniel J. Hopkins. 2014*b*. "Public Attitudes Toward Immigration." *Annual Review of Political Science* 17:225–249.

Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2013. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

Hainmueller, Jens and Dominik Hangartner. 2013. "Who Gets a Swiss Passport? A Natural Experiment in Immigration Discrimination." *American Political Science Review* 107(1):159–187.

Hainmueller, Jens, Dominik Hangartner and Teppei Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments Against Real-World Behavior." *Proceedings of the National Academy of Sciences* 112(8):2395–2400.

Hainmueller, Jens and Michael J Hiscox. 2010. "Attitudes Toward Highly Skilled and Low-Skilled Immigration: Evidence from a Survey Experiment." *American Political Science Review* 104(01):61–84.

Harrell, Allison, Stuart Soroka, Shanto Iyengar and Nicholas A. Valentino. 2012. "The Impact of Economic and Cultural Cues on Support for Immigration in Canada and the U.S." *Canadian Journal of Political Science* 45(3):499–530.

Helbling, Marc and Richard Traunmüller. 2018. "What is Islamophobia? Disentangling Citizens' Feelings Toward Ethnicity, Religion and Religiosity Using a Survey Experiment." *British Journal of Political Science* pp. 1–18.

Hellwig, Timothy and Abdulkader Sinno. 2017. "Different Groups, Different Threats: Public Attitudes Towards Immigrants." *Journal of Ethnic and Migration Studies* 43(3):339–358.

Hong, Yingyi, Michael W. Morris, Chi yue Chiu and Verónica Benet-Martinez. 2000. "Multicultural Minds: A Dynamic Constructivist Approach to Culture and Cognition." *American Psychologist* 55(7):709–720.

Ibarra, Herminia. 1993. "Network Centrality, Power, and Innovation Involvement: Determinants of Technical and Administrative Roles." *Academy of Management Journal* 36(3):471–501.

Ibarra, Herminia. 1995. "Race, Opportunity, and Diversity of Social Circles in Managerial Networks." *Academy of Management Journal* 38(3):673–703.

Inglehart, Ronald and Pippa Norris. 2002. *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge: Cambridge University Press.

Iyengar, Shanto, Simon Jackman, Solomon Messing, Nicholas Valentino, Toril Aalberg, Raymond Duch, Kyu S Hahn, Stuart Soroka, Allison Harell and Tetsuro Kobayashi. 2013. "Do Attitudes About Immigration Predict Willingness to Admit Individual Immigrants? A Cross-National Test of the Person-Positivity Bias." *Public Opinion Quarterly* 77(3):641–665.

Judis, John and Ruy Teixeira. 2002. *The Emerging Democratic Majority*. New York: Scribner.

Kalkan, Kerem Ozan, Geoffrey C. Layman and Eric M. Uslaner. 2009. "'Bands of Others'? Attitudes Toward Muslims in Contemporary American Society." *The Journal of Politics* 71(3):1–16.

Karing, Anne. 2019. "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone." *Working paper* .

Katz, Irwin and R Glen Hass. 1988. "Racial Ambivalence and American Value Conflict: Correlational and Priming Studies of Dual Cognitive Structures." *Journal of Personality and Social Psychology* 55(6):893.

Kevins, Anthony and Stuart N. Soroka. 2018. "Growing Apart? Partisan Sorting in Canada, 1992-2015." *Canadian Journal of Political Science* 51(1):103–133.

Kinder, Donald R and David O Sears. 1981. "Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life." *Journal of Personality and Social Psychology* 40(3):414.

Lambert, Alan J and Alison L Chasteen. 1997. "Perceptions of Disadvantage Versus Conventionality: Political Values and Attitudes Toward the Elderly Versus Blacks." *Personality and Social Psychology Bulletin* 23(5):469–481.

Langer Research Associates. 2020. "COVID Collaborative Survey: Coronavirus Vaccination Hesitancy in the Black and Latinx Communities." *Covid Collaborative Report* .

Larson, Heidi J. 2020. *Why Vaccine Rumours Stick—And Getting Them Unstuck*. Oxford: Oxford University Press.

Lau, Richard R., Richard A. Smith and Susan T. Fiske. 1991. "Political Beliefs, Policy Interpretations, and Political Persuasion." *The Journal of Politics* 53(3):646–675.

Lee, Taeku. 1991. "Language-of-Interview Effects and Latino Mass Opinion." *John F. Kennedy School of Government Faculty Research Working Paper Series* 41(1).

Leeper, Thomas J, Sara B Hobolt and James Tilley. 2019. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis* pp. 1–15.

Levy, Morris and Matthew Wright. 2020. *Immigration and the American Ethos*. New York: Cambridge University Press.

Lewis-Beck, Michael, William G. Jacoby, Helmut Norpoth and Herbert F. Weisberg. 2008. *The American Voter Revisited*. Ann Arbor: University of Michigan Press.

Li, Peggy and Lila Gleitman. 2002. "Turning the Tables: Language and Spatial Reasoning." *Cognition* 83(3):265–294.

Lorini, Chiara, Francesca Ierardi, Claudia Gatteschi, Giacomo Galletti, Francesca Collini, Laura Peracca, Patrizio Zanobini, Fabrizio Gemmi and Guglielmo Bonaccorsi. 2020. "Promoting Influenza Vaccination among Staff of Nursing Homes According to Behavioral Insights: Analyzing the Choice Architecture during a Nudge-Based Intervention." *Vaccines* 8(4):600–619.

Malik, Amyn A, SarahAnn M McFadden, Jad Elharake and Saad B Omer. 2020. "Determinants of COVID-19 Vaccine Acceptance in the U.S." *EClinicalMedicine* 26:100495.

Markus, Hazel R and Shinobu Kitayama. 1991. "Culture and the Self: Implications for Cognition, Emotion, and Motivation." *Psychological Review* 98(2):224–253.

Mayda, Anna Maria. 2006. "Who is Against Immigration? A Cross-Country Investigation of Individual Attitudes Toward Immigrants?" *The Review of Economics and Statistics* 88(3):510–530.

McClosky, Herbert and Dennis Chong. 1985. "Similarities and Differences Between Left-Wing and Right-Wing Radicals." *British Journal of Political Science* 15(3):329–363.

McCroskey, James C and Jason J Teven. 1999. "Goodwill: A Reexamination of the Construct and its Measurement." *Communications Monographs* 66(1):90–103.

McDermott, Rose. 2011. Internal and External Validity. In *Cambridge Handbook of Experimental Political Science*, ed. James N. Druckman, Donald P. Green, James H. Kuklinski and Arthur Lupia. Cambridge: Cambridge University Press pp. 27–40.

McGorry, Susan Y. 2000. "Measurement in a Cross-Cultural Environment: Survey Translation Issues." *Qualitative Market Research: An International Journal* 3(2):74–81.

Mesoudi, Alex. 2011. *Cultural Evolution.* Chicago: University of Chicago Press.

Milkman, Katherine L., Mitesh S. Patel, Linnea Gandhi, Heather Graci, Dena Gromet, Quoc Dang Hung Ho, Joseph Kay, Timothy Lee, Jonathan Bogard, Ilana Brody, Christopher Chabris, Edward Chang, Gretchen B. Chapman, Jennifer Dannals, Noah J. Goldstein, Amir Goren, Hal Hershfield, Alex Hirsch, Jillian Hmurovic, Samantha Horn, Dean Karlan, Ariella Kristal, Cait Lamberton, Michelle N. Meyer, Allison Oakes, Maurice E. Schweitzer, Maheen Shermohammed, Joachim H. Talloen, Caleb Warren, Ashley Whillans, Kuldeep Yadav, Julian Zlatev, Ron Berman, Chalanda Evans, Christopher Snider, Eli Tsukayama, Christophe Van den Bulte, Kevin Volpp and Angela Duckworth. 2021. "A Mega-Study of Text-Message Nudges Encouraging Patients to Get Vaccinated at their Pharmacy." *SSRN Working Paper No. 3780356* .

Milkman, Katherine L, Mitesh S Patel, Linnea Gandhi, Heather N Graci, Dena M Gromet, Hung Ho, Joseph S Kay, Timothy W Lee, Modupe Akinola, John Beshears et al. 2021. "A Megastudy of Text-Based Nudges Encouraging Patients to Get Vaccinated at an Upcoming Doctor's Appointment." *Proceedings of the National Academy of Sciences* 118(20):1–3.

Morgan, G Scott, Elizabeth Mullen and Linda J Skitka. 2010. "When Values and Attributions Collide: Liberals' and Conservatives' Values Motivate Attributions for Alleged Misdeeds." *Personality and Social Psychology Bulletin* 36(9):1241–1254.

Mowbray, Fiona, Afrodita Marcu, Cristina A Godinho, Susan Michie and Lucy Yardley. 2016. "Communicating to Increase Public Uptake of Pandemic Flu Vaccination in the UK: Which Messages Work?" *Vaccine* 34(28):3268–3274.

Ndugga, Nambi, Olivia Pham, Latoya Hill, Samantha Artiga and Noah Parker. 2021. "Latest Data on Covid-19 Vaccinations by Race/Ethnicity." *Kaiser Family Foundation* .

Nelson, Thomas E. and Donald R. Kinder. 1996. "Issue Frames and Group-Centrism in American Public Opinion." *The Journal of Politics* 58(4):1055–1078.

Newman, Benjamin J. 2013. "Acculturating Contexts and Anglo Opposition to Immigration in the United States." *American Journal of Political Science* 57(2):374–390.

Noe-Bustamante, Luis, Antonio Flores and Sono Shah. 2019. "Facts on Hispanics of Cuban Origin, 2017." *Pew Research Center* .

Nyhan, Brendan and Jason Reifler. 2015. "Does Correcting Myths About the Flu Vaccine Work? An Experimental Evaluation of the Effects of Corrective Information." *Vaccine* 33(3):459–464.

Nyhan, Brendan, Jason Reifler, Sean Richey and Gary L Freed. 2014. "Effective Messages in Vaccine Promotion: a Randomized Trial." *Pediatrics* 133(4):835–842.

Ogunnaike, Oludamini, Yarrow Dunham and Mahzarin R Banaji. 2010. "The Language of Implicit Preferences." *Journal of Experimental Social Psychology* 46(6):999–1003.

Olson, Olivia, Corinne Berry and Nirbhay Kumar. 2020. "Addressing Parental Vaccine Hesitancy Towards Childhood Vaccines in the United States: A Systematic Literature Review of Communication Interventions and Strategies." *Vaccines* 8(4):590–615.

Parsons, Talcott. 1960. *Structure and Process in Modern Societies*. New York: Free Press of Glencoe.

Pérez, Efrén O. 2015. The Language-Opinion Connection. In *The Oxford Handbook of Polling and Polling Methods*, ed. Mike Alvarez and Lonna Atkinson. New York: Oxford University Press.

Pérez, Efrén O. 2016. "Rolling off the Tongue into the Top-of-the-Head: Explaining Language Effects on Public Opinion." *Political Behavior* 38(3):603–634.

Pérez, Efrén O. and Margit Tavits. 2015. "His and Hers: How Language Shapes Public Attitudes Toward Gender Equality." Paper Presented at the Annual Meeting of the Midwest Political Science Association.

Popkin, Samuel L. 1994. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Quinn, Sandra, Amelia Jamison, Donald Musa, Karen Hilyard and Vicki Freimuth. 2016. "Exploring the Continuum of Vaccine Hesitancy Between African American and White Adults: Results of a Qualitative Study." *PLoS Currents* 8.

Ralston, David A, Mary K Cunniff and David J Gustafson. 1995. "Cultural Accommodation: The Effect of Language on the Responses of Bilingual Hong Kong Chinese Managers." *Journal of Cross-Cultural Psychology* 26(6):714–727.

Reyna, C, PJ Henry, W Korfmacher and A Tucker. 2006. "Attributional Stereotypes as Cues for Deservingness: Examining the Role of Principled Conservatism in Racial Policy." *Journal of Personality and Social Psychology* 90:109–128.

Rokeach, Milton. 1960. *The Open and Closed Mind: Investigations into the Nature of Belief Systems and Personality Systems*. New York: Basic Books.

Rokeach, Milton. 1973. *The Nature of Human Values*. New York: Free Press.

Romaniuc, Rustam, Andrea Guido, Nicholas Mai, Eli Spiegelman and Angela Sutan. 2021. "Increasing Vaccine Acceptance and Uptake: A Review of the Evidence." *Available at SSRN: https://ssrn.com/abstract=3839654 or http://dx.doi.org/10.2139/ssrn.3839654* .

Ross, Michael, WQ Elaine Xun and Anne E Wilson. 2002. "Language and the Bicultural Self." *Personality and Social Psychology Bulletin* 28(8):1040–1050.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5):688–701.

Saha, Somnath, Miriam Komaromy, Thomas D Koepsell and Andrew B Bindman. 1999. "Patient-Physician racial Concordance and the Perceived Quality and Use of Health Care." *Archives of Internal Medicine* 159(9):997–1004.

Savoia, Elena, Rachael Piltch-Loeb, Beth Goldber, Cynthia Miller-Idriss, Brian Hughes, Alberto Montron, Juliette Kayyem and Marcia A. Testa. 2021. "Predictors of Covid-19 Vaccine Hesitancy: Socio-Demographics, Co-Morbidity and Past Experience of Racial Discrimination." *Vaccines* 9(7):767–780.

Scheve, Kenneth F and Matthew J Slaughter. 2001. "Labor Market Competition and Individual Preferences Over Immigration Policy." *Review of Economics and Statistics* 83(1):133–145.

Schildkraut, Deborah J. 2011. *Americanism in the Twenty-First Century: Public Opinion in the Age of Immigration*. Cambridge: Cambridge University Press.

Schmuck, Desirée and Jörg Matthes. 2017. "Effects of Economic and Symbolic Threat Appeals in Right-Wing Populist Advertising on Anti-Immigrant Attitudes: The Impact of Textual and Visual Appeals." *Political Communication* 34(4):607–626.

Schuman, Howard and Otis D. Duncan. 1974. Questions About Attitude Survey Questions. In *Sociological Methodology 1973-1974*, ed. Herbert L. Costner. Hoboken: Jossey-Bass Publishers pp. 232–251.

Schuman, Howard and Stanley Presser. 1977. "Question Wording as an Independent Variable in Survey Analysis." *Sociological Methods and Research* 6:151–170.

Sears, David O. 1986. "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature." *Journal of Personality and Social Psychology* 51(3):515–530.

Semyonov, Moshe, Rebecca Raijman and Anastasia Gorodzeisky. 2006. "The Rise of Anti-Foreigner Sentiment in European Societies, 1988-2000." *American Sociological Review* 71(3):426–449.

Sides, John and Jack Citrin. 2007. "European Opinion about Immigration: The Role of Identities, Interests and Information." *British Journal of Political Science* 37(03):477–504.

Skitka, Linda J and Christopher W Bauman. 2008. "Moral Conviction and Political Engagement." *Political Psychology* 29(1):29–54.

Skitka, Linda J, Christopher W Bauman and Edward G Sargis. 2005. "Moral Conviction: Another Contributor to Attitude Strength or Something More?" *Journal of personality and social psychology* 88(6):895.

Skitka, Linda J, Elizabeth Mullen, Thomas Griffin, Susan Hutchinson and Brian Chamberlin. 2002. "Dispositions, ideological scripts, or motivated correction? Understanding ideological differences in attributions for social problems." *Journal of Personality and Social Psychology* 83(2):470–487.

Slobin, Dan I. 1996. From 'Thought and Language' to 'Thinking for Speaking'. In *Rethinking Linguistic Relativity*, ed. John J. Gumperz and Stephen C. Levinson. Cambridge: Cambridge University Press pp. 70–96.

Sniderman, Paul M. 2018. "Some Advances in the Design of Survey Experiments." *Annual Review of Political Science* 21(259-275).

Sniderman, Paul M., David A. Northrup, Joseph F. Fletcher and Peter H. Russell. 1993. "Psychological and Cultural Foundations of Prejudice: The Case of Anti-Semitism in Quebec." *Canadian Review of Sociology and Anthropology* 30(2):242–270.

Sniderman, Paul M. and Louk Hagendoorn. 2007. *When Ways of Life Collide: Multiculturalism and its Discontents in the Netherlands*. Princeton: Princeton University Press.

Sniderman, Paul M., Pierangelo Peri, Rui J.P. de Figueiredo and Thomas Piazza. 2002. *The Outsider: Prejudice and Politics in Italy.* Princeton: Princeton University Press.

Sniderman, Paul M. and Sean M. Theriault. 2004. The Structure of Political Argument and the Logic of Issue Framing. In *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*, ed. Willem E. Saris and Paul M. Sniderman. Princeton: Princeton University Press pp. 133–165.

Sniderman, Paul and Phillip E. Tetlock. 1986. "Reflections on American Racism." *Journal of Social Issues* 42(2):173–187.

Spence, Patric R, Kenneth A Lachlan, David Westerman and Stephen A Spates. 2013. "Where the Gates Matter Less: Ethnicity and Perceived Source Credibility in Social Media Health Messages." *Howard Journal of Communications* 24(1):1–16.

Spence, Patric R, Kenneth A Lachlan and Donyale R Griffin. 2007. "Crisis Communication, Race, and Natural Disasters." *Journal of Black Studies* 37(4):539–554.

Street, Richard L, Kimberly J O'Malley, Lisa A Cooper and Paul Haidet. 2008. "Understanding Concordance in Patient-Physician Relationships: Personal and Ethnic Dimensions of Shared Identity." *The Annals of Family Medicine* 6(3):198–205.

Torres, Carlos, Lucy Ogbu-Nwobodo, Marcella Alsan, Fatima Cody Stanford, Abhijit Banerjee, Emily Breza, Arun G. Chandrasekhar, Sarah Eichmeyer, Mohit Karnani, Tristan Loisel, Paul Goldsmith-Pinkham, Benjamin A. Olken, Pierre-Luc Vautrey, Erica Warner and Esther Duflo. 2021. "Experimental Evidence on the Effectiveness of Non-Experts for Improving Vaccine Demand." *NBER Working Paper No. 28593* .

Turgeon, Luc and Antoine Bilodeau. 2014. "Minority Nations and Attitudes Towards Immigration: The Case of Quebec." *Nations and Nationalism* 20(2):317–336.

Verkuyten, Maykel, Kumar Yogeeswaran and Levi Adelman. 2019. "Intergroup Toleration and Its Implications for Culturally Diverse Societies." *Social Issues and Policy Review* 13(1):5–35.

Walsh, Bryan. 2020. "Axios-Ipsos Poll: The Racial Gap on Coronavirus Vaccine." *Axios, Accessed July 13 2021* .

Wetherell, Geoffrey A, Mark J Brandt and Christine Reyna. 2013. "Discrimination Across the Ideological Divide: The Role of Value Violations and Abstract Values in Discrimination by Liberals and Conservatives." *Social Psychological and Personality Science* 4(6):658–667.

Whetten, Kathryn, Jane Leserman, Rachel Whetten, Jan Ostermann, Nathan Thielman, Marvin Swartz and Dalene Stangl. 2006. "Exploring Lack of Trust in Care Providers and the Government as a Barrier to Health Service Use." *American Journal of Public Health* 96(4):716–721.

Wright, Matthew, Morris Levy and Jack Citrin. 2016. "Public Attitudes Toward Immigration Policy Across the Legal/Illegal Divide: The Role of Categorical and Attribute-Based Decision-Making." *Political Behavior* 38(1):229–253.

Yancey, George. 2010. "Who Has Religious Prejudice? Differing sources of Anti-Religious Animosity in the United States." *Review of Religious Research* pp. 159–171.

Yglesias, Matthew. 202. "Trump's Surprising Resilience with Hispanic Voters, Explained." *Vox* .

Yokum, David, Julie C Lauffenburger, Roya Ghazinouri and Niteesh K Choudhry. 2018. "Letters Designed with Behavioural Science Increase Influenza Vaccination in Medicare Beneficiaries." *Nature Human Behaviour* 2(10):743–749.

Appendix: Chapter 1

French Translation of Survey Prompt for Canadian Sample

"Cette étude s'intéresse à l'immigration et aux personnes autorisées à vivre au Canada. Les opinions divergent quant au type de personnes d'autres pays qui devraient être autorisées à (y vivre en permanence/être citoyens). Nous aimerions connaître votre opinion. Certains immigrants ayant demandé (le statut de résident permanent au Canada/la citoyenneté) ont répondu aux mêmes questions que vous au sujet de l'avortement, de la réglementation environnementale, des dépenses d'aide sociale et de l'alignement des partis. Vous verrez ci-dessous leurs réponses à ces questions ainsi que d'autres informations de base à leur sujet. Pour chaque paire de personnes que vous voyez, veuillez indiquer laquelle des deux personnes vous préféreriez voir obtenir (l'autorisation officielle de vivre de manire permanente au Canada/la citoyenneté canadienne) ou si vous prfrez que ni l'une ni l'autre n'obtienne (l'autorisation/la citoyenneté). Premièrement, considérez ces deux immigrants appliquant pour la (résidence permanente/citoyenneté)."

Full Average Marginal Component Effects

Figure 1: Abortion AMCEs by Subgroup (USA)

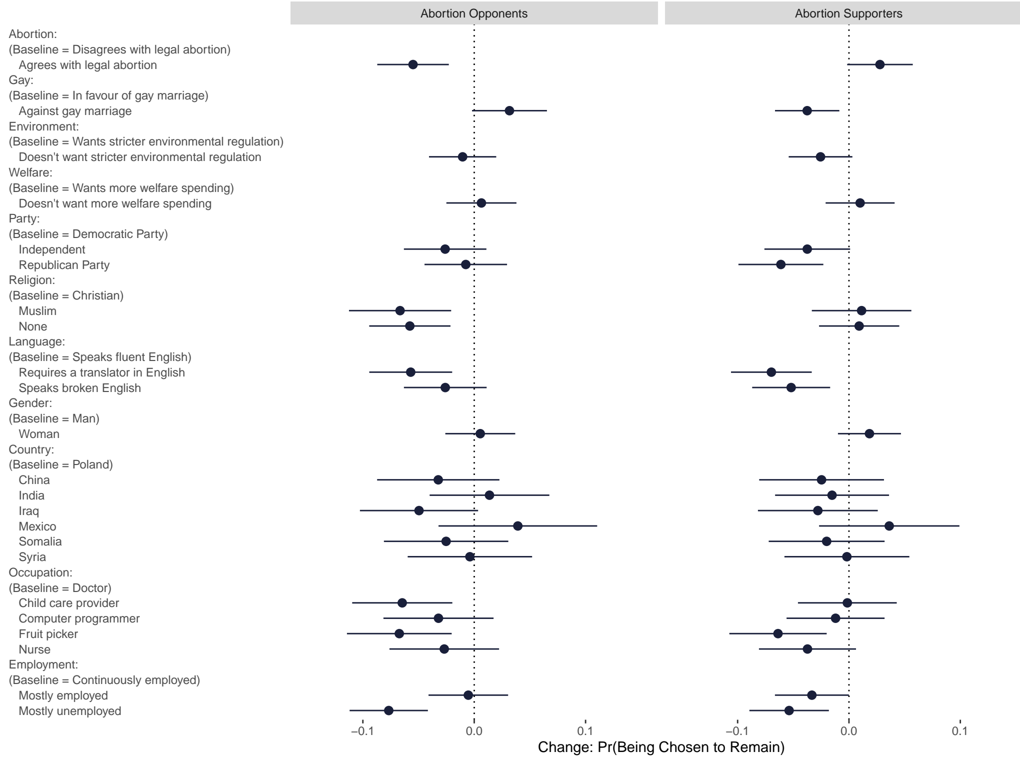


Figure 2: Welfare AMCEs by Subgroup (USA)

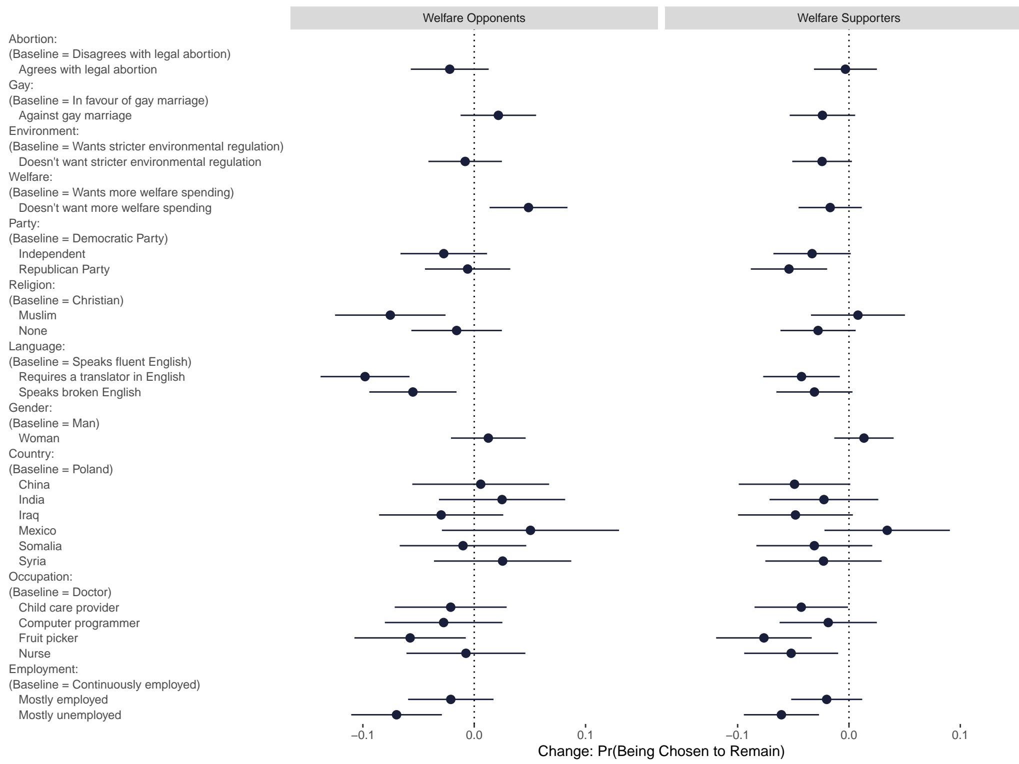


Figure 3: Environment AMCEs by Subgroup (USA)

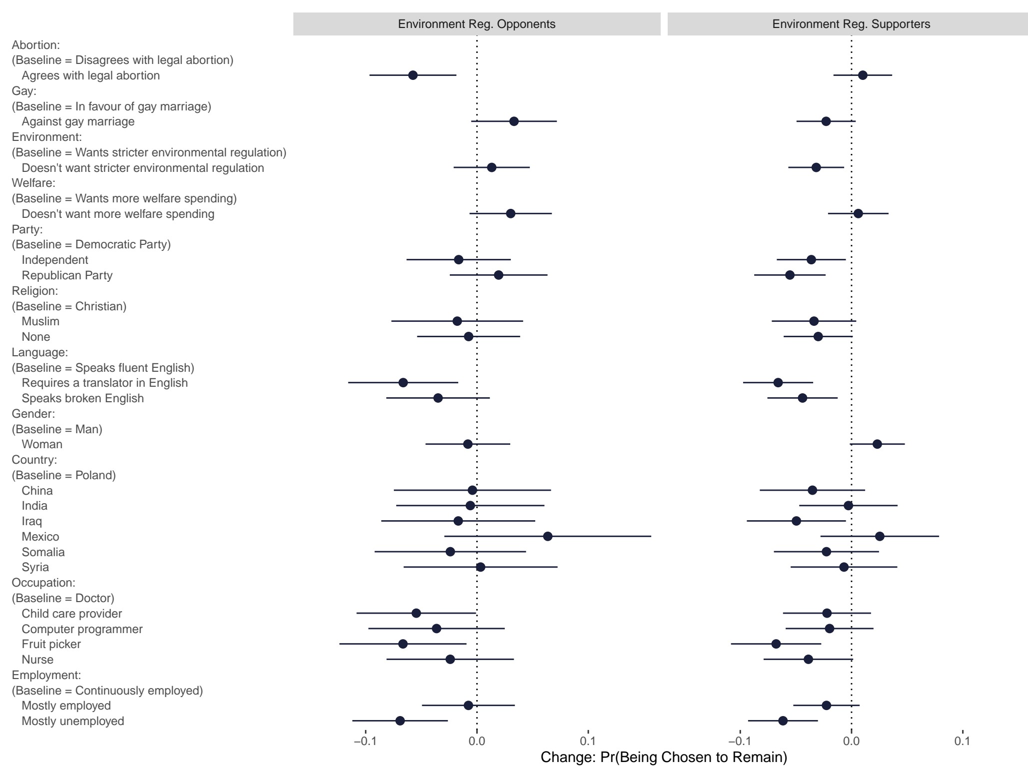


Figure 4: Abortion AMCEs by Subgroup (Canada)

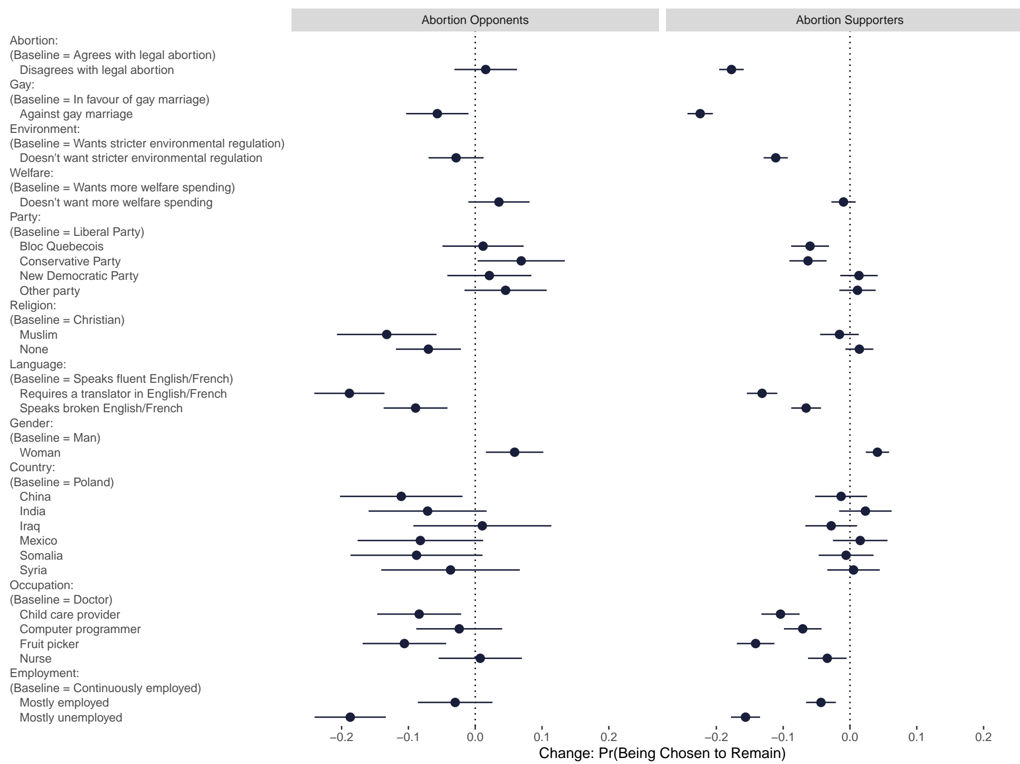


Figure 5: Welfare AMCEs by Subgroup (Canada)

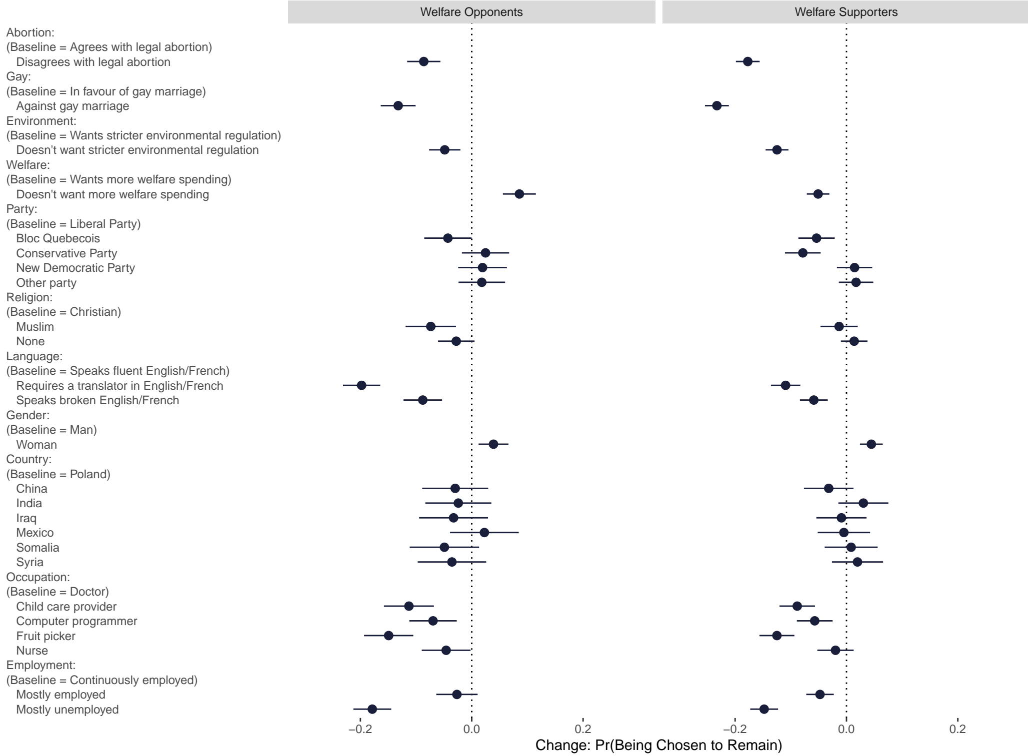
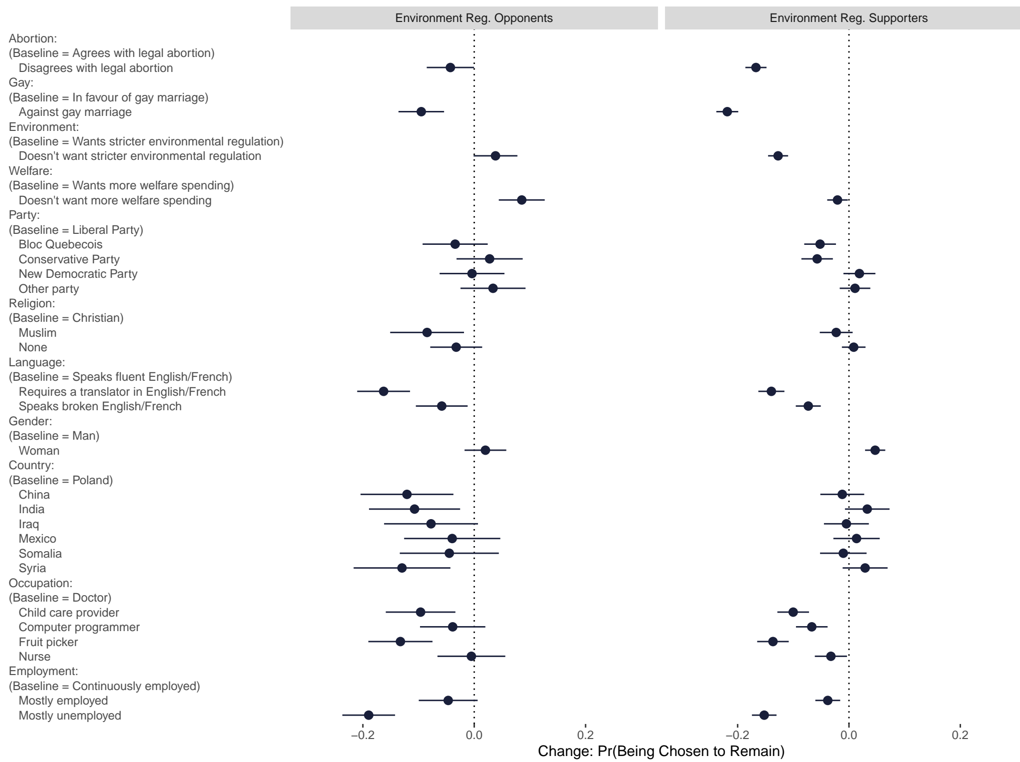


Figure 6: Environment AMCEs by Subgroup (Canada)



Robustness Checks

Randomization

In order to make sure that the randomization worked as it should, I regress various respondent characteristics on the putative immigrant attributes and conduct an F-test to assess whether the attributes are jointly insignificant predictors of respondent characteristics. In all cases, I obtain an F-statistic with a p-value greater than the conventional 0.05, suggesting that the attributes are jointly balanced across these characteristics. Detailed results of this test are available upon request.

Test of Coefficient Variance Across Prompt Wording Conditions

Results of formal tests of differences between treatment effects across "citizenship" and "permanent resident" wording prompts are presented below, where the treatment indicator "treat" is coded such that 0 = having received the permanent resident condition and 1 = having received the citizenship condition.

Table 1: United States Sample

treat × Woman	−0.067 (0.092)
treat × China	0.443*** (0.166)
treat × India	−0.102 (0.161)
treat × Iraq	0.273 (0.169)
treat × Mexico	0.264 (0.183)
treat × Somalia	0.179 (0.168)
treat × Syria	0.155 (0.162)
treat × Computer programmer	−0.173 (0.142)
treat × Doctor	−0.108 (0.135)
treat × Fruit picker	0.056 (0.145)
treat × Nurse	0.145 (0.136)
treat × Mostly employed	−0.070 (0.111)
treat × Mostly unemployed	0.206* (0.113)
treat × Muslim	0.343*** (0.118)
treat × No religion	0.024 (0.111)
treat × Speaks broken English	−0.292** (0.115)
treat × Speaks fluent English	−0.386*** (0.119)
treat × Disagrees with legal abortion	0.061 (0.092)
treat × In favor of gay marriage	−0.061 (0.102)
treat × Wants stricter environmental regulation	−0.079 (0.092)
treat × Wants more welfare spending	−0.057 (0.094)
treat × Independent	−0.019 (0.121)
treat × Republican Party	−0.019 (0.118)
Constant	−0.457*** (0.158)

Note: Respondent clustered standard errors in parentheses; *p<0.1; **p<0.05; ***p<0.01
Main effects omitted for economy of space

Table 2: Canadian Sample

treat × Woman	0.007
	(0.080)
treat × China	−0.105
	(0.153)
treat × India	−0.085
	(0.148)
treat × Iraq	−0.230
	(0.155)
treat × Mexico	0.057
	(0.173)
treat × Somalia	−0.082
	(0.158)
treat × Syria	−0.112
	(0.151)
treat × Computer programmer	−0.114
	(0.128)
treat × Doctor	−0.015
	(0.125)
treat × Fruit picker	0.032
	(0.128)
treat × Nurse	−0.089
	(0.123)
treat × Mostly employed	0.002
	(0.101)
treat × Mostly unemployed	0.133
	(0.104)
treat × Muslim	−0.043
	(0.098)
treat × No religion	0.011
	(0.095)
treat × Speaks broken English/French	0.134
	(0.096)
treat × Speaks fluent English/French	0.016
	(0.100)
treat × Disagrees with legal abortion	0.117
	(0.089)
treat × In favor of gay marriage	−0.054
	(0.089)
treat × Wants stricter environmental regulation	0.017
	(0.084)
treat × Wants more welfare spending	−0.062
	(0.081)
treat × Conservative Party	−0.140
	(0.137)
treat × Liberal Party	−0.177
	(0.129)
treat × New Democratic Party	−0.100
	(0.134)
treat × Other party	−0.157
	(0.130)
Constant	-1.372^{***}
	(0.146)

Note× Respondent clustered standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Main effects omitted for economy of space

Appendix: Chapter 2

Table 1: Balance Table

	Treatment Proportion (%)	Control Proportion (%)
Age: 18-34	25.0	24.8
Age: 35-64	53.7	54.9
Age: 65+	21.3	20.3
Education: Below high school	1.3	1.7
Education: High school	41.6	40.8
Education: University Degree	57.1	57.5
White	47.8	46.4
Income: Low	23	23.7
Income: Middle	42.7	42.4
Income: High	34.3	33.9
Religion: Catholic	29.8	31.9
Religion: Protestant	19.7	18.5
Religion: Other	2.8	2.8
Religion: No religion	47.8	46.8
Mean left-right ideology	42.99	42.89

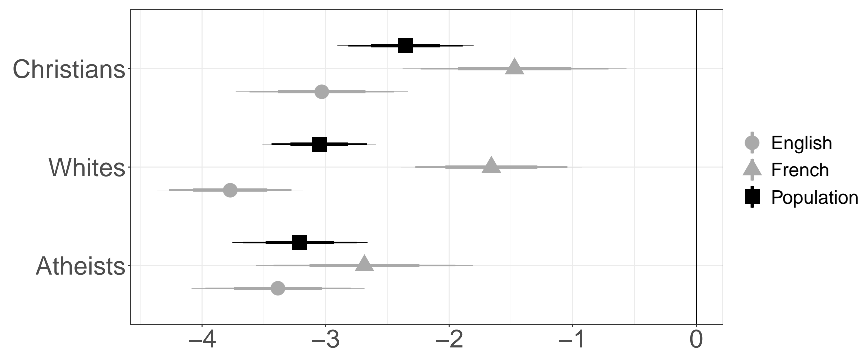
Table 2: ATE of "Like" Wording

	Combined Feeling Thermometers			
	(1)	(2)	(3)	(4)
Treatment (Like Version)	-2.61^{***}	-3.36^{***}	-4.05^{***}	-3.65^{**}
	(0.22)	(0.28)	(0.79)	(1.29)
French		-5.64^{***}	-5.19^{***}	-3.37^{***}
		(0.25)	(0.35)	(0.42)
Young			4.91^{***}	3.48^{***}
			(0.40)	(0.45)
University Degree			2.68^{***}	0.61
			(0.34)	(0.38)
White			0.53	0.01
			(0.35)	(0.40)
Poor			-1.00^{*}	-0.78
			(0.39)	(0.44)
Religious Minority			2.62^{*}	1.57
			(1.03)	(1.11)
No Religion			-0.23	-1.45^{***}
			(0.34)	(0.38)
Pro-Immigration				3.47^{***}
				(0.18)
Pro-Accommodation				2.78^{***}
				(0.20)
Treatment X French		1.12^{*}	1.82^{**}	2.05^{**}
		(0.46)	(0.66)	(0.78)
Treatment X Young			-1.27	-1.52
			(0.74)	(0.84)
Treatment X University Degree			-0.06	0.99
			(0.63)	(0.69)
Treatment X White			0.31	0.67
			(0.64)	(0.72)
Treatment X Poor			1.21	0.93
			(0.73)	(0.80)
Treatment X Religious Minority			0.38	0.83
			(1.86)	(2.05)
Treatment X No Religion			0.54	0.50
			(0.63)	(0.69)
Treatment X Pro-Immigration				-0.48
				(0.34)
Treatment X Pro-Accommodation				0.12
				(0.36)
_constant	72.39^{***}	74.72^{***}	71.10^{***}	55.43^{***}
	(0.12)	(0.16)	(0.44)	(0.71)
n	35,671	35,671	17,974	13,842
R^2	0.004	0.02	0.04	0.11

Source: Vox Pop Labs, 2015. $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Method: Ordinary Least Squares regression.

Figure 1: ATEs by Racial Group Question (Christians, Whites, and Atheists)



Source: Vox Pop Labs, 2015.

Appendix: Chapter 3

OLS Regression Estimates of ATEs

Table 1: ATEs on Likelihood of Taking a Covid-19 Vaccine

	(1)	(2)	(3)
Tyler v. Control	−0.012		
	(0.028)		
Lindsay v. Control		−0.041	
		(0.027)	
Tyler v. Lindsay			0.029
			(0.028)
Constant	0.487***	0.487***	0.446***
	(0.019)	(0.019)	(0.020)
Observations	667	681	630
R^2	0.0003	0.003	0.002
Adjusted R^2	−0.001	0.002	0.0001

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 2: ATEs on Delaying to Taking Vaccine

	(1)	(2)	(3)
Tyler v. Control	0.020		
	(0.033)		
Lindsay v. Control		−0.041	
		(0.033)	
Tyler v. Lindsay			0.061*
			(0.034)
Constant	0.475***	0.475***	0.435***
	(0.022)	(0.022)	(0.024)
Observations	502	508	462
R^2	0.001	0.003	0.007
Adjusted R^2	−0.001	0.001	0.005

Note: Standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table 3: ATEs on Expecting Negative Vaccine Side Effects

Tyler v. Control	−0.001		
	(0.015)		
Lindsay v. Control		0.012	
		(0.015)	
Tyler v. Lindsay			−0.012
			(0.015)
Constant	0.208***	0.208***	0.220***
	(0.010)	(0.010)	(0.011)
Observations	458	440	420
R^2	0.00000	0.001	0.001
Adjusted R^2	−0.002	−0.001	−0.001

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4: ATEs on Recommending to Others to Get Vaccine

	(1)	(2)	(3)
Tyler v. Control	−0.006		
	(0.028)		
Lindsay v. Control		−0.061**	
		(0.027)	
Tyler v. Lindsay			0.055*
			(0.029)
Constant	0.474***	0.474***	0.413***
	(0.019)	(0.018)	(0.020)
Observations	669	683	630
R^2	0.0001	0.008	0.006
Adjusted R^2	−0.001	0.006	0.004

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 5: ATEs on Believing Vaccines are Safe

	(1)	(2)	(3)
Tyler v. Control	−0.010		
	(0.025)		
Lindsay v. Control		−0.039	
		(0.024)	
Tyler v. Lindsay			0.028
			(0.025)
Constant	0.427***	0.427***	0.388***
	(0.017)	(0.017)	(0.017)
Observations	668	681	631
R^2	0.0002	0.004	0.002
Adjusted R^2	−0.001	0.002	0.0005

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 6: ATEs on Believing Vaccines are Effective

	(1)	(2)	(3)
Tyler v. Control	−0.023		
	(0.025)		
Lindsay v. Control		−0.016	
		(0.024)	
Tyler v. Lindsay			−0.007
			(0.025)
Constant	0.442***	0.442***	0.426***
	(0.017)	(0.016)	(0.018)
Observations	669	681	632
R^2	0.001	0.001	0.0001
Adjusted R^2	−0.0002	−0.001	−0.001

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 7: ATEs on Believing Vaccines are Safe for Black People

	(1)	(2)	(3)
Tyler v. Control	−0.013		
	(0.025)		
Lindsay v. Control		−0.022	
		(0.025)	
Tyler v. Lindsay			0.009
			(0.026)
Constant	0.433***	0.433***	0.410***
	(0.017)	(0.017)	(0.019)
Observations	671	682	631
R^2	0.0004	0.001	0.0002
Adjusted R^2	−0.001	−0.0003	−0.001

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Questionnaire and Experimental Stimuli

1. How closely, if at all, are you following news and information about the coronavirus pandemic?
    - Very closely
    - Somewhat closely
    - Not very closely
    - Not at all
2. Which of these, if any, is your main source of news and information about the pandemic?
    - Social media
    - Broadcast news (ABC, CBS, NBC)
    - FOX News
    - CNN or MSNBC
    - Public television or radio
    - Newspaper/newspaper websites
    - Other news websites
    - Government websites
    - Friends or family
    - None of these
3. How likely do you think it is that you will get covid-19?
    - Very likely
    - Somewhat likely
    - Not so likely
    - Not at all likely
    - Already had it
4. How often, if at all, do you use natural home remedies instead of medicines to try to prevent or cure illnesses?
    - Very often
    - Often
    - Sometimes
    - Occasionally
    - Never
5. Generally speaking, how much do you trust the federal government in Washington?
    - Completely
    - Mostly
    - Somewhat
    - Not much
    - Not at all

6. How much do you trust the Biden administration in specific?
    - Completely
    - Mostly
    - Somewhat
    - Not much
    - Not at all
7. How often do you think the government in this country can be trusted to look out for the interests of Black people?
    - Almost always
    - Often
    - Occasionally
    - Rarely
    - Never
8. How much, if at all, has racial discrimination interfered with your own ability to get good health care?
    - A great deal
    - A good amount
    - Just some
    - Only a little
    - Not at all
9. How important is being Black to your self-image?
    - Extremely important
    - Very important
    - Somewhat important
    - Not so important
    - Not important at all
10. How strong is your sense of belonging to the Black community?
    - Extremely strong
    - Very strong
    - Somewhat strong
    - Not so strong
    - Not strong at all
11. Screening Question: Have you received a covid-19 vaccine yet?
    - Yes
    - No
    - (⇒ If No, respondent continues to experiment. If Yes, respondent is rerouted to outcome questions)

12. Survey respondents are fully randomly assigned to view one of the videos corresponding to three treatment conditions:

    - *Condition 1:* Tyler Perry vaccine video (6 mins) (Link: `https://s3.valeria.science/videos-anja/GayleKingTylerPerrySmall.mp4`) [*NB:* the version of this video used as a treatment stops at exactly the 6:00 minute mark for brevity]
    - *Condition 2:* Sandra Lindsey vaccine video (3:38 mins) (Link: `https://s3.valeria.science/videos-anja/SandraLindsey.mp4`)
    - *Condition 3:* Unrelated Tyler Perry video (5:13 mins) (Link: `https://s3.valeria.science/videos-anja/GayleKingTylerPerryMovieStudio.mp4`)

13. Prior to viewing the videos, respondents will see the following prompts:

    - *Conditions 1 and 2:* "Some leaders in the Black community have recently received a covid-19 vaccine and have publicly voiced their perspective on covid-19 vaccines in order to share their views with others. Below, you will see a video clip in which (film and TV personality Tyler Perry)/(New York nurse Sandra Lindsey) receives a covid-19 vaccine and talks a little bit about it. The video is about (6/3) minutes long. After watching it, the survey will continue and you will see some more questions we would like your opinion on."
    - *Condition 3:* "Some Black celebrities have recently talked about how they have advanced in their respective industries. Below, you will see a video clip in which film and TV personality Tyler Perry talks about how he came to create his own film studio. The video is about 6 minutes long. After watching it, the survey will continue and you will see some more questions we would like your opinion on."

14. Manipulation Check: Had you seen this video before now?

    - Yes
    - No

15. How often do you watch TV interviews with interesting people or celebrities?

    - Often
    - Occasionally
    - Rarely
    - Never

16. Do you think celebrities have been unfairly prioritized to receive covid-19 vaccines?

    - Yes
    - No
    - I don't know

17. How likely are you to get a covid-19 vaccine when you are eligible to get it?

    - Certainly would get it
    - Probably would get it
    - Probably would not get it
    - Certainly would not get it

- (People who chose any category *other* than 'Certainly would not get it' in Q15 are presented with Q16)

18. If you were to get a covid-19 vaccine when you're eligible, would you wait before getting it? If so, for how long?
    - I would get it immediately
    - I would wait a few days
    - I would wait a few weeks
    - I would wait a few months
    - I would wait six months or more
19. If you got a covid-19 vaccine, how likely do you think it is that you would experience severe side effects?
    - Very likely
    - Somewhat likely
    - Not so likely
    - Not at all likely
20. How likely would you be to recommend to others that they get a covid-19 vaccine?
    - Very likely
    - Somewhat likely
    - Not so likely
    - Not at all likely
21. Overall, how much do you trust that covid-19 vaccines are safe?
    - Completely
    - Mostly
    - Somewhat
    - Not much
    - Not at all
22. Overall, how much do you trust that covid-19 vaccines really do prevent people from getting covid-19?
    - Completely
    - Mostly
    - Somewhat
    - Not much
    - Not at all
23. How confident are you that the covid-19 vaccines were sufficiently tested for safety and effectiveness among Black people specifically?
    - Very confident
    - Somewhat confident
    - Not so confident
    - Not confident at all