Can Leaders Persuade?
Examining Movement in Immigration Beliefs

Hassan Afrouzi Carolina Arteaga Emily Weisburst*

December 2021

Abstract

Can leaders change beliefs? If so, is it their rhetoric, identity, or the interaction of the two that matters? We construct an experiment where participants are exposed to leader speeches, where treatments include an anti-immigrant and pro-immigrant speech from Presidents Obama and Trump. We benchmark these treatments to versions recorded by an actor to control for speech messages. Our findings show that leader messages and sources influence beliefs. Leaders persuade when participants hear unanticipated messages from sources perceived as reliable, consistent with a Bayesian framework. This evidence supports the hypothesis that individuals will "follow their leader" to new policy positions.

JEL Classification: D83, C90

Keywords: Leaders, Political Beliefs, Partisan Identity, Polarization, Immigration

*Afrouzi: Columbia University Economics, ha2475@columbia.edu; Arteaga: University of Toronto Economics, carolina.arteaga@utoronto.ca; Weisburst: UCLA Luskin School of Public Affairs, weisburst@ucla.edu. We thank Jonathan Weisburst for sound editing and Michael Floyd for voice recording. We gratefully acknowledge support from ISERP seed grant from Columbia University, SSHRC Institutional Grant and REG from UCLA. We thank seminar participants at UC Merced, UCLA, the University of Toronto and the CEA.

Over the last several decades, political polarization and party tribalism in the United States have increased dramatically, creating divisions in society and stifling policy progress (Boxell et al., 2020; Schultz, 1996). At the same time, support for partisan *leaders* has divided along party lines; the difference in presidential approval ratings across parties was 85 (75) points for President Donald Trump (Barack Obama) relative to approximately 30 points for Presidents Jimmy Carter or Gerald Ford in the 1970s (Jones, 2021). These trends could stem from multiple factors, including the rise of social media and the segmentation of media exposure, which has reduced the overlap of information viewed by partisans (e.g. Di Tella et al., 2021; Levy, 2021; Allcott et al., 2020; Jo, 2017; Barberá et al., 2015), the introduction of widely available decentralized propaganda or "fake news" (Azzimonti and Fernandes, 2018), and new internet platforms for leaders to share their ideas.

The concurrent rise of political polarization and allegiance to partisan leaders raise fundamental questions about the ways that political beliefs are formed in the first place. The political information that we consume contains a bundle of characteristics; however, information silos mean that certain types of signals tend to come from sources with particular political reputations. A large body of work predicts that news media will slant their messages to cater to segmented consumer political preferences, potentially exacerbating polarization (Gentzkow and Shapiro, 2006; Baron, 2006; Mullainathan and Shleifer, 2005). A practical implication of these models is that news outlets and political leaders are unlikely to share views that are inconsistent with their partisan reputations. As a result, it is challenging to assess the separate role of political messages and leader sources in the determination of beliefs.

Understanding leader influence matters because the impact of leaders in a polarized democratic society is theoretically unclear. Partisan citizens may choose to elect candidates who already serve their existing political preferences. Alternatively, partisans may choose leaders they admire to act as their delegates, adhering to their policy positions even as they change over time. Do leaders reflect our existing beliefs or do they have the power to directly influence followers?

In this paper, we investigate three research questions to understand the influence of leaders. First, do partisan or political messages of leaders persuade people to change their beliefs? Second, does the brand identity of a leader prime people to change their beliefs? And third, do political messages have differing persuasive power when they come from different leaders?

We ask these questions in the context of beliefs about immigration in the United States. Immigration is a meaningful policy topic for voters; 52% (70%) of voters in the 2020 (2016) presidential election characterized immigration as being "very important" to their vote (Doherty et al., 2020, 2016). Further, views on immigration policy are highly polarized by political party, with Democrats holding generally pro-immigrant views and Republicans holding generally anti-immigrant views. Recent survey evidence shows that 50% of Democrats would like to see immigration levels increase, relative to only 13% of Republicans (Younis, 2020).

Our analysis leverages a novel randomized controlled experiment that independently varies the content and sources of political information signals. We utilize audio recordings of actual speeches given by Presidents Barack Obama and Donald Trump to construct these information treatments, simulating the ways in which voters receive information signals in the real world. First, we use original speeches on the topic of immigration to edit and extract an *anti-immigrant* segment and a *pro-immigrant* segment for *both* presidents. We then hired a voice actor to record replicate versions of each of these four immigration messages. The actor versions of the anti-immigrant and pro-immigrant speech messages (for each president) allow us to test the direct impact of partisan messages, given an ambiguous source. Relative to the president versions of speeches, the actor versions also allow us to test for the persuasive power of a particular president source, holding the content of a message fixed. Lastly, we include audio segments of non-ideological speeches for each president (ceremonial "turkey pardon" speeches on Thanksgiving) to test whether the brand identity or mere reputation of a president primes individuals to update their immigration beliefs. We embed these treatments in an online survey using a sample of Republican and Democrat participants, and stratify the treatment randomization within party.

Our experiment design is guided by a Bayesian framework in which an agent has a prior immigration belief and may update this belief when exposed to new political messages or sources. This framework provides several useful predictions. First, the model predicts that agents will update when they believe that a message conveys new information about the state of the world regarding immigration, as opposed to an uninformative political signal. Second, the framework predicts that agents will only be primed by the brand reputation of leaders to change their beliefs if they subjectively link the identity of a leader to the true state of the world regarding immigration. Lastly, the model provides useful insights about when particular sources increase the persuasive power of a particular message type. Here, persuasiveness can be decomposed into two multiplicative factors: 1) how unexpected a message is when it comes from a particular source and 2) the agent's subjective view of the reliability of that source. The framework predicts that leaders will have the most influence on beliefs when they express surprising or unexpected messages (ex. a pro-immigrant message from President Trump) to an audience of supporters who are likely to find the leader to be reliable (ex. Republican participants for President Trump messages).

We find that both political messages and sources influence beliefs. Participants from both parties update their beliefs based on political messages; they become more anti-immigrant (pro-immigrant) when they hear a message that is anti-immigrant (pro-immigrant), with effects that are larger for messages that oppose the party prior. This finding has weight because the speech treatments do not contain any facts about immigration or its consequences for society, but instead contain a mixture of emotional statements and policy proposals. Second, we do not find any independent priming effect related to the brand identity of leaders; participants do not change their views on immigration based on hearing a non-ideological message from either president. Lastly, leaders are more *persuasive* for particular messages and participant groups. Consistent with the Bayesian framework, we find that Republicans become more pro-immigrant when they hear pro-immigrant messages from President Trump and Democrats become more anti-immigrant when they hear anti-immigrant messages from President Obama, relative to participants who hear identical messages from the actor.

The total effect of hearing a political message from a president's voice, relative to the baseline prior in the control group that hears no audio message, moves beliefs by $\approx 2-10\%$ in the direction of the message. Because political messages move individuals toward the position of the message, speeches tend to polarize participants when messages align with the participant priors and reduce polarization when they do not align with the participant priors. The message portion of this result, estimated using the actor versions of the speeches relative to the no audio control group, drives this effect in nearly all cases. The implication of this finding is that partisans who only see messages from their own party may become more polarized than they otherwise would be if they were exposed to a variety of political information sources.

In two symmetrical cases, the presidential source directly persuades people to change their beliefs. For the anti-immigrant message from Obama and the pro-immigrant message from Trump, we find that the persuasion effect of the leader source comprises 53% and 39% of the total movement in beliefs for Democrats and Republicans, respectively. These leader source effects serve to reduce polarization relative to the control group because these messages oppose the party prior. The overall implication of the results is that *supporters of a leader are willing to follow their leaders to a new position.* In our context, this leads to a 30% reduction in polarization in beliefs about immigration, measured as the distance between Republicans and Democrats relative to the control group. However, in our general framework, any surprising position taken by a trusted leader could serve to sway beliefs, which could have the effect of increasing political division in a different context.

Our study builds on a large literature in behavioral economics and political economy about the nature of bias in information sources and the impact that new information can have on beliefs. When individuals are exposed to new information, they are aware of possible biases and perspectives of the information source (Gentzkow et al., 2018). At the same time, Baum and Gussin (2008) show that people perceive bias from identical messages when the source assigned to the content differs. Our work probes the relative importance of information message bias versus perceived bias of information sources, by separating these two determinants of information signals empirically.

The first contribution of our paper is to test whether partisan rhetoric from leaders, rather than factual information, can sway political beliefs. This question is unsettled in the economics literature on information signaling. A large literature has shown theoretically and empirically that partisan individuals often assign their own biases to messages, and can interpret identical neutral signals in differing ways in accordance with their existing priors (e.g. Baysan, 2021; Fryer Jr et al., 2019; Benoît and Dubra, 2019; Andreoni and Mylovanov, 2012). Given the finding that individuals can place so much of their own bias on identical messages, it is ex ante unclear whether political messaging could serve to move beliefs. A second strand of literature in economics and political science finds mixed results of the importance of messages intended to move beliefs in a particular direction (e.g. Levy, 2021; Song, 2021; Durante et al., 2019; Kalla and Broockman, 2018; Martin and Yurukoglu, 2017; Adena et al., 2015; Enikolopov et al., 2011; DellaVigna and Kaplan, 2007; Gerber et al., 2009). Work in this literature has also found that individuals have large misperceptions about facts relevant to policy issues, but that they update their beliefs when provided with new factual information (e.g. Grigorieff et al., 2020; Alesina et al., 2018; Bursztyn et al., 2018; Cruces et al., 2013). Some of this work finds meaningful impacts for *fact-based* information treatments on the policy topic of immigration (e.g. Alesina et al., 2018), but this research may or may not translate to our research question testing the messaging effects of *emotion-based* partisan speeches from leaders. Our experiment uses actual audio from presidential speeches in lieu of listed statistics or text narratives, and attempts to approximate the biased, incomplete and sometimes inaccurate messages common in the real world.

Our second contribution is to empirically test whether partisan leaders can persuade individuals to change beliefs, separate from the content of a political message. The fact that individuals perceive and assign bias to information signals could mean that they are more likely to distrust information that comes from partisan sources, and these types of sources may not be able to sway beliefs. Prior work has also explored the question of the persuasive power of partisan sources. Chiang and Knight (2011) find that newspaper endorsements of political candidates are most influential when they come from an unexpected source; for example, an endorsement of a Democratic candidate is most effective when it comes from a right-leaning newspaper, a finding that is consistent with our model framework and pattern of results. We complement and build on the findings of Chiang and Knight (2011) by building a randomized controlled trial that is able to both experimentally vary the ideological source and content of messages and examine the pattern of effects for different party groups.

Our work on leader persuasion also builds on a literature in political science and political communication. In this field, a number of papers have conducted analyses that experimentally vary the source of a policy statement within a survey and then test how beliefs may or may not change given a particular source (e.g. Barber and Pope, 2019; Broockman and Butler, 2017;

Boudreau and MacKenzie, 2014; Druckman et al., 2013; Nicholson, 2012; Bullock, 2011). This work finds mixed impacts of the persuasive power of partisan sources, typically focusing on a brief policy statement or a single partisan figure and then using one survey question to measure belief outcomes. Here, we contribute by building both a comprehensive experiment and model framework for interpreting our findings. Our experiment allows us to vary message ideology for two different leaders with very different partisan leanings, and also allows us to compare these treatments to identical messages with an ambiguous source, as well as to non-ideological messages from these leaders. Further, we reduce noise in our outcome measure of immigration beliefs by combining information from a large number of post-treatment questions that ask about different aspects of immigration policy. This robust range of treatments and outcome questions allows us to credibly isolate the effect of both partisan messages and partisan leader sources on belief formation.

1 Experiment Design

Our experiment is embedded within an online survey, where participants are exposed to different audio segments of presidential speeches within the survey environment. Participants are asked a series of background questions on demographics, political views, and news consumption prior to treatment. After treatment, participants are asked questions on their views on immigration. We recruited participants who identified as either Republicans or Democrats and stratified the randomization within these groups. The experiment was pre-registered with the American Economic Association.[1] The experiment was conducted during the period of the 2020 presidential election between October 16 - November 10, 2020.

1.1 Treatment

The experiment contains 11 treatment arms for each political party. These arms include 4 president immigration speeches, a pro-immigrant and an anti-immigrant speech for both Presidents Donald Trump and Barack Obama, and replicate versions of these 4 speeches recorded by a voice actor. We additionally include 2 presidential speeches with no ideological content; these speeches are ceremonious "turkey pardon" addresses for Trump and Obama. The last arm of the study is a control group that is not exposed to any audio speech treatment. Figure 1 depicts the experiment design.

The source material for the speeches are actual speeches delivered by each president. The audio for these speeches are extracted excerpts edited by a sound editor. It is important to note that there is no deception used in this study, as speeches are always introduced to study participants as "an excerpt of a presidential speech from President [Donald Trump or Barack Obama]."

[1] The pre-registration number is AEARCTR-0006552.

After the edited speeches were constructed, we hired a voice actor to record replicate versions of the edited speeches. The actor versions of speeches are introduced to study participants as "an excerpt of a presidential speech read by an actor."

We use an address given by Barack Obama on November 20, 2014 and an address given by Donald Trump on January 1, 2019 as source material. The purpose of the Obama speech was to introduce new protections for undocumented immigrants, including the Deferred Action for Childhood Arrivals (DACA) program, which created new safeguards for individuals who arrived in the U.S. illegally as children. Coupled with these reforms, the Obama speech also outlined additional provisions for border security. The purpose of the Trump speech was to provide a proposal to end a government shutdown related to immigration policy negotiations with Congress; this proposal included both new border security programs as well as a concession to protect DACA recipients. Both speeches were televised from the White House, prepared in advance, and delivered using a teleprompter. We are able to extract both pro-immigrant and anti-immigrant segments from each speech because each speech contains proposals to provide protections for immigrants as well as proposals to curb illegal immigration. The full text of the original speeches and treatment excerpts, as well as links to videos of the original speeches and audio files for the treatment segments are included in the Supplemental Materials Appendix S2. Similarly, we use original "turkey pardon" speeches to compose a non-ideological treatment speech for each president.[2]

Our experiment does not directly compare treatment effects across presidents, as the extracted segments are sourced from different speeches and cannot be constructed to be identical in content. However, the Trump and Obama speeches are similar in their rhetoric and ideology. Appendix Section A1 shows the opening paragraphs of each speech. The text of the speeches show similar sentiment within the anti-immigrant or pro-immigrant treatments, as well as several common phrases used by both presidents.

As a direct test of the relative strength of the messages, we ask participants how anti-immigrant or pro-immigrant they felt the treatment speech was, on a scale of 0 to 100. Figure 2 shows the perceived degree of message *strength* for the actor versions of speeches, or the perception that an anti-immigrant (pro-immigrant) speech is anti-immigrant (pro-immigrant). We plot perceptions from the actor speech groups so that responses are not colored by the president's reputation (Table A1 and Figure A2 shows corresponding estimates for all treatment groups). Strikingly, the Trump and Obama speeches are perceived nearly identically within

[2]The presidential tradition of a ceremonial turkey pardon typically consists of an event where a president gives a speech on the day before Thanksgiving, which includes the presentation of a live turkey who the president spares from being killed for a Thanksgiving dinner. We use the first Thanksgiving turkey pardon for each president, delivered on November 25, 2009 by Obama and on November 21, 2017 by Trump. These speeches are edited only for length and are also included in Supplemental Materials Appendix S2.

message type, for both Republicans and Democrats. Further, the message strength across party is also similar, with both groups viewing the pro-immigrant messages as somewhat stronger than the anti-immigrant messages. Again, while the experiment design does not rely on any direct comparisons *across* president or *across* party, these results are reassuring in that they show that the constructed speeches are similar for both Presidents and message types.

1.2 Survey and Sample Restrictions

We recruited participants who are eligible voters and identify as either Republicans or Democrats through a survey aggregation company called Cloud Research. Cloud Research partners with a number of different online survey panels to compile a sample that targets particular demographic groups and ensures quality standards for participant responses.

The survey begins with several screener questions to ensure a high level of quality of participant responses. These include asking for consent to participate, verification that the participant is not a robot and is able to listen to audio content on their device, and assurance that the participant is a U.S. citizen who is eligible to vote. Further, we include a quality check that simulates comprehension of the audio speech treatment. This check asks participants to listen to an audio segment of a weather forecast, and asks participants questions about what was discussed in the weather clip. For participants randomized to listen to an audio speech in the experiment, we also ask comprehension questions about the subject matter of the treatment speech. The full survey instrument is available in Supplemental Materials Appendix S1.

Figure A1 depicts the sample restrictions that were used to arrive at the final sample used in analysis. There were 23,639 individuals who attempted the online survey. Some individuals attempted the survey more than once and we keep only the first attempts for these participants, dropping 1,683 observations. Next, we geocode the locations of the IP addresses of survey takers and keep only the respondents located in the U.S., dropping 1,112 individuals. From this pool, we exclude individuals who did not consent to the study, did not have audio capability on their devices, answered that they were not U.S. citizens, or declared in the survey that they were affiliated with the opposite party from the party they said they were affiliated with in the recruitment screen; these restrictions remove 3,293 individuals. Within the survey, we also drop individuals who fail the quality check related to the weather forecast audio clip, or 3,549 individuals. We also remove a small number of individuals who are treated and did not answer the treatment comprehension questions correctly, or 443 individuals. Lastly, we remove 685 individuals who took the survey exceptionally quickly or slowly, keeping those who completed the survey in 4 to 30 minutes. The final sample contains 12,874 individuals, of which 6,993 are Democrats and 5,881 are Republicans. This translates to treatment group sizes of approximately 635 individuals in the Democrat sample and 535 individuals in the Republican sample.

Prior to treatment, we ask a series of demographic questions, including gender, race, age, education, employment status, party affiliation, and the candidate that the participant supported in the 2016 presidential election. We additionally ask a background question on four different political issues; immigration, gun control, abortion, healthcare, and taxes, as well as which of these issues is most important to the respondent's vote. We also ask how often participants receive their news from different modes (e.g. newspaper, TV, facebook), and different sources (e.g. Fox News, MSNBC). Lastly, we ask participants to state whether they are "fans of" a list of public figures and celebrities, which includes both Donald Trump and Barack Obama. Our questions about views on immigration and approval of Presidents Trump and Obama are embedded within larger lists of issues and public figures so as not to prime participants prior to treatment.

After treatment, we ask several questions about participant immigration views. We ask whether participants favor or oppose proposals to: expand construction of a wall on the U.S./Mexico border, hire more border patrol agents, require businesses to check the immigration status of workers, deport all immigrants living in the U.S. illegally or deport the subset of this population with a criminal record, allow immigrants living in the U.S. illegally to become citizens or allow the subset of this population who came to the U.S. illegally as children to become citizens (DACA recipients). We also ask whether participants view the following groups as positively or negatively contributing to U.S. society: immigrants working legally, immigrants not working legally, immigrants from English-speaking countries, immigrants from Spanish-speaking countries, and "dreamers" or undocumented immigrants brought to the U.S. by their parents as children (DACA recipients). Lastly, we ask participants whether immigrants benefit the economy, commit a disproportionate share of violent crimes, and if they feel immigration should be increased, decreased, or kept at present levels, as well as their overall perception of the contribution of immigrants to the U.S.,

Each of these questions have responses that are collapsed into either a pro-immigrant response, an anti-immigrant response, or a neutral response. We scale each question to have responses that range from 0 to 1, where 0 is a pro-immigrant answer, 1 is an anti-immigrant answer, and 0.5 is a neutral response. We then average across 16 questions to construct our anti-immigration views index. We also utilize a pro-immigrant index in the paper; this is simply calculated as 1 minus the anti-immigrant index.

Our post-treatment questions also include several questions about the treatment itself. We ask participants who heard an immigration treatment speech how anti-immigrant or pro-immigrant they perceived the speech to be. We also ask participants which president they thought gave the original speech out of a choice of the four most recent presidents (and an other option).[3] For individuals in the ambiguous source group, or those that received the actor

[3] This question serves as an attention check for individuals who received a speech with a revealed president

versions of the speeches, this question solicits the participants' best guess of the original source of the speech.

1.3 Sample Characteristics and Balance

Political party affiliation is an increasingly important group identifier in the U.S., where the political climate is highly polarized. We deliberately constructed our experiment to separately measure effects by party affiliation to account for these differences in ideology and group identity. Figure A3 plots the distribution of the key outcome, the anti-immigration beliefs index, for the control group that received no audio treatment. Republicans clearly hold views that are more anti-immigrant than Democrats in our sample, with the mass of each party separated along the distribution of this outcome.

We targeted a participant pool that would mimic the demographic characteristics of the Democratic and Republican parties, such that our results would best represent responses of individuals in these parties. Figure A4 shows that our sample is quite similar to the make-up of the national Democratic and Republican parties. One of our pre-treatment questions is copied from the national Gallup survey on immigration attitudes and asks whether participants think that immigration levels should be increased, decreased, or be kept at present levels. We can directly compare responses to this question to the national Gallup responses. Panel A and B show that our sample is similar to but slightly more moderate than the national parties, with larger share of Republicans and Democrats who feel that immigration should be kept at present levels. Panels C and D show that the study sample is quite similar to the national parties in demographics, though the sample slightly younger, more likely to be white, more likely to be female, and more educated. These differences may partly reflect the fact that the study pool was recruited for an online survey, which tends to attract individuals who have greater access and comfortability with technology and may also be more likely to be young and educated.

The summary statistics for both samples are shown in Table 1. As noted above, the Republican and Democrat samples are designed to reflect the demographics of their respective national parties. Likewise, the Republican sample has a higher share of respondents who are white, older, and from the southern U.S., while a smaller fraction are college educated. 36% of the Republican sample regularly watches FOX News, compared to 14.5% of the Democrat sample. We also leverage our survey data to estimate a prediction of whether participants are highly polarized on the topic of immigration prior to treatment (far right for Republicans or far left for Democrats), this measure shows that 36% of Democrats and 25% of Republicans are highly polarized.[4]

---

source; and in fact, nearly all participants in these groups guess the president correctly in this case (Figure A2).

[4]This prediction uses a highly granular set of controls from pre-treatment questions. We first determine who is in the top 25% of anti-immigrant (among Republicans) or pro-immigrant views (among Democrats) in the "No Audio" control group. We then estimate a probit regression that predicts this outcome separately for Democrats

15% of Republicans view immigration as their top political issue, compared to 5% of Democrats. Consistent with predicted party views, 45% of Republicans believe that immigration should be decreased, compared with only 14% of Democrats. Our sample has a high voter participation and is likely engaged in politics, as over 80% of both sample groups voted in the 2016 election. The parties are predictably stratified in their support of Trump and Obama; 92% (9%) of Republicans (Democrats) voted for Trump in 2016, while 78% (16%) of Republicans (Democrats) are fans of Donald Trump. Likewise, 89% of Democrats are fans of Obama, as compared to only 17% of Republicans.

Table 1 also shows that the study sample is balanced across treatment groups. For each characteristic and party sample, we regress the characteristic on indicators for the 11 treatment arms in the study and calculate the joint significance of these indicators. Successful randomization will be associated with a lack of joint significance for these indicators. Nearly all tests pass balance and do not show statistical significance. Only 5 of 62 tests are significant at the 10% level and 3 at the 5% level, similar to the number we should expect to fail due to chance alone (≈ 6 and 3 tests, respectively).

2 Bayesian Framework

What does theory predict about how beliefs should respond to information signals from leaders? This section outlines three distinct mechanisms that emerge from a Bayesian framework, which we use as a blueprint for our experiment design. The first mechanism is message content; messages potentially communicate information that are relevant for beliefs. The second mechanism is priming; independent of any message content, exposure to leaders might prime individuals to change their beliefs directly due to the identity of the leader. The third mechanism is persuasion; fixing the content of a message, the identity of a leader might amplify or diminish the impact of a message.

2.1 Setting

Let Ω denote the sample space for an individual's beliefs about immigration. We begin with the assumption that this set is binary and includes two outcomes: whether immigration is *favorable* or *unfavorable*. The goal is to define a probability space on *events* that contain messages as well as sources. We define the set of possible messages about immigration as $\mathscr{M}$ and the set of possible sources for a message as $\mathscr{S}$. The idea is that $\mathscr{M}$ contains all the messages that an individual might expect to receive from a leader in the context of immigration, and the $\mathscr{S}$ is the set of all the leaders who can be the sources of such messages. Given the discrete nature of outcomes, messages and sources, we assume all the three sets Ω, $\mathscr{M}$ and $\mathscr{S}$ are finite sets.

and Republicans using pre-treatment controls. We use coefficient estimates from this regression to extrapolate a predicted variable of being "polarized" for the full sample.

We can then define the augmented outcome space $\bar{\Omega}$ as $\Omega \times \mathscr{M} \times \mathscr{S}$ and model an individual's subjective beliefs as a probability space on $\bar{\Omega}$. Since $\bar{\Omega}$ is finite, to model this probability space, it suffices to assign a probability to every outcome on this set. Let $\mathbb{P} : \bar{\Omega} \to \mathbb{R}_+$ denote such a probability measure. Then, for any triple $(\omega, m, s) \in \bar{\Omega}$, $\mathbb{P}(\omega, m, s)$ captures the subjective beliefs of the individual about the joint probability of the outcome ω (for instance, immigration being favorable) with message m coming from leader s.

In our experiment, we elicit $\mathbb{P}$ from our subjects, but under different treatments. In all treatments, we will elicit the probability of an outcome $\omega \in \Omega$—for instance, measuring individual's belief about how *favorable* immigration is—for different treatment groups along with a control group.

1. *Control group*: We ask individuals about an outcome $\omega \in \Omega$ without exposing them to any message or source. In our notation, this corresponds to the marginal probability of ω, $\mathbb{P}(\omega)$.
2. *Treatment with message*: We elicit beliefs about an outcome $\omega \in \Omega$ after treating individuals with a message $m \in \mathscr{M}$ but without revealing the source—this treatment uses a voice actor to convey a message and is discussed in Section 1. In our notation, this corresponds to the conditional probability of ω on realization of the message m:
$$\mathbb{P}(\omega|m) = \frac{\sum_{s' \in \mathscr{S}} \mathbb{P}(\omega, m, s')}{\mathbb{P}(m)}$$
where $\mathbb{P}(m) = \sum_{s' \in \mathscr{S}} \sum_{\omega' \in \Omega} \mathbb{P}(\omega', m, s')$ is the marginal probability of the message m.
3. *Treatment with source*: Similarly, we elicit beliefs about an on outcome $\omega \in \Omega$ after treating individuals with a source $s \in \mathscr{S}$ without necessarily exposing them to any messages from that source about immigration—this treatment uses a non-ideological presidential speech about a "turkey pardon" on Thanksgiving, and is discussed in Section 1. In our notation, this belief corresponds to:
$$\mathbb{P}(\omega|s) = \frac{\sum_{m' \in \mathscr{M}} \mathbb{P}(\omega, m', s)}{\mathbb{P}(s)}$$
where $\mathbb{P}(m) = \sum_{m' \in \mathscr{M}} \sum_{\omega' \in \Omega} \mathbb{P}(\omega', m', s)$ is the marginal probability of the source s.
4. *Treatment with source and message*: Finally, we elicit beliefs about an outcome $\omega \in \Omega$ after treating an individual with a message about immigration from a *known* source. In our model, this corresponds to a double treatment with both a source s and a content m, and identifies:
$$\mathbb{P}(\omega|m, s) = \frac{\mathbb{P}(\omega, m, s)}{\mathbb{P}(m, s)}$$
where $\mathbb{P}(m, s) \equiv \sum_{\omega' \in \Omega} P(\omega', m, s)$ is the marginal probability of the double treatment (m, s) from the individual's perspective.

2.2 Message Content Effects

If messages contain relevant information with regards to the true state of the world, then they should move beliefs, independent of who communicates them. We call this the message content effect. Bayesian updating of beliefs provides a structure for how beliefs should move with respect to such content. In particular, a simple application of Bayes' rule implies that:

$$\mathbb{P}(\omega|m) = \mathbb{P}(\omega) \times \underbrace{\frac{\mathbb{P}(m|\omega)}{\mathbb{P}(m)}}_{\text{odds ratio}\equiv\Theta(m|\omega)} \tag{1}$$

Mapping this to the treatments described above, the left hand side of this equation captures the belief of an individual treated with the message m, while the $\mathbb{P}(\omega)$ on the right hand side corresponds to the belief of an individual in the control group. Given randomization of the treatments, beliefs in the control group can be interpreted similarly to *prior* beliefs relative to *posterior* beliefs in the treatment groups.

This expression leads to the well-known prediction of Bayesian models that the treatment effect should be related to the *odds ratio* of the treatment, $\Theta(m|\omega)$; i.e., how likely the message m is conditional on the outcome ω relative to the unconditional probability of m. It is straightforward to observe that if this odds ratio is larger than 1—i.e., the individual believes that observing m when ω holds is more likely that observing m unconditionally—then the individual revises their belief about ω upwards and vice versa.

Moreover, to map Equation (1) to an empirical specification, we need to recognize the issue that in its stated form, the treatment effect is multiplicative in the control group's belief, which is inconsistent with a linear regression model. In fact, the analogous linear model would be to consider this equation in logs and *log-odds ratios.* However, a simple log model would lead to the complication that for some individuals, the measured prior or posterior probability might be zero. Therefore, we consider a minor adjustment to account for these values by adding 1 to measured probability outcomes. With slight manipulation of Equation (1) one can derive the following expression:

$$\underbrace{\ln(1+\mathbb{P}(\omega|m))}_{\text{belief under treatment } m} = \underbrace{\ln(1+\mathbb{P}(\omega))}_{\text{belief among control group}} \underbrace{+\ln\left(1+\frac{\mathbb{P}(\omega)}{1+\mathbb{P}(\omega)}\times(\Theta(m|\omega)-1)\right)}_{\text{treatment effect of } m\equiv\Delta(m|\omega)} \tag{2}$$

which also characterizes the treatment effect of the message m in this linear specification as $\Delta(m|\omega)$. While the expression for $\Delta(m|\omega)$ is a bit more complicated than the corresponding log odds ratio of the message m, it is still a monotonic transformation of this object and therefore identifies this effect within this linear specification.

Remark 1. The treatment effect $\Delta(m|\omega)$ is increasing in the odds ratio $\Theta(m|\omega)$ and is zero if this ratio is equal to one.

A simple prediction from this framework is that individuals will update their beliefs toward a particular state of the world if they perceive a message to be more likely under that state of the worldAs an example, an individual who hears an anti-immigrant message will only update her beliefs to become more anti-immigrant if she believes that the message is more likely under the state of the world where immigration is *unfavorable*. If the individual does not believe that partisan or political messages are linked to the underlying state of the world, which may be the case if the individual is skeptical about the information value of partisan messages, then the individual should not update her beliefs.

2.3 Source Effects

In principle, the identity of a source delivering a message can matter in two different ways. The first potential effect of a source is *priming*, which captures the possibility that the mere identity of a leader affects an individual's belief, *independent* of the content communicated by that leader. The second possible effect of a source is *persuasion*. Persuasion is concerned with the possibility a particular message coming from a partisan leader should have a different effect on beliefs than the same message coming from an alternate source.

2.3.1 Priming

Consider a treatment where an individual is exposed to a message from the source that is not informative about immigration. Formally, such a treatment is captured by a pair (m, s) with the property that $\mathbb{P}(m|\omega, s) = \mathbb{P}(m|s), \forall \omega \in \Omega$; i.e., the message does not reveal any information about immigration, conditional on the identity of the source s. We can then measure the belief of the treated individual as $\mathbb{P}(\omega|m, s)$ and use Bayes' rule to write:

$$\mathbb{P}(\omega|m, s) = \mathbb{P}(\omega|s)\frac{\mathbb{P}(m|\omega, s)}{\mathbb{P}(m|s)} = \mathbb{P}(\omega|s) \tag{3}$$

where the second equality follows from the fact that the message is purposely selected to be irrelevant to immigration; $\mathbb{P}(m|\omega, s) = \mathbb{P}(m|s)$. This equation formally establishes that under this structure, the treatment effect of such a message would reveal the sole effect of the priming of a source's identity or reputation on the individual's belief about immigration.

We can also apply Bayes' rule one more time to write this priming effect in terms of the odds ratio for the source,

$$\mathbb{P}(\omega|s) = \mathbb{P}(\omega) \times \underbrace{\frac{\mathbb{P}(s|\omega)}{\mathbb{P}(s)}}_{\text{odds ratio}\equiv\Theta(s|\omega)} \tag{4}$$

where the left hand side captures the beliefs of the treated individual, $\mathbb{P}(\omega)$ on the right hand side represents the belief of the control group (or prior), and the third term—the odds ratio of the source—captures the priming effect for the source. It is, however, more useful in this case to think about the treatment effect in terms of *independence* of outcome and source, rather than the odds ratio itself. In particular, notice that the odds ratio in Equation (4) can be written as:

$$\Theta(s|\omega) = \frac{\mathbb{P}(s|\omega)}{\mathbb{P}(s)} = \frac{\mathbb{P}(s,\omega)}{\mathbb{P}(s)\mathbb{P}(\omega)} \tag{5}$$

This expression then delivers the following intuitive prediction:

Remark 2. The beliefs of the primed individual should be the same as the belief of a subject in the control group only if the identity of the source and the outcome ω are subjectively *independent*:

$$\mathbb{P}(\omega, s) = \mathbb{P}(\omega)\mathbb{P}(s) \Leftrightarrow \Theta(s|\omega) = 1 \tag{6}$$

Furthermore, similar to content effects, we can identify the odds ratio in Equation (5) in linear regression model once we transform it to a log-scale (or log of the probability plus 1). Formally, we can re-write Equation (4) as

$$\underbrace{\ln(1+\mathbb{P}(\omega|s))}_{\text{belief under treatment } s} = \underbrace{\ln(1+\mathbb{P}(\omega))}_{\text{belief among control group}} + \underbrace{\ln(1+\frac{\mathbb{P}(\omega)}{1+\mathbb{P}(\omega)} \times (\Theta(s|\omega)-1))}_{\text{treatment (priming) effect of } s \,\equiv\, \Delta(s|\omega)} \tag{7}$$

where the term $\Delta(s|\omega)$ characterizes the priming effect of the source: it is increasing in the odds ratio $\Theta(s|m)$ and is equal to zero if ω and s are perceived to be independent, i.e. if $\Theta(s|\omega) = 1$.

Here, we can easily see another straightforward prediction. Individuals who subjectively believe that the identity of a particular leader is linked to a state of the world will be subject to priming effects when they are exposed to *any* message from that leader. An example could be an individual who updates their beliefs to become more anti-immigrant when they are exposed to a uninformative message from President Trump, if this individual subjectively links Trump to the likelihood that immigration is *unfavorable*. However, if the individual believes that the likelihood of being exposed to an uninformative message from President Trump is independent from the true state of the world regarding immigration, she will not update her beliefs as a result of a source priming effect.

2.3.2 Persuasion

Next, we consider changes in beliefs that occur due to the interaction of particular sources with particular messages. Here we ask whether the identity of source could amplify or dampen the effect of a message on an individual's belief, which we refer to as *persuasion*.

Since persuasion specifically relies on the interaction of source and content, it is essential here to focus on treatments in which subjects jointly observe source and content. In this setting, how can we identify the sole effect of the persuasiveness of a source on beliefs? We address this by relying on a control group that observes the same content but does not observe the source, or the actor versions of our treatments, which are discussed in detail in Section 1. Intuitively, the joint treatment effect of source and content relative to such a control group should identify whether revealing the identity of the source amplifies or dampens the effect of the content.

In our framework, a belief under joint treatment (m, s) is given by $\mathbb{P}(\omega|m, s)$ and the belief under treatment m without revealing s is given by the marginal probability $\mathbb{P}(\omega|m)$. Our effect of interest is then the ratio of these two probabilities:

$$\Theta(s|\omega, m) \equiv \frac{\mathbb{P}(\omega|s, m)}{\mathbb{P}(\omega|m)} = \frac{\mathbb{P}(s|\omega, m)}{\mathbb{P}(s|m)} \tag{8}$$

where the second equality follows from Bayes' law and relates $\Theta(s|\omega, m)$ to the odds ratio of the sources' identity for outcome ω conditional on the content m. If this odds ratio is equal to one, then revealing the identity of the source is irrelevant to the perception of the content and corresponds to the absence of persuasion effects.

It is important to note that not revealing the source does not mean that the individual assigns no source to the content. On the contrary, the Bayesian subject recognizes that the source is in the set $\mathscr{S}$ and assigns probabilities to each source, given the content. Specifically, in our experiment we introduce treatments with ambiguous sources, or those recorded by an actor, as coming from a "presidential speech," without revealing the identity of the leader (see Section 1). This can be seen from the definition of marginal probability for the control group:

$$\mathbb{P}(\omega|m) = \sum_{s' \in \mathscr{S}} \mathbb{P}(s'|m) \times \mathbb{P}(\omega|s', m) \tag{9}$$

Using this equation and for a given source s, we can characterize the conditions under which persuasion exists—i.e., the odds ratio deviates from unity, $\Theta(s|\omega, m) \neq 1$.

Remark 3. For $s \in \mathscr{S}$, let $\neg s$ denote the event that the source is *not* s. Then, there is persuasion if:

$$0 \neq \Theta(s|\omega, m) - 1 = \frac{1}{\mathbb{P}(\omega|m)} \times \underbrace{(1 - \mathbb{P}(s|m))}_{\text{surprise}} \times \underbrace{(\mathbb{P}(\omega|s, m) - \mathbb{P}(\omega|\neg s, m))}_{\text{subjective reliability of source}} \tag{10}$$

Equation (10) shows that the odds ratio $\Theta(s|\omega, m)$ will be larger than unity if two conditions are simultaneously satisfied. First, there must be *surprise* in revealing the identity of the source, or the individual should not expect the message to come from the particular source. If there is no surprise, the content m by itself fully reveals the identity of the source s, meaning that $\mathbb{P}(s|m) = 1$, and we should not observe any differences between the treatment and control group in this experiment.

The second condition is more directly related to how the identity of the source alters the perception of the content. The term called the *subjective reliability of the source* measures how the subjects' belief about ω changes when the same message m is delivered by the source s relative to other sources $(\neg s)$. Here, an individual must find the focal source to be more subjectively reliable than alternative possible sources for the same message. If this perception is the same for all sources, meaning that $\mathbb{P}(\omega|s, m) = \mathbb{P}(\omega|s', m)$ for all s and s', then the source's identity is irrelevant to the perception of the message m and there will be no persuasion.

Finally, we can test the presence of persuasion in linear regression framework by considering the same log transformation in the previous sections. In particular, we can re-formulate Equation (8) to write

$$\underbrace{\ln(1+\mathbb{P}(\omega|s,m))}_{\text{belief under treatment }(s,m)} = \underbrace{\ln(1+\mathbb{P}(\omega|m))}_{\text{belief under treatment }m} \underbrace{+\ln(1+\frac{\mathbb{P}(\omega|m)}{1+\mathbb{P}(\omega|m)}\times(\Theta(s|\omega,m)-1))}_{\text{treatment (persuasion) effect of }s\equiv\Delta(s|m,\omega)} \tag{11}$$

where $\Delta(s|m,\omega)$ characterizes the treatment effect identified in this specification and identifies the persuasion effects. Persuasion is increasing in the odds ratio $\Theta(s|m,\omega)$ and is equal to zero when there is no persuasion. More generally, a positive $\Delta(s|m,\omega)$ implies that the source amplifies the effect of the message m.

Again, we can characterize the prediction using an example. Persuasion effects will only induce updating for surprising messages that come from sources that the individual views as reliable. A surprising message could be one where the immigration message contrasts with the partisan reputation of the leader, i.e. an anti-immigrant message that comes from President Obama, or a pro-immigrant message that comes from President Trump. An individual may find a partisan leader to be more subjectively reliable than an ambiguous source if that leader represents their political party, i.e. Republicans may find President Trump to be subjectively reliable while Democrats may find President Obama to be subjectively reliable.

3 Empirical Framework

This study utilizes a randomized controlled experiment, and the empirical approach will use simple comparisons of treatment arms to one another that leverage this randomization.

Our first objective is to test the impact of an anti-immigration or a pro-immigration message treatment on beliefs about immigration. To do this, we compare the actor recordings of a particular message to the control group that received no audio treatment, separately by political party, p.

$$\ln(1+\mathbb{P}(\text{Outcome}))_{ip} = \alpha + \beta \times \text{Actor}_{ip} + \gamma X_{ip} + \varepsilon_{ip} \tag{12}$$

$\mathbb{P}(\text{Outcome})$ represents the anti-immigration index for Republicans, $\mathbb{P}(\text{Anti})$, and represents the pro-immigrant index for Democrats, $\mathbb{P}(\text{Pro})$, where $\mathbb{P}(\text{Pro}) = 1 - \mathbb{P}(\text{Anti})$. As discussed above, the immigration indices are constructed on a probability scale, averaging the responses of 16 post-treatment questions in the survey for each participant, i. Following the Bayesian framework, we consider movement in the natural log of these outcomes in order to linearize the ratio form of Bayes' law. We add 1 to each probability in order to retain individuals where $\mathbb{P}(\text{Outcome}) = 0$. In robustness tests in Section 5 we show that our results do no change with alternative definitions of the outcome variable or alternate functional forms.

The treatment "Actor_{ip}" is an indicator for being assigned to an actor version of the immigration speeches, where each speech type (anti-immigrant (Trump), anti-immigrant (Obama), pro-immigrant (Trump), pro-immigrant (Obama)) are tested in separate regressions. β thus represents the average change in beliefs between the group treated with a message and the control group that did not listen to any audio in the experiment.

In our preferred specifications, we also include a vector of pre-treatment control variables selected from a double lasso procedure, X, to increase precision of the estimates (see Online Appendix A4).[5] As discussed below and shown in Figures A8 and A9, the results are robust to excluding these controls.

Next, we are interested in priming effects of the source on beliefs. Here, we test whether a non-ideological message delivered by a president with an anti-immigrant or pro-immigrant reputation will prime individuals to become more anti- or pro-immigrant.

$$\ln(\mathbb{P}(1+\text{Outcome}))_{ip} = \alpha + \beta \times \text{Turkey}_{ip} + \gamma X_{ip} + \varepsilon_{ip} \tag{13}$$

This test is identical to that above, but considers the treatment as the president turkey pardon message relative to the no audio control group. For completeness, we also examine the impact of president messages versus either the no audio control group or the turkey pardon message as a control group. Using the turkey pardon group as a control allows us to measure the impact of a pro-immigrant or anti-immigrant message on beliefs, fixing the presidential source.

Third, we investigate the channel of leader persuasion effects: whether partisan sources influence beliefs, fixing the content of a message. Here, we compare individuals treated with a president message to individuals treated with the identical message recorded by an actor.

$$\ln(\mathbb{P}(1+\text{Convinced}))_{ip} = \alpha + \beta \times \text{President}_{ip} + \gamma X_{ip} + \varepsilon_{ip} \tag{14}$$

In these regressions, we alter the outcome to be a measure of how convinced an individual is by a particular message. $\mathbb{P}$(Convinced) corresponds to the anti-immigrant (pro-immigrant) index for anti-immigrant (pro-immigrant) messages, for both presidents and parties. This structure allows us to easily portray the impact of sources on persuasion. β corresponds to the difference in persuasiveness or influence between a message delivered by a partisan source and an identical message from an unknown source.

[5] The procedure consists of first including all potential control variables from baseline and converting categorical variables into sets of indicators. Our set of potential controls is quite rich given the large number of survey questions we ask to participants prior to the treatment. We keep the controls selected from the LASSO minimization of a regression which predicts treatment assignment. We then repeat this exercise for a regression that predicts the outcome variable and take the union of controls selected from both procedures as our baseline controls. The list of controls included in the models is outlined in Online Appendix A4.

4 Results

4.1 Partisan Messages

Our first set of results relate to the impact of partisan messages on immigration beliefs. Table 2 and Figure 3 display the differences between treatment groups that heard an actor version of an anti-immigrant or pro-immigrant speech relative to the control group that did not hear any speech. Overall, these results show that participants update in the *direction* of the message that they hear, or a pro-immigrant message increases pro-immigrant beliefs, while an anti-immigrant message increases anti-immigrant beliefs. Movements *away from the party prior* appear to be larger and more significant; pro-immigrant effects are larger for Republicans and anti-immigrant effects are larger for Democrats. Only one of the anti-immigrant messages moves beliefs for Republicans (significant at the 10% level) and one of the pro-immigrant messages moves beliefs for Democrats. The effect sizes are symmetric, for Republicans (Democrats), pro-immigrant (anti-immigrant) messages change their subjective probability that immigration is favorable (unfavorable) by 3-5%. In addition to finding differences in the average effect across groups, the differences in the distributions of the immigration index that are significant are also statistically different across groups using Kolmogorov-Smirnov tests of equality.

These results are notable because the treatment speeches contain political and emotional language as well as policy proposals, but do not inform participants about any new impartial or factual information about immigration that they may not have known. This set of findings shows that the partisan messages that leaders use do affect the public, and that these messages can impact beliefs.

The Bayesian framework predicts that participants will update in a particular direction if they believe that the message that they heard is more likely under a particular state of the world. Ex ante, it is not clear whether there should be any evidence of updating from exposure to partisan messages about immigration. Participants may have strong priors and be unswayed by political messages of any type, believing that they do not hold any information about the true state of the world. However, this is not what we find: partisan messages do have the power to change beliefs.

4.2 Source Reputation Effects (Priming) and President Messages

Next, we test whether exposure to leaders primes participants to change their views based only on the leader's reputation on an issue. In our setting, Trump has a reputation for being anti-immigrant and Obama has a reputation for being pro-immigrant. We test this hypothesis by treating participants with a non-ideological message, a ceremonial turkey pardon speech delivered for the Thanksgiving holiday. We then observe whether participants who hear a non-ideological speech from Trump (Obama) become more anti-immigrant (pro-immigrant).

The results in Table 2 and Figure 3 show that respondents do not update their beliefs based on being primed by a leader's reputation alone. Neither party changes their immigration views after viewing a non-ideological (Turkey) message from either leader, when compared to the control group that does not hear any speech. These findings show that source priming does not meaningfully alter beliefs. Given the framework, this means that participants do not subjectively link a leader's identity to the true state of the world regarding immigration.

We also consider the direct impact of partisan messages delivered by presidents themselves, relative to either the control group that did not hear any speech or the group that heard a non-ideological (Turkey) message from the same president. These results are shown in the bottom panel of Table 2 and Figure 4. Similar to the messages delivered by the actor, the president messages also move participants in the *direction* of the message. This is directionally true even when the control group is the non-ideological turkey pardon message from a particular president, a test that allows us to fix the leader source and vary the message treatment. Again, messages that oppose the party prior are more impactful in influencing beliefs. Additionally, for Democrats, the effects of speeches voiced by President Trump have smaller or insignificant impacts on changing beliefs, relative to speeches voiced by President Obama.

4.3 Leader Persuasion: Source Effect Conditional on Message

In this section, we examine whether the identity of a leader can persuade for particular partisan messages, holding the content of a message fixed. We structure this test as a comparison of participants exposed to a president speech treatment relative to participants exposed to the identical speech message delivered by the actor. As noted in Section 1.1, the actor versions of immigration speeches are always introduced as "an excerpt of a presidential speech read by an actor." Consequently, the presumed source in the actor speeches is an ambiguous president.

The Bayesian framework predicts that the strength of leader persuasion will be a function of two different factors, the surprise of the message and the subjective reliability of the source of the statement. Surprise of the message is the participant's belief that the actual source was unlikely given the message, or whether the participant *would not expect* that a particular source would deliver a message of a certain type. Specifically, this is characterized as 1 minus the perceived probability of the actual source conditional on the message. Ex ante, it is not clear that participants will find any message to be surprising, as it could be the case that participants think that all politicians can easily change positions and any message type from any leader is equally likely.

In our setting, we can examine actual data on the surprise of the message using participant guesses about the identity of the true source when they heard an actor message. Panel A of Figure 2 shows our measure of message surprise: the share of respondents in the actor groups who guessed the president *incorrectly* when asked which president they thought gave the speech (out of four recent presidents and an other option). The plot shows that the most surprising messages are those where a leader delivers a message that opposes their reputation on immigration. The anti-immigrant Obama speech and the pro-immigrant Trump speech have more surprising sources from the perspective of participants in the study.

The Bayesian framework predicts that the second factor that drives leader persuasion is the subjective reliability of the source. Again, prior to conducting this experiment, it is not clear that participants will view some leaders as more reliable messengers than others. If individuals are persuaded solely by the information in messages, it could be the case that they prefer particular leaders only because those leaders have platforms that reflect what the individuals already believe.

We are not able to directly measure subjective reliability in our data, but we are able to illustrate differences in the favorability of leaders across parties which are likely correlated with reliability. Panel B of Figure 2 summarizes pre-treatment support for presidents across parties. Republicans overwhelmingly voted for Trump in 2016, with a share of 91.5%, in contrast to only 9% of Democrats. Similarly, when participants are asked whether they are fans of Obama and Trump, the vast majority of Republicans state that they are fans of Trump (78%) and are not fans of Obama (17%). A similar story is present for Democrats, with 89% stating they are fans of Obama, versus only 6% that are fans of Trump. This pattern suggests that Republicans may find Trump to be more subjectively reliable than Obama, and vice versa for Democrats.

Our empirical results from are consistent with the predictions of the Bayesian framework, with the highest level of leader persuasion coming from own party leaders when they deliver surprising messages. Table 3 shows the results of the comparison between president speeches and the identical speech delivered by the actor. The top panel of Table 3 uses the anti-immigration (pro-immigration) index outcome for Republicans (Democrats), as in the results examining the impact of partisan messages. These results show that Democrats become are persuaded by Obama to become more anti-immigrant and Republicans are persuaded by Trump to become more pro-immigrant, when either group hears these respective messages. The results are symmetric and equivalently sized, with a 3-4% decrease in subjective probability that immigration is favorable (Democrats) or unfavorable (Republicans).

An alternative way to display the results is to consider the outcome of whether a participant is *convinced* by a message, or whether the participant becomes more anti-immigrant (pro-immigrant) when she hears an anti-immigrant (pro-immigrant) message. This formulation is a convenient way to show persuasion effects and is shown in the bottom panel of Table 3 and in Figure 5. Again, the results are symmetric across parties, with leader persuasion increasing the likelihood that Republicans (Democrats) are convinced by 5% (8%) when Trump (Obama) delivers a pro-immigrant (anti-immigrant) speech.[6]

For both panels in Table 3, not only are these mean differences across groups significant, but the differences in the distributions of beliefs are also significantly different using Kolmogorov-Smirnov tests. Appendix Figures A5 and A6 plot the distribution of the outcome $ln(Convinced+1)$ for the president and actor groups for different messages, separately by party. Panel C of Figure A5 for Republicans (A6 for Democrats) shows that the distribution of the President Trump treatment group for the pro-immigrant (anti-immigrant) message is to the right (or more convinced) of the corresponding distribution for the actor version of this speech. In both cases, there appears to be a full distributional shift, which is not driven by a sub-segment of either group.

4.4 Decomposition of Source and Message Effects

Next, we decompose the total difference between a President treatment and no audio control group into a leader persuasion or source effect and the impact of the partisan message. To do this, we measure the distance between no audio control group and the actor version of a speech as the message component, and the remaining distance between the actor and the president as the source component. We estimate this decomposition using regressions that include the no audio group, actor group, and president group for each message and party.

Figure 6 plots this decomposition using the outcome of likelihood of being convinced by a message, defined as becoming more anti-immigrant (pro-immigrant) for an anti-immigrant (pro-immigrant) message. The darker bars show the message effects, which are positive for all messages and both parties, though they are larger for the messages opposing the presumed prior for each party. This is consistent with the discussion of message effects above, pro-immigrant rhetoric tends to make participants more pro-immigrant and vice versa.

The leader persuasion or source effects are represented by the lighter bars in both figures. The significant source effects in this study are for messages that oppose the party position on immigration and are delivered by the party leader, or the pro-immigrant (anti-immigrant) message from Trump (Obama) for Republicans (Democrats). The estimates imply that leader persuasion comprises 53% (39%) of the total effect for Democrats (Republicans) for these types of speeches. These effects are large, and suggest that leader persuasion can have a meaningful impact on the political beliefs of the public.

Interestingly, Figure 6 also suggests a potential secondary effect. While neither estimate is significant, the point estimates imply that Republicans are less likely to be convinced by Obama

[6] The magnitude of the changes in underlying probability are higher here than in the top panel of Table 3 because the bases differ. With the likelihood convinced, the coefficients measure the increase in Democrats' anti-immigration index or Republicans' pro-immigration index, both of which have a lower mean value in the reference group as compared to indices that would accord with the majoritarian views of each party (pro-immigrant for Democrats, anti-immigrant for Republicans.

when the message is pro-immigrant and Democrats are less likely to be convinced by Trump when the message is anti-immigrant. Notably, these messages are both more convincing when delivered by the actor, but are relatively less convincing when delivered by the opposing party leader (Table 2). Returning to the Bayesian framework, it may be the case that participants view the actor, or the ambiguous source, as a more reliable source of information than the opposition leader in this case.

4.5 Implications for Polarization

Leveraging the randomization of the experiment, we can also explore how the distance between Republican and Democrat immigration beliefs, or polarization, would be affected under different alternative scenarios. We estimate the change in polarization by estimating a regression using the anti-immigration index as the outcome that includes both parties in the sample. Each regression includes a treatment group for Democrats and Republicans relative to a corresponding control group for both parties, and records the change in distance between the two groups in the treatment groups versus the control group.

Table A2 and Figure 7 display the results of this exercise in terms of total polarization, and Figure A7 shows the party specific contributions to polarization. Given our findings that people move in the direction of the messages that they hear, we show that polarization increases when participants hear messages consistent with their priors and decreases when participants hear messages that oppose their priors. The most commonplace counterfactual in this exercise is what occurs when party leaders deliver messages to their followers that are consistent with the priors of their followers (Own Leader, With Prior). These messages increase polarization by 9-11%, but there is no difference between the actor and the president, or there is no additional leader persuasion or source effect conditional on the message. Messages consistent with party priors that come from the opposition leader also increase polarization, but these point estimates are not significant.

Partisan messages that are contrary to the priors of participants decrease polarization. The opposition message effects from the party leader or opposition leader both decrease polarization to a similar degree, ranging from 14-18%. As discussed above, the leader persuasion or source effect amplifies the impact of the message for party leaders, leading to an extra decrease in polarization of 19%, or a total decrease in polarization of $\approx 30\%$. In contrast, opposition leaders are not more persuasive than the ambiguous source for opposition messages. In fact, the figures show that for both Republicans and Democrats, the actor is more persuasive than the president for the opposition message from the opposition leader. This is likely because the actor is the more reliable source for these treatments, as discussed above. Collectively, these results imply that party leaders have the capacity to move their followers in a new alternative direction, and in this setting, yields the potential to reduce political polarization.

5 Robustness

In this section, we probe the robustness of the baseline findings. Figures A8 and A9 show that the results are stable across different specifications and sample restrictions. The first specification excludes the demographic covariates that are included in the baseline model using the double procedure (See Appendix A4) and shows very similar results.

Next, we adjust the sample to exclude observations collected after the 2020 election. We purposefully conducted our experiment during the period leading up to the national election in order to capture political beliefs and attitudes at this time, beginning on October 16, 2020. Due to a slower recruitment pace than anticipated, the study ran until November 10, 2020. The results excluding all responses collected on election day (November 3) or later, produces quite similar results to the baseline estimates.

In our baseline sample, we include the first attempts of the survey for individuals who attempted to take the survey multiple times. Next, we check that the results are consistent when dropping any attempt of a duplicate responder, and again find very similar results.

The last specifications in Figures A8 and A9 vary the construction of the immigration index outcome. Some of the questions in the survey have 5 option responses (e.g. Strongly Agree, Agree, Disagree, Strongly Disagree, No Opinion) and others have 3 option responses (e.g. Yes, No, No Opinion). In our baseline index, we translate all questions into 3 options; a pro-immigrant answer, an anti-immigrant answer, or a neutral response, collapsing the variation in the 5 option questions. In the robustness figures, we show the results when we consider the full variation from 5 option questions and find very similar results. Lastly, when we initially pre-registered the experiment, we specified an index that would use only 14 of the 16 post-treatment questions in the survey, because we thought that some of the questions might produce ambiguous responses, given that they related to a potentially more favored immigrant populations. These two questions asked whether respondents thought that immigrants from English-speaking countries or immigrants living in the U.S. legally contribute positively or negatively to society. In practice, these questions did not produce anomalous results and thus we added them to our baseline index to increase its information content. However, using only the 14 question version also produces similar results to the baseline estimates.

In Figures A10 and A11 we estimate the model using a simpler outcome of $ln(P(Outcome))$ rather than our baseline transformation of $ln(P(Outcome)+1)$. We utilize the latter transformation in our baseline regressions so as not to exclude individuals with purely anti-immigrant or pro-immigrant views, where $P(Outcome) = 0$. As additional checks, we include a specification using the inverse hyperbolic sine of $P(Outcome)$ and a simple linear regression using

$P(Outcome)$ as the dependent variable. The results of each of these alternative specifications produces results that are similar in direction and significance to the baseline estimates, given that each specification is a monotonic transformation of the key outcome, though the scale of the values differ given the difference of arguments in the expressions.

Our baseline regressions consider the log of an immigration index that has a continuous support of [0,1], given that it is an average of multiple questions each which can assume values of 0, 0.5, or 1. This construction has the benefit of incorporating gradations in beliefs about immigration for each participant. However, it is also useful to consider alternative specifications that discretize this outcome and employ logit or probit models.[7] Figure A12 and A13 finds results that are quite similar in direction and significance to the baseline estimates using a logit and probit model, as well as when using a linear probability model.

6 By Question Results

Our central results rely on an index composed of 16 different questions to incorporate multiple dimensions of immigration attitudes. Here, we test the robustness of the index to excluding individual questions, as well as which questions drive the effects that we find.

First, we consider the sensitivity of the index by re-calculating the estimates leaving out one question at a time in Figures A14 and A15 and estimating the models separately for each question in Figures A16 and A17. The results are more disperse when separately estimated by question, but across both sets of tests the findings are quite consistent with the baseline index outcome.

Figures A18 and A19 display the results by individual question, focusing on questions that move the results most for each party for display purposes.[8] The results are noisier when estimated at the individual question level, but interesting suggestive patterns do emerge. Republicans become more pro-immigrant on questions related to their "overall view" of immigrants, providing a path to citizenship for immigrants residing in the U.S., especially for immigrants who came to the U.S. illegally as children or "Dreamers," and their views on whether all immigrants residing in the U.S. illegally should be deported. This pattern suggests that when Republicans become more pro-immigrant their beliefs move on questions related to providing a legislative solution to citizenship, with a focus on providing this path to individuals who came to the U.S. as children, and not unilaterally deporting all immigrants. Arguably, these questions capture aspects of the immigration debate that are the most modest, or most open to support from right-leaning individuals.

For Democrats, we observe a somewhat symmetric pattern whereby partisans are more likely

[7] We discretize each outcome using the median value of the corresponding index in the no audio control group for each party.

[8] All possible individual question effects can be viewed in the following Figures A20 and A21.

to become anti-immigrant on immigration issues that penalize immigrants who may be viewed as the most egregious rule-breakers. Here, we observe anti-immigrant movement on questions related to the normative contribution of immigrants living in the U.S. illegally to society, the contribution of immigrants to the economy, interest in deporting immigrants living in the U.S. illegally, especially those with a criminal record, expanding immigration background checks for workers, and increasing funding for U.S. border patrol agents. Broadly, this pattern of findings suggests that Democrats are prone to adopting anti-immigrant attitudes towards immigrants that may be commonly viewed as the most severe "bad actors," and these policy areas may also be the most open to support from left-leaning individuals.

The "by question" results also highlight a crucial methodological strength of the study: the fact that we are able to aggregate responses of multiple dimensions of immigration attitudes into a summary index. Immigration is a complex policy topic, and partisans may change their beliefs on some aspects of immigration and not others. Additionally, using information from multiple distinct questions increases the total precision of our outcome. Had we taken the approach of earlier work and considered only 1-2 outcome questions with a general focus, we very well could have missed the robust and symmetric effects that we observe for both parties in our study.

7 Heterogeneity

Our experiment was designed to specifically address differences in reactions to treatment along political party lines, which we hypothesized would be the critical dimension of heterogeneity for our study. In line with our hypothesis, the results are symmetrical and mirrored by party, as party affiliation appears to be the key characteristic that predicts both prior beliefs about immigration as well as responses to different leaders.

There are several additional dimensions of heterogeneity that are worth exploring within party group, and we examine these dimensions to the extent possible given our sample size and statistical power. First, we test the importance of the strength of party affiliation as well as the strength of prior beliefs about immigration in Figures A22 and A23. We do this by including results for alternate samples that exclude individuals who voted for the opposite party in 2016 ("Flip Voters"), individuals who answered the pre-treatment immigration question by stating that immigration should be increased (decreased) who were Republicans (Democrats) ("Anti-Party Views"), individuals who were recruited as Democrats or Republicans but then identified as independents within the survey (Independents), individuals who stated that they were not a fan of their own party president (Non-Fan President), and lastly individuals who were predicted to have far right or left views on immigration using information on all pre-treatment questions in the survey ("Polarized").[9] Across each of these sub-groups of the data, the estimates are nearly

[9]"Polarized" is an estimated characteristic assigned to participants in the sample. We do this by predicting

identical to the baseline sample that includes all participants. An exception is that the message effects tend to move participants leftward for polarized Republicans, a feature that could be a result of these individuals having extreme right priors before treatment that do not permit them to move rightward to the same degree as other groups (Panel A, Figure A22). Overall however, this stability across groups suggests that neither moderates nor extreme members of each party are driving the results. Rather, consistent with the plots in Figures A5 and A6, it appears that there is a full shift of the distribution as a result of treatment, that is not stronger at either tail.

Next we test whether differences in political engagement or news consumption matter in Figures A24 and A25. Here we re-estimate the models excluding individuals who consume news from at least two platforms (Newspaper, TV, Twitter, Facebook) at least daily ("Multiple News Types"), individuals who consume news from both a left-leaning and a right-leaning source at least weekly ("Bipartisan News"), and individuals who did not vote in the 2016 election. Again, the estimates are remarkably stable across groups, suggesting that neither highly informed/engaged nor uniformed/unengaged participants are driving the findings of the study.

Lastly, we explore a number of demographic dimensions of heterogeneity in Figures A26 and A27. Overall, we find very few notable differences in the estimates according to gender, race, employment status, or educational attainment. A potential exception is age; the effect of leader persuasion conditional on the message (Panel C) appears to be slightly stronger for older Republicans and younger Democrats. However, even here, the confidence intervals of the estimates overlap.

In sum, we find limited variation in our findings across several dimensions of heterogeneity *within* party. This result could be attributable to a lack of power to detect these kinds of subgroup differences, as we did not design our experiment to discern these effects given resource constraints. At the same time, while the standard errors do moderately increase with the smaller sub-groups, the point estimates for each cut of the data are quite consistent. It appears that the shifts in beliefs that we observe are present for all segments of the distribution within party. The pattern of findings supports the hypothesis that *party affiliation* is the most important factor that determines responsiveness to treatment.

8 Discussion

Do leaders have the capacity to change beliefs? We leverage a novel randomized experiment to investigate this question in the context of U.S. immigration policy, a topic that is both important to voters and is characterized by a high level of political polarization. Our experiment uses audio

---

whether an individual has views in the top quartile of the party position (pro-immigrant for Democrats, anti-immigrant for Republicans) using only the control group and pre-treatment question responses. We use the coefficients from the estimation to extrapolate who would be polarized (without treatment) for the whole sample.

segments of speeches from Presidents Barack Obama and Donald Trump as well as replicate versions of these speeches recorded by an actor to isolate: (1) the importance of partisan rhetoric or messages, (2) the priming effects of partisan leader identity, and (3) the persuasive power of leaders for particular messages, holding the content of a message fixed. The experiment is grounded in a Bayesian framework that provides predictions that inform the results.

Our first key finding is that political messages cause participants to change their views. We deliberately test messages that do not contain facts and are instead composed of emotion-based arguments and policy proposals. Ex ante, it is not obvious that such messages will change beliefs, given that they contain no new substantive information about immigration. Both Republicans and Democrats are swayed by the pro-immigrant and anti-immigrant messages in this study, with effects that are stronger for messages that oppose party priors. An implication of these results is that the highly polarized nature of political beliefs in the U.S. could be related to the segmentation of information exposure by party. Information silos or "echo chambers" may be exacerbating polarization by shielding partisans from alternative information signals.

We also find new evidence that leaders can persuade individuals as messengers of information. While we do not find that participants change their beliefs through simple exposure to the brand identity of a leader, we do find that certain messages are more persuasive when voiced by particular leaders. These persuasion effects are measured as the effect of a leader source, holding fixed the substance of a message. We find that a leader is most persuasive when expressing statements that are unexpected to an audience of individuals who find the leader to be credible. Specifically, President Obama is most persuasive when voicing an anti-immigrant speech to Democrats and President Trump is most persuasive when voicing a pro-immigrant speech to Republicans.

This pattern of results illustrates that supporters will "follow their leader" to new political positions. In our context, these effects lead to a reduction in party polarization, but more generally, theory suggests that this need not be the case. Our framework implies that surprising or new positions for leaders will be most persuasive to followers, which may or may not increase longer-term polarization depending on the circumstance. A more abstract implication of our general findings is that a leader has the potential to create a "cult of personality" and use his or her persuasive power to shape policy decisions.

Our work is the first to credibly and comprehensively isolate the determinants of leader persuasion. Future research should continue to dissect the ways in which prominent partisan figures may shape, alter, or disrupt public opinion. Of particular interest is how leaders may not only change the beliefs of supporters but also change the actions of supporters. Understanding these dynamics will provide new insights into the strengths and fragilities of democratic governments as well as potential obstacles to policy progress.

References

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2015). Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics 130*(4), 1885–1939.

Alesina, A., A. Miano, and S. Stantcheva (2018). Immigration and redistribution. Technical report, National Bureau of Economic Research.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020). The welfare effects of social media. *American Economic Review 110*(3), 629–76.

Andreoni, J. and T. Mylovanov (2012). Diverging opinions. *American Economic Journal: Microeconomics 4*(1), 209–32.

Azzimonti, M. and M. Fernandes (2018). Social media networks, fake news, and polarization. Technical report, National Bureau of Economic Research.

Barber, M. and J. C. Pope (2019). Does party trump ideology? disentangling party and ideology in america. *American Political Science Review 113*(1), 38–54.

Barberá, P., J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science 26*(10), 1531–1542.

Baron, D. P. (2006). Persistent media bias. *Journal of Public Economics 90*(1-2), 1–36.

Baum, M. A. and P. Gussin (2008). In the eye of the beholder: How information shortcuts shape individual perceptions of bias in the media. *Quarterly Journal of political science 3*(1), 1–31.

Baysan, C. (2021). Persistent polarizing effects of persuasion: Experimental evidence from turkey. Technical report.

Benoît, J.-P. and J. Dubra (2019). Apparent bias: What does attitude polarization show? *International Economic Review 60*(4), 1675–1703.

Boudreau, C. and S. A. MacKenzie (2014). Informing the electorate? how party cues and policy information affect public opinion about initiatives. *American Journal of Political Science 58*(1), 48–62.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2020). Cross-country trends in affective polarization. Technical report, National Bureau of Economic Research.

Broockman, D. E. and D. M. Butler (2017). The causal effects of elite position-taking on voter attitudes: Field experiments with elite communication. *American Journal of Political Science 61*(1), 208–221.

Bullock, J. G. (2011). Elite influence on public opinion in an informed electorate. *American Political Science Review 105*(3), 496–515.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2018). Misperceived social norms: Female labor force participation in saudi arabia. Technical report, National Bureau of Economic Research.

Chiang, C.-F. and B. Knight (2011). Media bias and influence: Evidence from newspaper endorsements. *The Review of economic studies 78*(3), 795–820.

Cruces, G., R. Perez-Truglia, and M. Tetaz (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics 98*, 100–112.

DellaVigna, S. and E. Kaplan (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics 122*(3), 1187–1234.

Di Tella, R., R. H. Gálvez, and E. Schargrodsky (2021). Does social media cause polarization? evidence from access to twitter echo chambers during the 2019 argentine presidential debate. Technical report, National Bureau of Economic Research.

Doherty, C., J. Kiley, and N. Asheer (2020). In Changing U.S. Electorate, Race and Education Remain Stark Dividing Lines. Technical report, Pew Research Center.

Doherty, C., J. Kiley, N. Asheer, and C. Jordan (2020). Election 2020: Voters are Highly Engaged, but Nearly Half Expect Difficulties to Voting. Technical report, Pew Research Center.

Doherty, C., J. Kiley, and B. Johnson (2016). 2016 Campaign: Strong Interest, Widespread Dissatisfaction. Technical report, Pew Research Center.

Druckman, J. N., E. Peterson, and R. Slothuus (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review 107*(1), 57–79.

Durante, R., P. Pinotti, and A. Tesei (2019). The political legacy of entertainment tv. *American Economic Review 109*(7), 2497–2530.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011). Media and political persuasion: Evidence from russia. *American Economic Review 101*(7), 3253–85.

File, T. (2018). Characteristics of Voters in the Presidential Election of 2016. Technical report, U.S. Census.

Fryer Jr, R. G., P. Harms, and M. O. Jackson (2019). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. *Journal of the European Economic Association 17*(5), 1470–1501.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of political Economy 114*(2), 280–316.

Gentzkow, M., M. B. Wong, and A. T. Zhang (2018). Ideological bias and trust in information sources. *Working Paper*.

Gerber, A. S., D. Karlan, and D. Bergan (2009). Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics 1*(2), 35–52.

Grigorieff, A., C. Roth, and D. Ubfal (2020). Does information change attitudes toward immigrants? *Demography 57*(3), 1117–1143.

Igielnik, R. and A. Budiman (2020). The Changing Racial and Ethnic Composition of the U.S. Electorate. Technical report, Pew Research Center.

Jo, D. (2017). Better the devil you know: An online field experiment on news consumption.

Jones, J. (2019). New High in U.S. Say Immigration Most Important Problem. Technical report, Gallup.

Jones, J. (2021). Last Trump Job Approval 34%; Average Is Record-Low 41%. Technical report, Gallup.

Kalla, J. L. and D. E. Broockman (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review 112*(1), 148–166.

Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review 111*(3), 831–70.

Martin, G. J. and A. Yurukoglu (2017). Bias in cable news: Persuasion and polarization. *American Economic Review 107*(9), 2565–99.

Mullainathan, S. and A. Shleifer (2005). The market for news. *American economic review 95*(4), 1031–1053.

Nicholson, S. P. (2012). Polarizing cues. *American journal of political science 56*(1), 52–66.

Schultz, C. (1996, 04). Polarization and Inefficient Policies. *The Review of Economic Studies 63*(2), 331–344.

Song, L. (2021). The heterogeneous effects of social media content on racial attitudes.

Younis, M. (2020). Americans Want More, Not Less, Kmmigration for First Time. Technical report, Gallup.

Table 1: Summary Statistics and Balance Tests

	Republicans				Democrats			
	Mean	S.D.	F-Test	P-Value	Mean	S.D.	F-Test	P-Value
Female	0.469	(0.499)	1.002	0.439	0.511	(0.500)	0.588	0.825
White	0.892	(0.311)	0.701	0.724	0.649	(0.477)	0.747	0.681
Black	0.029	(0.169)	0.868	0.563	0.186	(0.389)	1.140	0.327
Hispanic	0.040	(0.195)	1.261	0.247	0.083	(0.276)	0.621	0.798
Asian	0.029	(0.169)	0.415	0.940	0.065	(0.247)	0.489	0.898
18-24 years	0.081	(0.272)	0.732	0.695	0.154	(0.361)	0.908	0.525
25-34 years	0.184	(0.388)	1.193	0.290	0.257	(0.437)	1.240	0.260
35-44 years	0.219	(0.414)	0.904	0.528	0.219	(0.413)	0.828	0.601
45-64 years	0.324	(0.468)	1.004	0.437	0.235	(0.424)	0.804	0.625
65+ years	0.192	(0.394)	0.233	0.993	0.136	(0.343)	0.497	0.893
Northeast	0.166	(0.372)	0.678	0.746	0.220	(0.414)	0.529	0.871
Midwest	0.241	(0.427)	0.650	0.772	0.213	(0.410)	1.356	0.194
South	0.432	(0.495)	0.883	0.548	0.364	(0.481)	2.107	0.021
West	0.161	(0.368)	0.582	0.830	0.203	(0.403)	0.986	0.453
College or More	0.460	(0.498)	0.597	0.817	0.520	(0.500)	1.099	0.359
Full-time Employed	0.425	(0.494)	1.397	0.175	0.446	(0.497)	1.016	0.427
News (Weekly+): Facebook	0.397	(0.489)	0.992	0.448	0.386	(0.487)	0.572	0.839
News (Weekly+): Twitter	0.175	(0.380)	0.666	0.757	0.263	(0.440)	1.690	0.077
News (Weekly+): TV	0.584	(0.493)	0.781	0.647	0.601	(0.490)	0.755	0.672
News (Weekly+): Newspaper	0.250	(0.433)	1.974	0.032	0.320	(0.467)	1.266	0.244
News (Weekly+): FOX News	0.363	(0.481)	1.060	0.390	0.146	(0.353)	0.827	0.603
News (Weekly+): MSNBC	0.083	(0.276)	0.842	0.588	0.193	(0.394)	0.771	0.657
Party: Independent	0.077	(0.266)	1.672	0.081	0.054	(0.226)	2.173	0.017
Polarized (Estimated)	0.247	(0.431)	1.446	0.153	0.358	(0.479)	1.137	0.329
Voted (2016)	0.842	(0.365)	0.724	0.703	0.836	(0.371)	0.781	0.647
Voted for Trump (2016)	0.916	(0.277)	1.388	0.179	0.086	(0.280)	0.536	0.865
Fan of Trump	0.780	(0.414)	1.479	0.140	0.060	(0.237)	1.116	0.345
Fan of Obama	0.168	(0.374)	0.779	0.649	0.889	(0.314)	0.844	0.586
Immigration: Top Issue	0.156	(0.363)	1.139	0.328	0.051	(0.219)	1.191	0.291
Immigration: Should Increase	0.114	(0.318)	1.032	0.413	0.365	(0.482)	0.875	0.556
Immigration: Should Decrease	0.452	(0.498)	0.898	0.534	0.142	(0.349)	1.432	0.159
N	5881				6993			

Notes: Table summarizes demographic characteristics of sample from questions asked prior to intervention. Balance tests use separate regressions of the demographic characteristic on the full set of treatment dummies, F-tests refer to joint significance of treatment assignment. “Party: Independent” refers to individuals recruited to the survey as Democrats or Republicans who indicate within the survey that they are Independents. “Polarized (Estimated)” is a probability of being in the top 25th percentile of anti-immigrant (pro-immigrant) views for Republicans (Democrats) constructed in the following way: predict this outcome in the post-treatment immigration index in the no audio control group using only pre-treatment characteristics and use the coefficients to predict this outcome for the full sample. “Fan of Obama/Trump” comes from a question where we ask whether participants are a fan of these presidents. Prior to treatment, we ask participants to identify the top issue relevant to their vote (“Immigration: Top Issue”) and whether they think immigration should be increased, decreased or remain constant (“Immigration: Should Increase (Decrease)”).

Table 2: Impact of Messages

Message	Source	Control	Republicans: ln(Probability Anti-Immigrant) β	(S.E.)	%Diff.	P(Dist.)	N	Democrats: ln(Probability Pro-Immigrant) β	(S.E.)	%Diff.	P(Dist.)	N
Baseline Messages												
Anti	Actor, Trump	No Audio	0.010*	(0.006)	2.86%	*0.082*	1076	-0.017***	(0.005)	-4.05%	*0.015*	1274
Anti	Actor, Obama	No Audio	0.009	(0.006)	2.44%	*0.144*	1082	-0.018***	(0.005)	-4.27%	*0.005*	1255
Pro	Actor, Trump	No Audio	-0.011*	(0.006)	-2.95%	*0.047*	1095	0.003	(0.005)	0.67%	*0.187*	1294
Pro	Actor, Obama	No Audio	-0.019***	(0.006)	-5.15%	*0.007*	1131	0.011**	(0.005)	2.67%	*0.097*	1313
Turkey	Trump	No Audio	0.008	(0.006)	2.20%	*0.549*	1086	-0.002	(0.005)	-0.42%	*0.642*	1337
Turkey	Obama	No Audio	0.006	(0.006)	1.77%	*0.187*	1056	-0.002	(0.005)	-0.37%	*0.879*	1281
President Messages												
Anti	Trump	No Audio	0.012**	(0.006)	3.31%	*0.070*	1103	-0.010**	(0.005)	-2.47%	*0.106*	1267
Anti	Trump	Turkey	0.005	(0.006)	1.40%	*0.389*	1079	-0.007	(0.005)	-1.74%	*0.521*	1282
Anti	Obama	No Audio	0.015**	(0.006)	4.18%	*0.004*	1061	-0.032***	(0.005)	-7.65%	*0.000*	1283
Anti	Obama	Turkey	0.007	(0.006)	2.00%	*0.615*	1007	-0.030***	(0.005)	-7.11%	*0.000*	1242
Pro	Trump	No Audio	-0.026***	(0.006)	-7.00%	*0.000*	1106	-0.004	(0.005)	-0.97%	*0.305*	1302
Pro	Trump	Turkey	-0.032***	(0.006)	-8.71%	*0.000*	1082	-0.001	(0.005)	-0.35%	*0.797*	1317
Pro	Obama	No Audio	-0.014**	(0.006)	-3.80%	*0.045*	1080	0.008*	(0.005)	1.88%	*0.004*	1336
Pro	Obama	Turkey	-0.021***	(0.006)	-5.76%	*0.001*	1026	0.009*	(0.005)	2.19%	*0.025*	1295

Notes: *p<0.1, **p<0.05, ***p<0.01. This table presents the estimates of the different message treatments, relative to either the no audio control group or the Turkey Pardon message for a particular president. Outcomes are constructed as $ln(P(Outcome)+1)$, where the outcome is the index probability anti-immigrant (pro-immigrant) for Republicans (Democrats). All regressions additionally include covariates selected by the double lasso procedure to improve precision (See Appendix A4). "% Diff." calculates the implied change in outcome *probability*, or the untransformed probability index, due to treatment, relative to the mean for the control group. "$P(Dist.)$" is the p-value from a Kolmogorov-Smirnov test of equality of the distributions of the control and treatment groups.

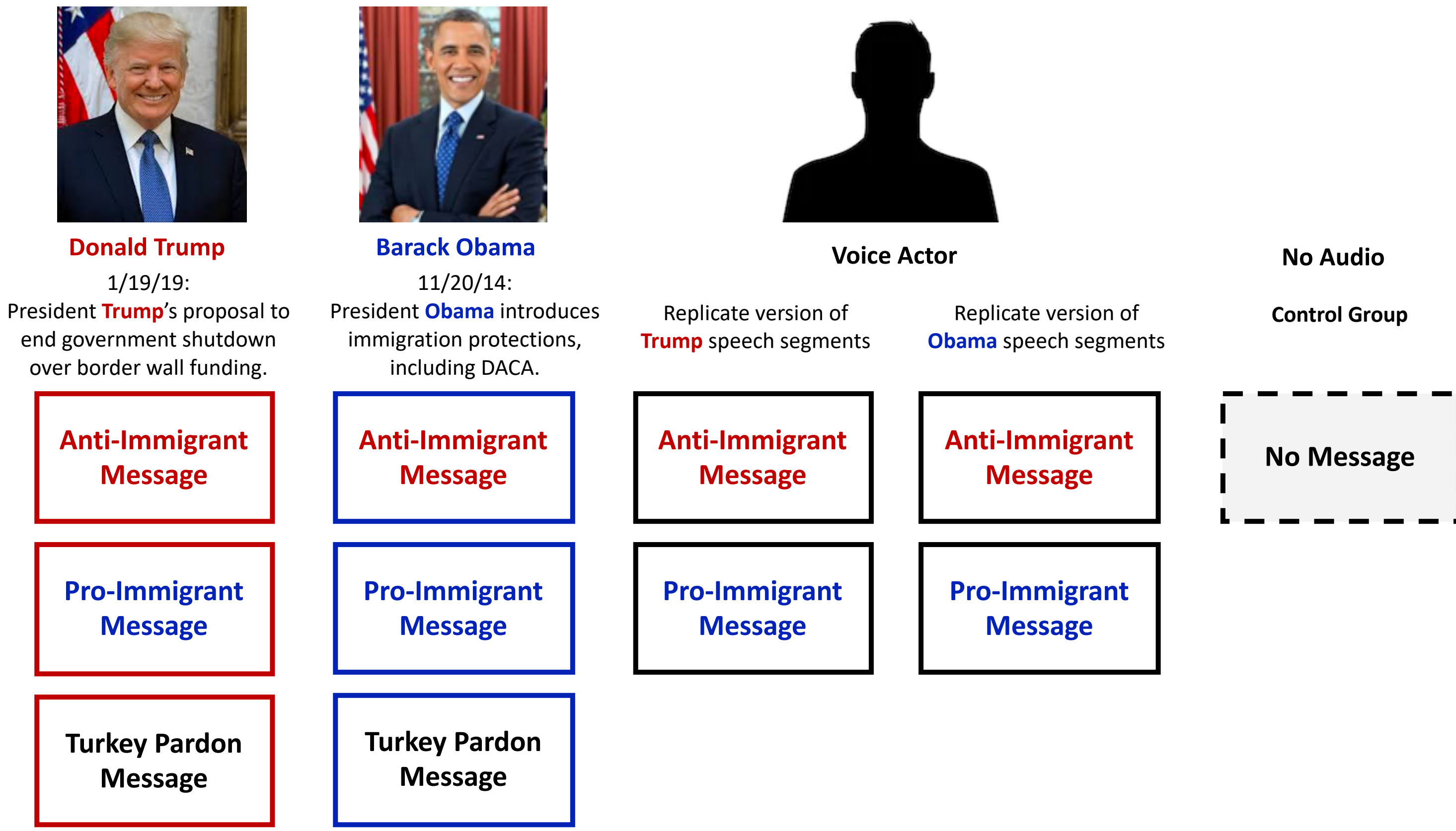
Table 3: Impact of Source within Fixed Message

			Republicans *ln(Probability Anti)*					**Democrats** *ln(Probability Pro)*				
Message	**Source**	**Control**	β	(S.E.)	%Diff.	*P(Dist.)*	N	β	(S.E.)	%Diff.	*P(Dist.)*	N
Anti	Trump	Actor	0.003	(0.006)	0.91%	*0.665*	1069	0.007	(0.005)	1.64%	*0.363*	1219
Anti	Obama	Actor	0.006	(0.006)	1.72%	*0.567*	1033	-0.014***	(0.005)	-3.43%	*0.001*	1216
Pro	Trump	Actor	-0.014**	(0.006)	-3.92%	*0.073*	1091	-0.007	(0.005)	-1.56%	*0.329*	1274
Pro	Obama	Actor	0.007	(0.006)	2.04%	*0.275*	1101	-0.004	(0.005)	-0.95%	*0.141*	1327

			Republicans *ln(Probability Convinced)*					**Democrats** *ln(Probability Convinced)*				
Message	**Source**	**Control**	β	(S.E.)	%Diff.	*P(Dist.)*	N	β	(S.E.)	%Diff.	*P(Dist.)*	N
Anti	Trump	Actor	0.003	(0.006)	0.91%	*0.665*	1069	-0.007	(0.006)	-3.33%	*0.385*	1219
Anti	Obama	Actor	0.006	(0.006)	1.72%	*0.567*	1033	0.020***	(0.006)	8.83%	*0.001*	1216
Pro	Trump	Actor	0.017***	(0.006)	5.58%	*0.075*	1091	-0.007	(0.005)	-1.56%	*0.329*	1274
Pro	Obama	Actor	-0.009	(0.007)	-2.91%	*0.505*	1101	-0.004	(0.005)	-0.95%	*0.141*	1327

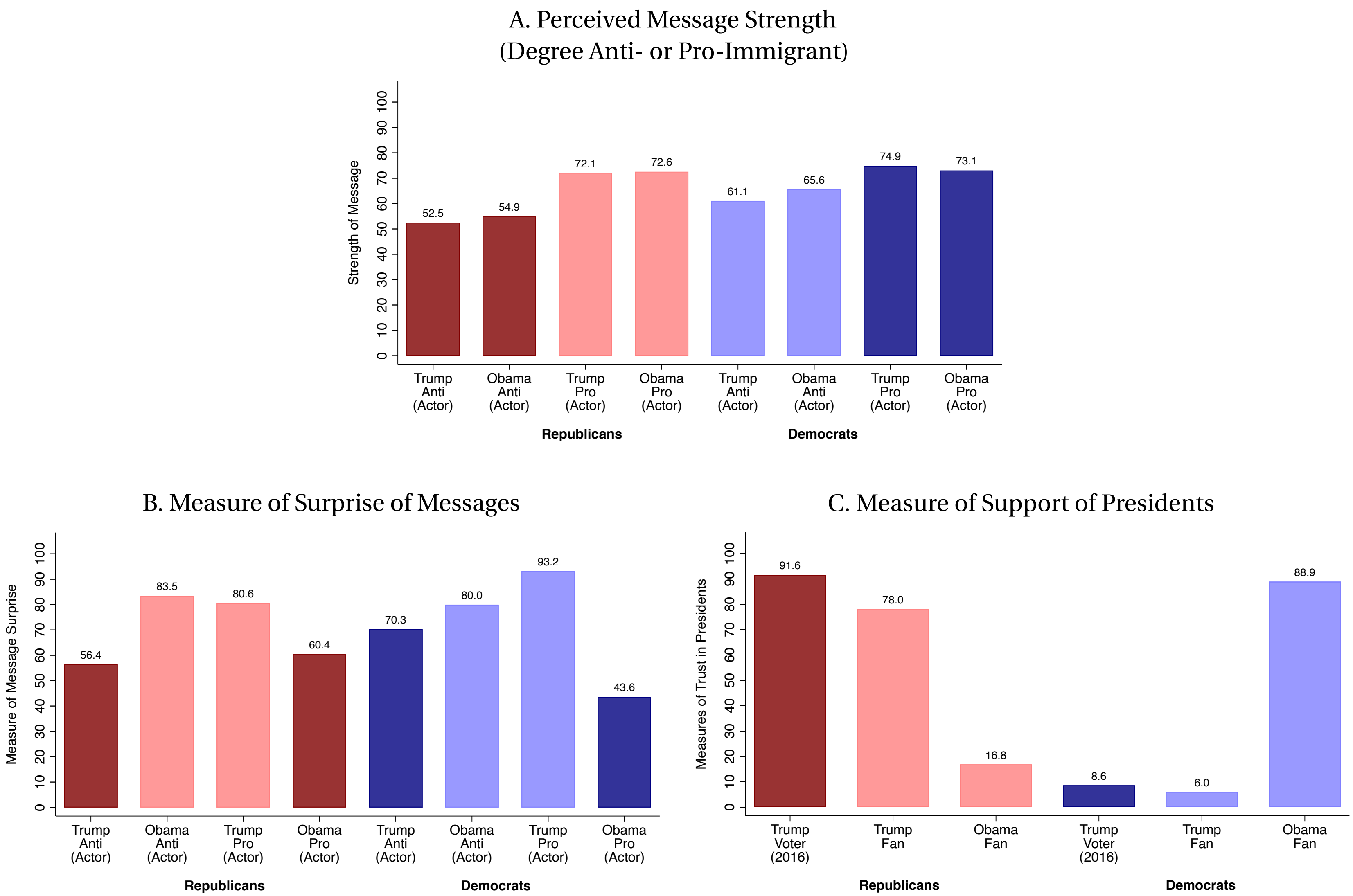
Notes: *p<0.1, **p<0.05, ***p<0.01. This table presents the estimates of the different source treatments, relative to the group that heard the same message from the actor. Outcomes in the top panel are constructed as $ln(P(Outcome)+1)$, where the outcome is the index probability anti-immigrant (pro-immigrant) for Republicans (Democrats). All regressions additionally include covariates selected by the double lasso procedure to improve precision (See Appendix A4). In the bottom panel, outcomes are the index probability anti-immigrant or pro-immigrant to match the message type, changing the outcome to $ln(P(Convinced)+1)$. "% Diff." calculates the implied change in outcome *probability*, or the untransformed probability index, due to treatment, relative to the mean for the control group. "$P(Dist.)$" is the p-value from a Kolmogorov-Smirnov test of equality of the distributions of the control and treatment groups.

Figure 1: Experiment Design



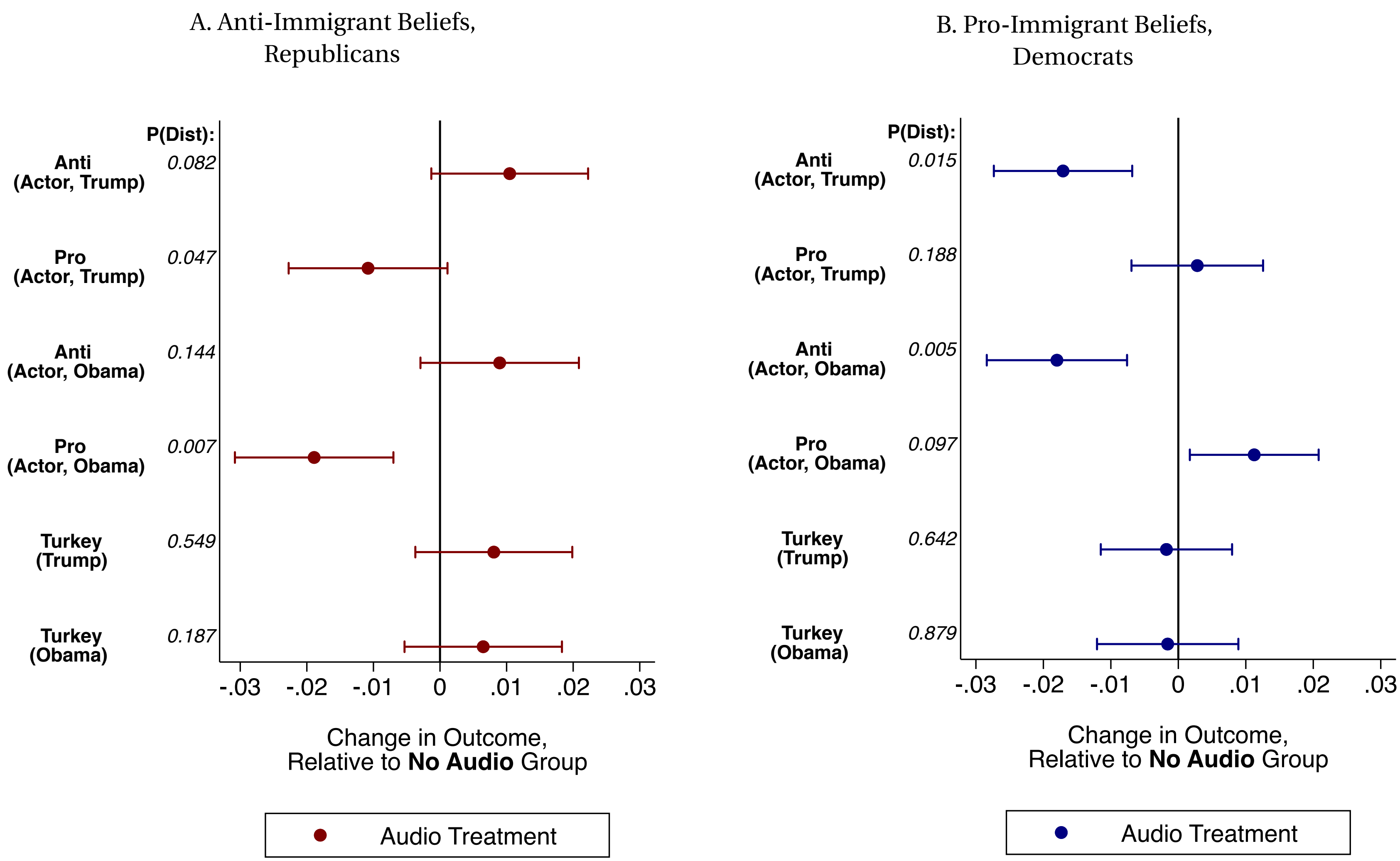
Notes: This figure depicts the treatment arms of the experiment. The sample is stratified by party, Republican or Democrat, and then within party all participants are randomized into the 11 treatment arms shown above.

Figure 2: Measures of Message Similarity, Surprise of Messages, and Source Favorability



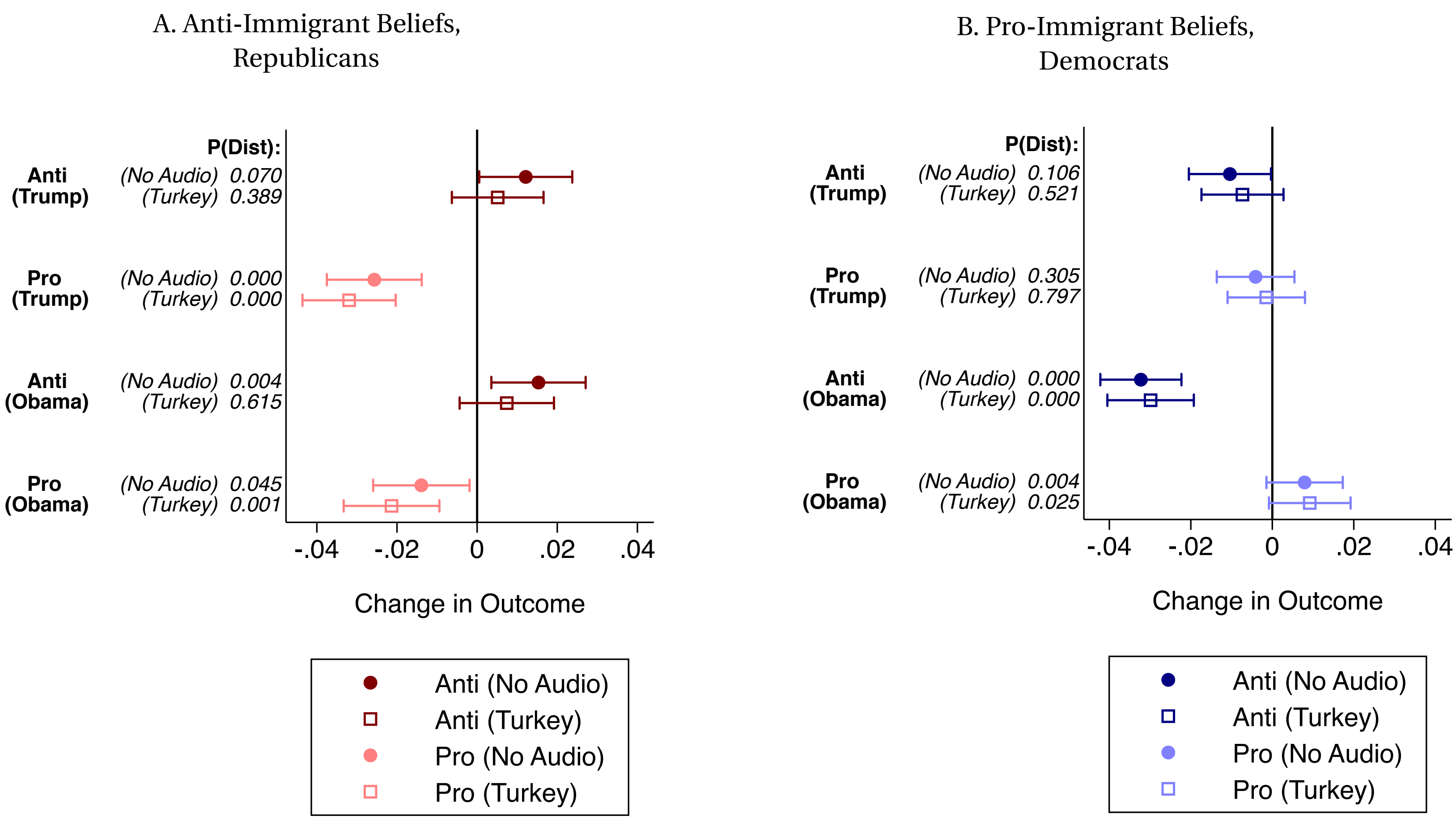
Notes: This figure compiles participant responses to a question that asks participants how pro- or anti- immigrant they think the speech was after treatment. Panel A plots the degree pro-immigrant for pro-immigrant speeches and the degree anti-immigrant for anti-immigrant speeches. The actor speeches shown in this figure are not colored by a stated source and can be compared across presidents to measure the "strength" of different speeches. Panel B contains responses from a multiple choice question asked post-intervention about which president the participant thought was the source of the speech, from a list of the four most recent presidents (and an "Other" option). This plot shows the share of participants who guessed the *incorrect* president for each actor version of the treatments. This measure is a share of participants who would be surprised by the source of the speech. Panel C includes pre-treatment responses from questions asking whether participants voted for Trump in 2016 and whether participants are a fan of Trump and Obama.

Figure 3: Impact of Baseline Messages



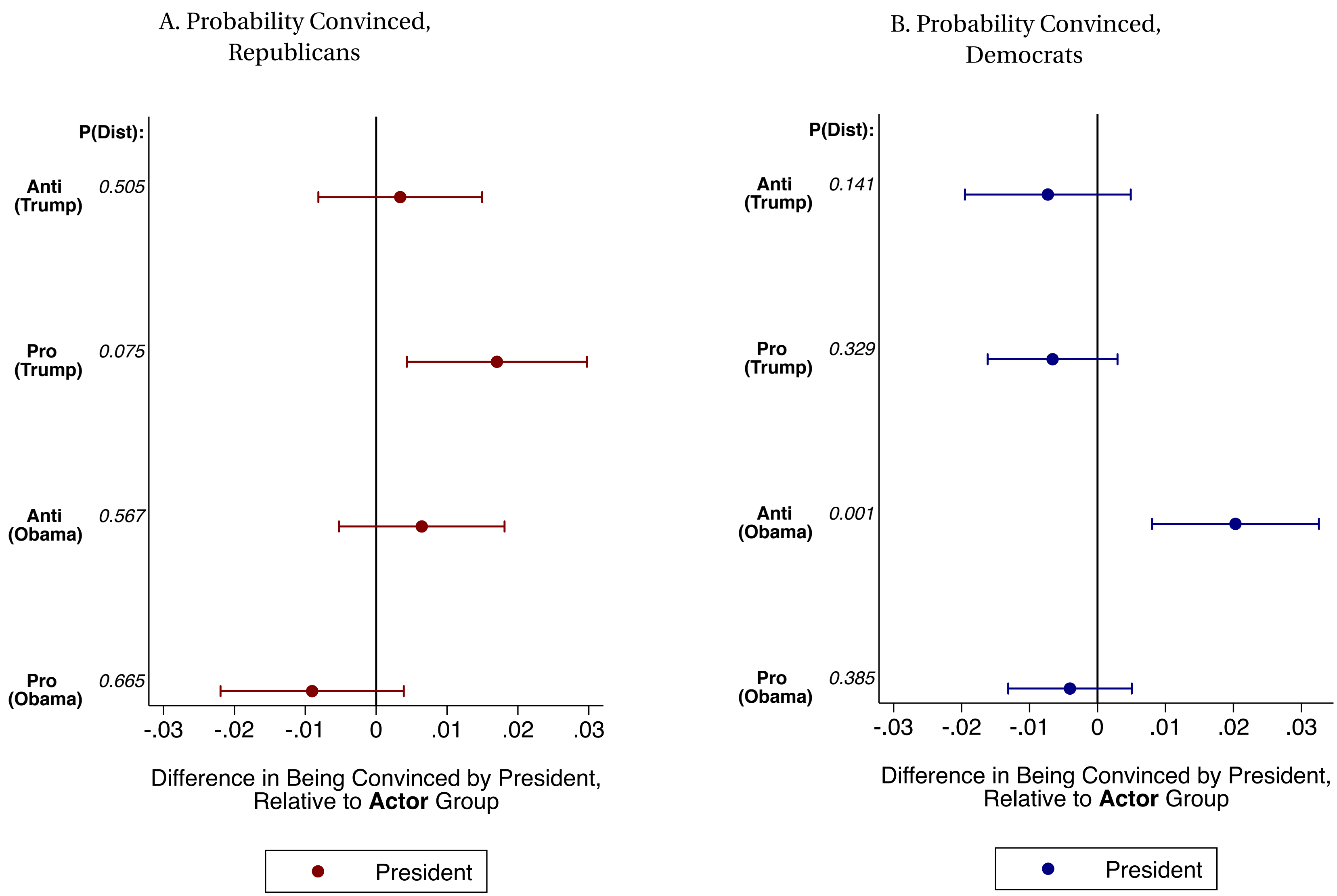
Notes: This figure plots the coefficient estimates and 95% confidence intervals of the treatments sourced from the actor and the president turkey pardon speeches. All estimates are equivalent to those in Table 2. Each treatment is compared to the no audio control group. Outcomes are measured as $ln(P(Outcome)+1)$, where the outcome is the anti-immigrant (pro-immigrant) index for Republicans (Democrats). All regressions additionally include covariates selected by the double lasso procedure to improve precision (See Appendix A4). "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure 4: Impact of President Messages



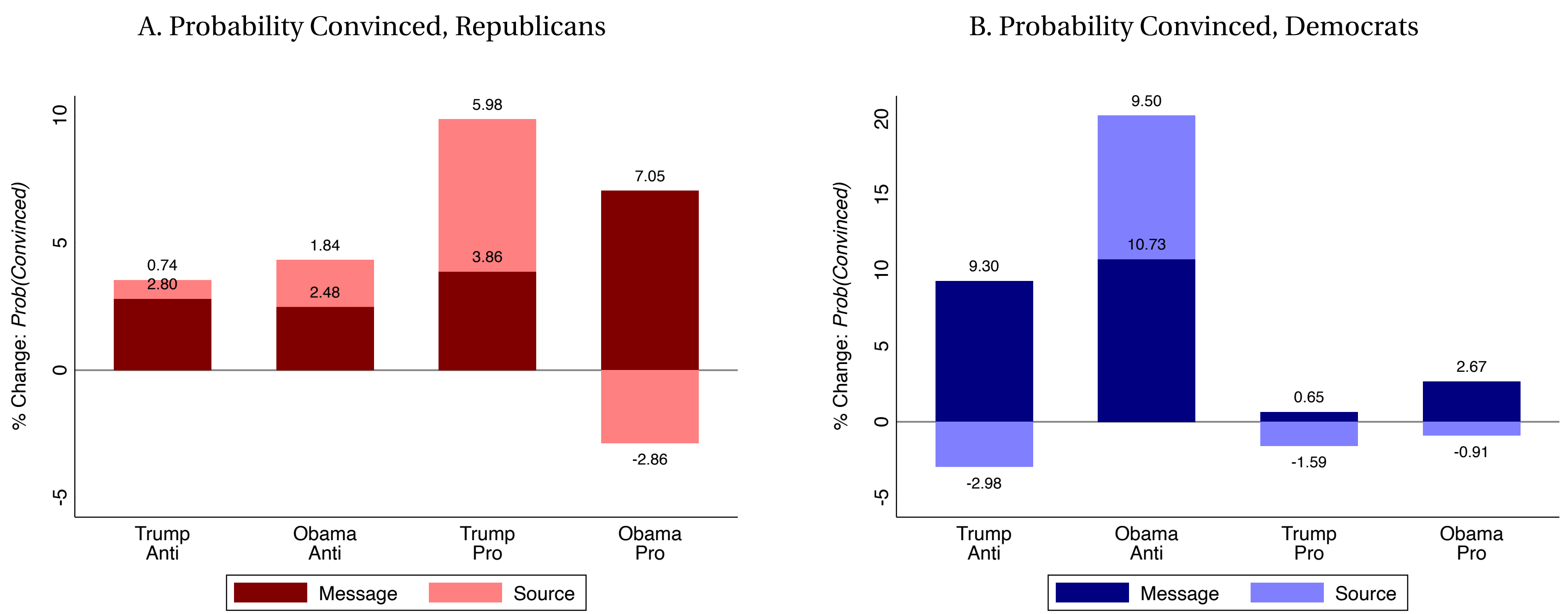
Notes: This figure plots the coefficient estimates and 95% confidence intervals of the treatments sourced from the presidents. All estimates are equivalent to those in Table 2. Each treatment is compared to either the no audio control group or the turkey pardon message. Outcomes are measured as $ln(P(Outcome)+1)$, where the outcome is the anti-immigrant (pro-immigrant) index for Republicans (Democrats). All regressions additionally include covariates selected by the double lasso procedure to improve precision (See Appendix A4). "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure 5: Difference in Probability of Being Convinced, Across Sources within Message



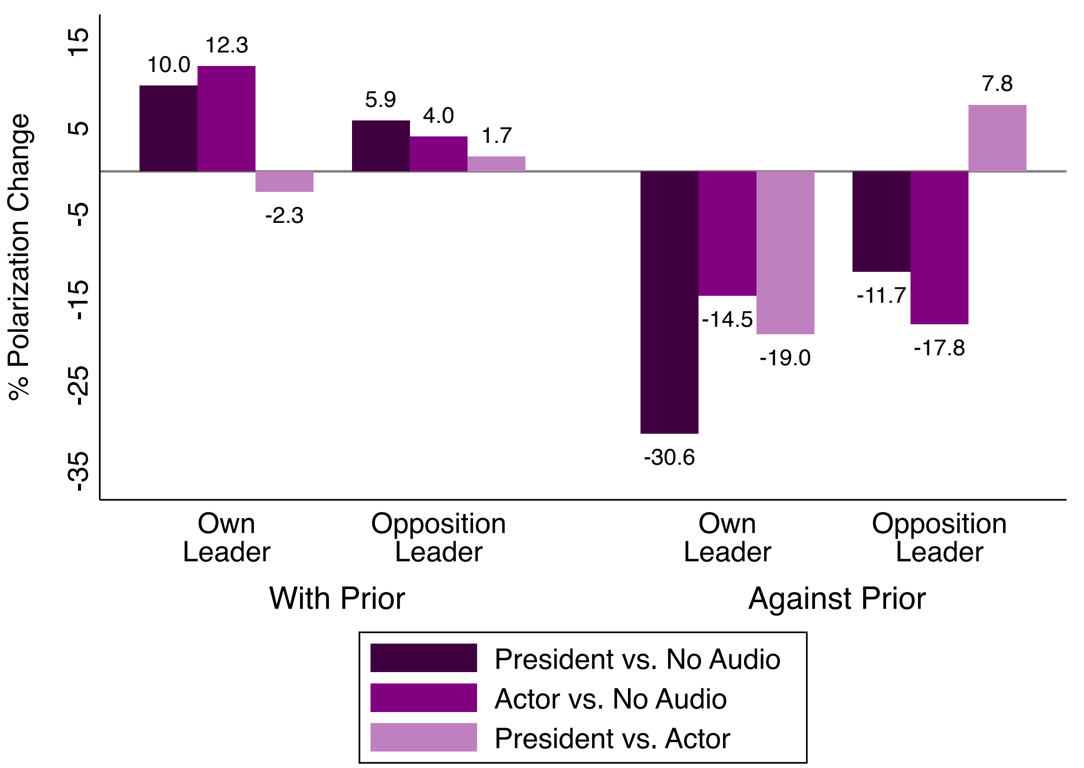
Notes: This figure plots the coefficient estimates and 95% confidence intervals of the treatments sourced from the presidents, relative to the replicate versions recorded by the actor. All estimates are equivalent to those in Table 3. Outcomes are measured as $ln(P(Convinced)+1)$, where the outcome is the immigrant index associated with the direction of the message, ex. anti-immigrant index for an anti-immigrant speech. All regressions additionally include covariates selected by the double lasso procedure to improve precision (See Appendix A4). "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure 6: Decomposition of Source and Message Effects



Notes: This plot decomposes the change in the untransformed *probability convinced* into source and message effects. Probability convinced is measured in the direction of the message, ex. the anti-immigrant (pro-immigrant) index for an anti-immigrant (pro-immigrant) message. Changes are measured as percent increases or decreases relative to the mean of the outcome in the no audio control group. Message effects are the incremental change between the actor speech and the no audio group. Source effects are the incremental change between the president speech and the replicate actor version of the speech. All estimates are determined from a regression that includes the president group, actor group, and no audio group, with indicators for the two speech treatments. These regressions are also adjusted for the covariates selected by the double lasso procedure and described in Appendix A4.

Figure 7: Change in Polarization in Counterfactual Scenarios



Notes: This plot corresponds to the estimates in Table A2. The estimates represent percent changes in the distance in the anti-immigrant index probability between Republicans and Democrats in the treatment versus control group. Specifically, the corresponding regression is $ln(P(Anti)+1) = \alpha + \beta_1 President + \beta_2 Republican + \beta_3 President * Republican + \gamma X + \varepsilon$, where the interaction coefficient β_3 represents a change in polarization. The covariates X correspond to those in Appendix A4 for both party groups. A positive estimate is an increase in polarization. Estimates are plotted as percent changes in the party gap in *probability anti* relative the average party gap in the control group. The control group in the plot above is either the no audio group or the actor version of a speech. "Own Leaders" are treatments that are Trump for Republicans and Obama for Democrats, while "Opposition Leaders" are the converse. "With Prior" messages are anti-immigrant speeches for Republicans and pro-immigrant speeches for Democrats, and "Against Prior" messages are the converse.

ONLINE APPENDIX

A1 Opening Paragraphs of Immigration Speech Treatments

Trump Anti-Immigrant Speech "We believe in a safe and lawful system of immigration, one that upholds our laws, our traditions and our most cherished values. Unfortunately, our immigration system has been badly broken for a very long time. Over the decades, many presidents and many lawmakers have come and gone, and no real progress has been made on immigration. We are now living with the consequences – and they are tragic – brought about by decades of political stalemate, partisan gridlock, and national neglect. Illegal immigration reduces wages and strains public services. The lack of border control provides a gateway, and a very wide and open gateway, for criminals and gang members to enter the United States. I want this to end; it's got to end now. These are not talking points. These are the heartbreaking realities that are hurting innocent, precious human beings every single day on both sides of the border."

Obama Anti-Immigrant Speech "Today, our immigration system is broken, and everybody knows it. Families who enter our country the right way and play by the rules watch others flout the rules. Business owners who offer their workers good wages and benefits see the competition exploit undocumented immigrants by paying them far less. All of us take offense to anyone who reaps the rewards of living in America without taking on the responsibilities of living in America. Millions of us, myself included, go back generations in this country, with ancestors who put in the painstaking work to become citizens. So we don't like the notion that anyone might get a free pass to American citizenship. I know that some worry immigration will change the very fabric of who we are, or take our jobs, or stick it to middle-class families at a time when they already feel like they've gotten the raw end of the deal for over a decade. I hear these concerns."

Trump Pro-Immigrant Speech "Just a short time ago, I had the honor of presiding over the swearing in of five new great American citizens. It was a beautiful ceremony and a moving reminder of our nation's proud history of welcoming immigrants from all over the world into our national family. I told them that the beauty and majesty of citizenship is that it draws no distinctions of race or class or faith or gender or background. All Americans, whether first generation or tenth generation, are bound together in love and loyalty, friendship and affection. We're all equal. We are one team and one people proudly saluting one great American flag. Unfortunately, our immigration system has been badly broken for a very long time. Over the decades, many presidents and many lawmakers have come and gone, and no real progress has been made on immigration. The good news is these problems can all be solved, but only if we have the political courage to do what is just and what is right. Both sides in Washington must simply come together, listen to each other, put down their armor, build trust, reach across the aisle and find solutions."

Obama Pro-Immigrant Speech "For more than 200 years, our tradition of welcoming immigrants from around the world has given us a tremendous advantage over other nations. It's kept us youthful, dynamic, and entrepreneurial. It has shaped our character as a people with limitless possibilities – people not trapped by our past, but able to remake ourselves as we choose. But today, our immigration system is broken, and everybody knows it. It's been this way for decades. And for decades, we haven't done much about it. When I took office, I committed to fixing this broken immigration system. We need more than politics as usual when it comes to immigration; we need reasoned, thoughtful, compassionate debate that focuses on our hopes, not our fears."

Full speech transcripts are viewable in Supplemental Materials S2.

A2 Appendix Tables & Figures

Table A1: Perception of Treatment Content and Source

		Republicans			Democrats		
Message	**Source**	Mean	(S.D.)	N	Mean	(S.D.)	N
Perception Anti-Immigrant							
Anti-Immigrant	Trump	52.15	(29.40)	548	75.03	(27.01)	606
Anti-Immigrant	Obama	49.19	(27.44)	506	41.20	(26.77)	622
Anti-Immigrant	Actor, Trump	52.46	(27.24)	521	61.06	(28.87)	613
Anti-Immigrant	Actor, Obama	54.91	(28.44)	527	65.60	(29.25)	594
Perception Pro-Immigrant							
Pro-Immigrant	Trump	74.58	(27.04)	551	53.69	(28.95)	641
Pro-Immigrant	Obama	77.79	(26.92)	525	82.49	(24.73)	675
Pro-Immigrant	Actor, Trump	72.09	(26.08)	540	74.92	(27.20)	633
Pro-Immigrant	Actor, Obama	72.55	(26.99)	576	73.06	(27.76)	652
Perception Correct President							
Anti-Immigrant	Trump	98.91	(10.42)	548	97.36	(16.05)	606
Anti-Immigrant	Obama	97.04	(16.98)	506	98.88	(10.56)	622
Anti-Immigrant	Actor, Trump	43.57	(49.63)	521	29.69	(45.73)	613
Anti-Immigrant	Actor, Obama	16.51	(37.16)	527	20.03	(40.06)	594
Pro-Immigrant	Trump	97.64	(15.19)	551	95.01	(21.80)	641
Pro-Immigrant	Obama	97.91	(14.34)	525	98.96	(10.14)	675
Pro-Immigrant	Actor, Trump	19.44	(39.61)	540	6.79	(25.18)	633
Pro-Immigrant	Actor, Obama	39.58	(48.95)	576	56.44	(49.62)	652
Turkey	Trump	98.49	(12.19)	531	96.89	(17.36)	676
Turkey	Obama	94.81	(22.20)	501	97.26	(16.34)	620

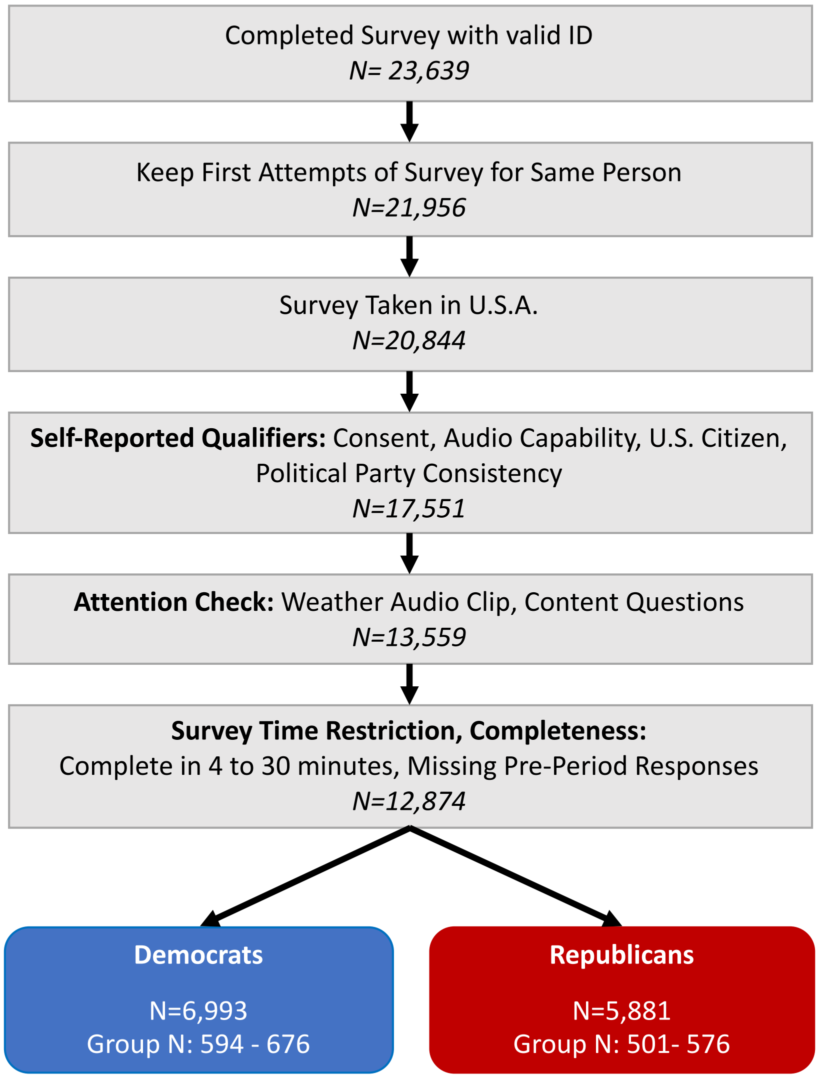
Notes: This table shows the results of the post-treatment questions asking participants the degree that they perceive treatment to be anti-immigrant or pro-immigrant, as well as their guess of which president originally gave the speech. The latter results are presented as the share of participants in a group who guess the *correct* president for a particular treatment.

Table A2: Change in Polarization in Counterfactual Scenarios

			Polarization Change *ln(Probability Anti)* *Diff.: Republican-Democrat*			**Baseline** *Probability Anti*		
Message	**Source**	**Control**	β	(S.E.)	%Diff.	Republican Mean	Democrat Mean	N
Own Leader Messages								
With Prior	Actor	No Audio	0.027***	(0.009)	11.38%	0.579	0.271	2389
With Prior	President	No Audio	0.022**	(0.009)	9.27%	0.579	0.271	2439
With Prior	President	Actor	-0.006	(0.008)	-2.25%	0.597	0.251	2396
Opposition Leader Messages								
With Prior	Actor	No Audio	0.009	(0.009)	3.75%	0.579	0.271	2376
With Prior	President	No Audio	0.013	(0.009)	5.46%	0.579	0.271	2363
With Prior	President	Actor	0.004	(0.008)	1.61%	0.597	0.264	2307
Own Leader Messages								
Against Prior	Actor	No Audio	-0.032***	(0.009)	-14.49%	0.579	0.271	2350
Against Prior	President	No Audio	-0.067***	(0.009)	-30.60%	0.579	0.271	2389
Against Prior	President	Actor	-0.035***	(0.009)	-18.96%	0.560	0.298	2307
Opposition Leader Messages								
Against Prior	Actor	No Audio	-0.039***	(0.009)	-17.80%	0.579	0.271	2405
Against Prior	President	No Audio	-0.025***	(0.009)	-11.68%	0.579	0.271	2347
Against Prior	President	Actor	0.015*	(0.009)	7.75%	0.550	0.281	2320

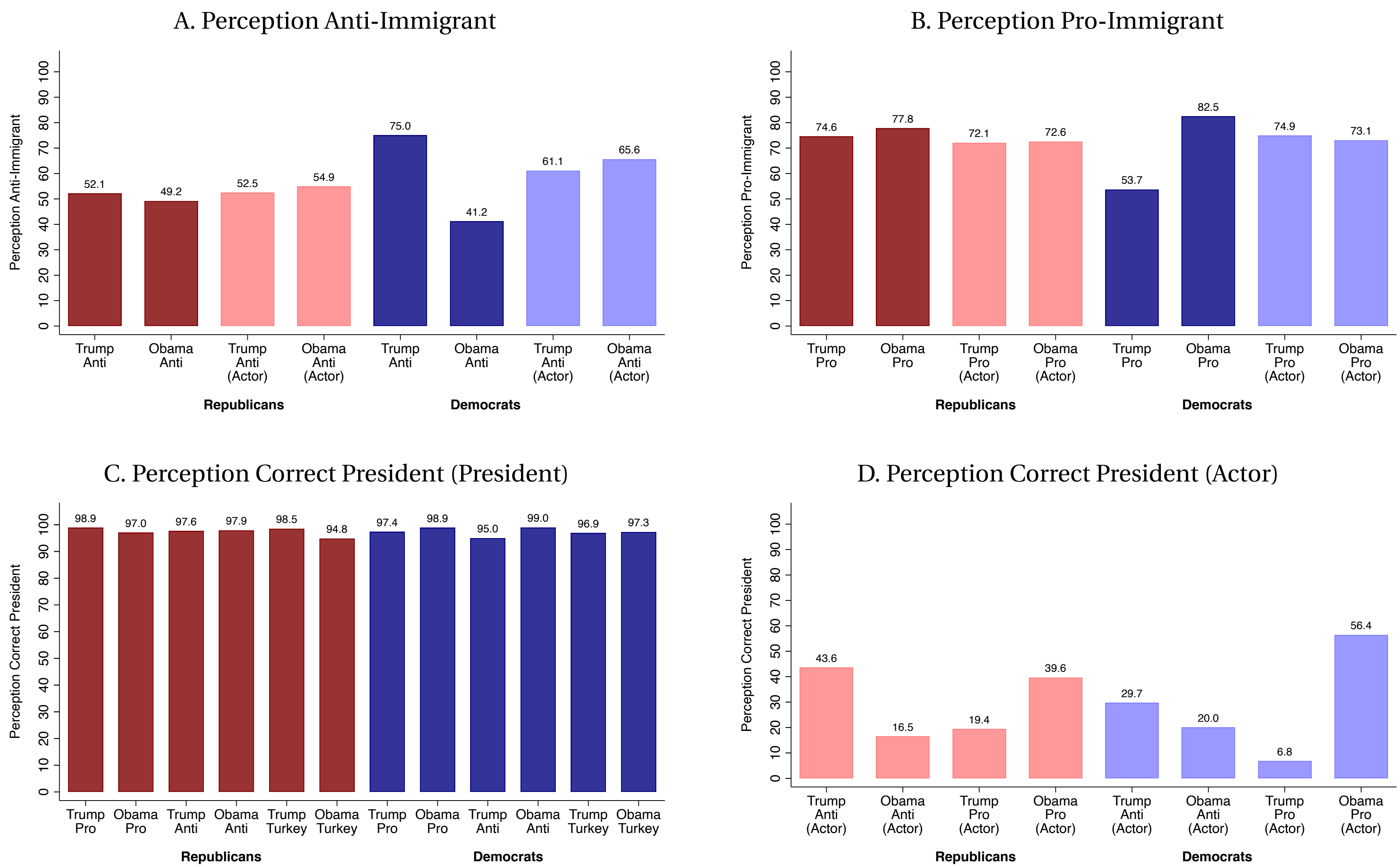
Notes: *p<0.1, **p<0.05, ***p<0.01. This table compares the change in distance between Republicans and Democrats from a treatment, relative to a control group, where each regression includes both Democrats and Republicans. The treatment conditions depend on party; for example "With Prior", "President", corresponds to Republicans (Democrats) hearing an anti-immigrant (pro-immigrant) message from Trump (Obama). Likewise, actor control groups correspond to replicate messages for a given treatment. Outcomes are constructed as $ln(P(Anti)+1)$, where the outcome is the index probability anti-immigrant. Specifically, the corresponding regression is $ln(P(Anti)+1) = \alpha + \beta_1 President + \beta_2 Republican + \beta_3 President * Republican + \gamma X + \varepsilon$, where the interaction coefficient β_3 represents a change in polarization. A positive β corresponds to an increase in polarization. All regressions additionally include covariates selected by the double lasso procedure to improve precision for both party groups (See Appendix A4)."% Diff." calculates the implied change in the difference of outcome *probability* across party, or the untransformed probability index, due to treatment, relative to the baseline difference across party in the control group.

Figure A1: Sample Restrictions



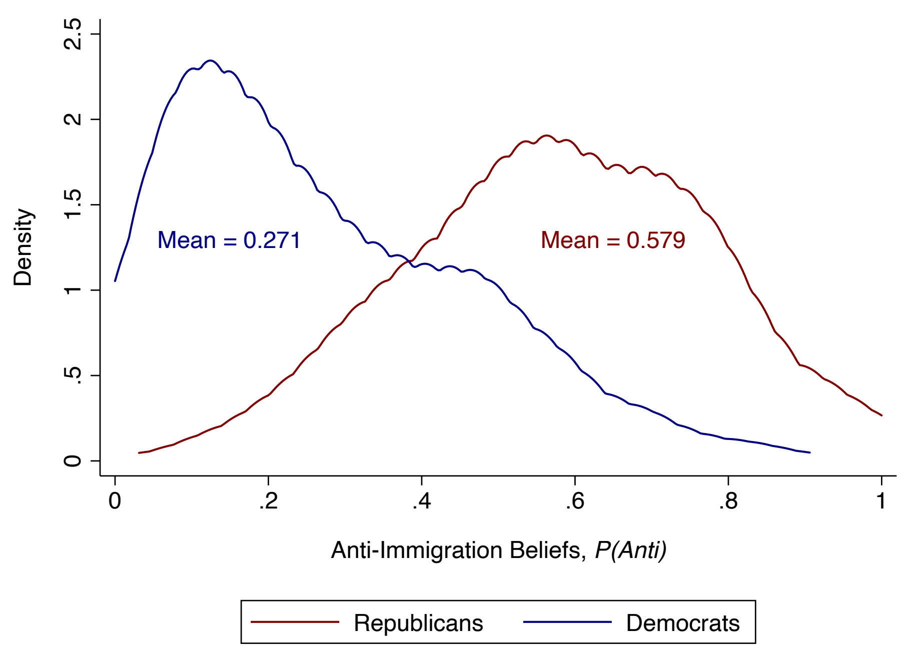
Notes: This figure displays the sample restrictions used to construct the study data set. The first restriction identifies individuals who took the survey more than once using IP addresses and retains only the first attempt. Next, we geocode the survey taker location and remove respondents who took the survey outside of the U.S. The third restriction cuts people who did not consent to the study, stated that their device did not have audio capability, were not U.S. citizens, or were recruited to the survey as a member of particular political party and then stated that their party affiliation was the opposite party within the survey. Next, we include two sets of attention checks to screen individuals for quality. First, we have all participants listen to an audio clip of a weather forecast and answer comprehension related questions. Second, for participants who listen to a treatment audio clip (immigration or turkey pardon speech) we also ask comprehension related questions. Lastly, we exclude individuals who take the survey exceptionally fast or slow, restricting the time in survey to 6 to 25 minutes.

Figure A2: Perception Treatment Content and Source



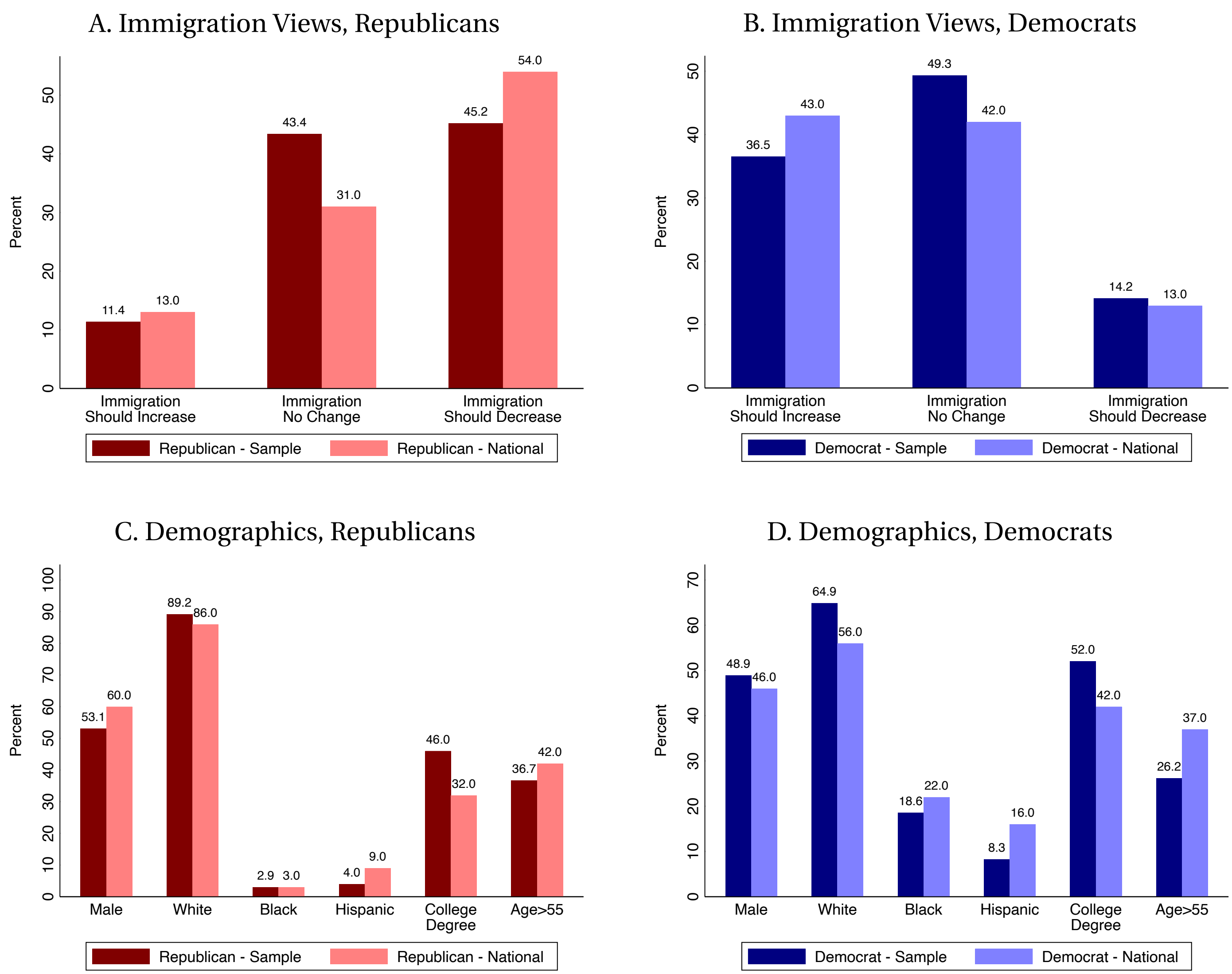
Notes: This set of plots correspond to the estimates in Table A1. The bars correspond to average shares within groups.

Figure A3: Anti-Immigration Beliefs in No Audio Control Group



Notes: This plot shows the baseline distribution of the Anti-Immigration index or $P(Anti-Immigration)$ in the no audio control group. This index is composed of 16 questions about immigration beliefs asked in the second part of the survey.

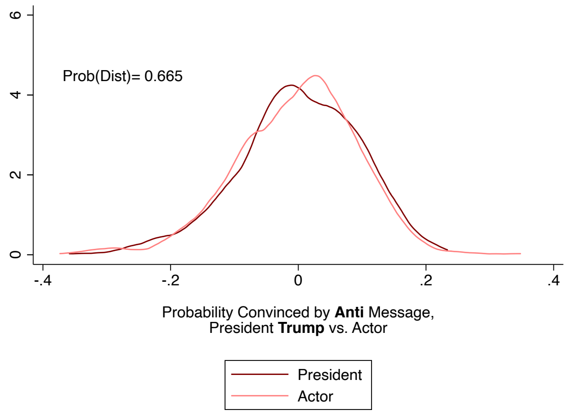
Figure A4: Sample Comparison to Party Demographics



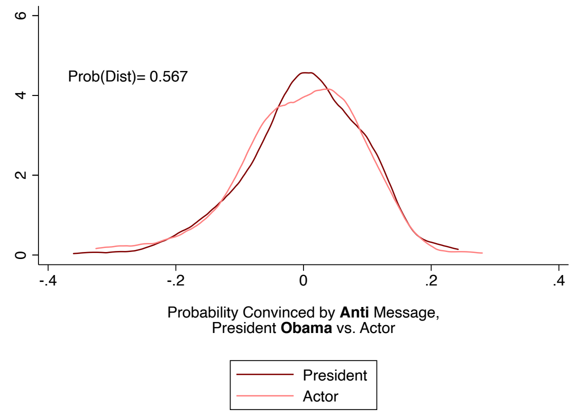
Notes: This set of plots compares characteristics of the study sample to national data for political parties. The top panel compares pre-treatment survey responses to a Gallup Survey question about whether immigration should be increased, decreased or stay the same (Jones, 2019). The national comparison is compiled from survey evidence and data on eligible voters from the U.S. Census, and Pew Research (File, 2018; Igielnik and Budiman, 2020; Doherty et al., 2020).

Figure A5: Difference in Probability of Being Convinced, Republicans

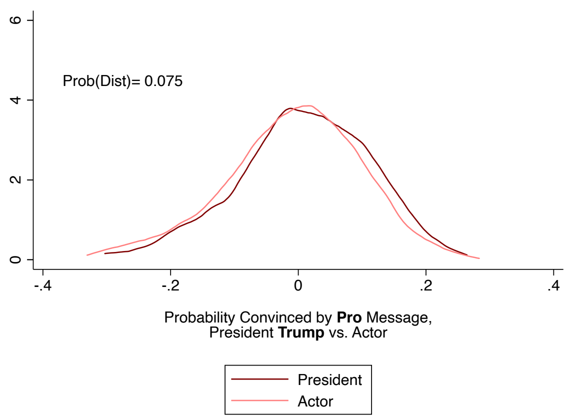
A. Anti-Immigrant Message, Trump



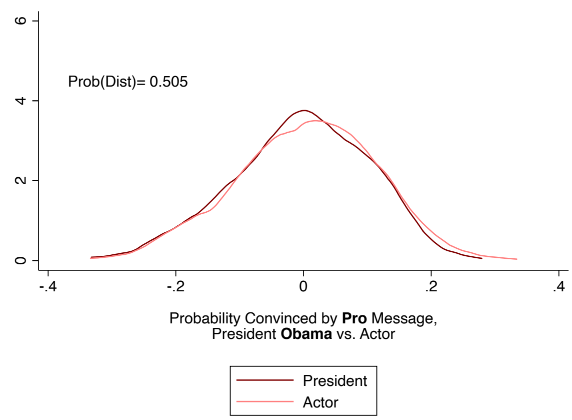
B. Anti-Immigrant Message, Obama



C. Pro-Immigrant Message, Trump

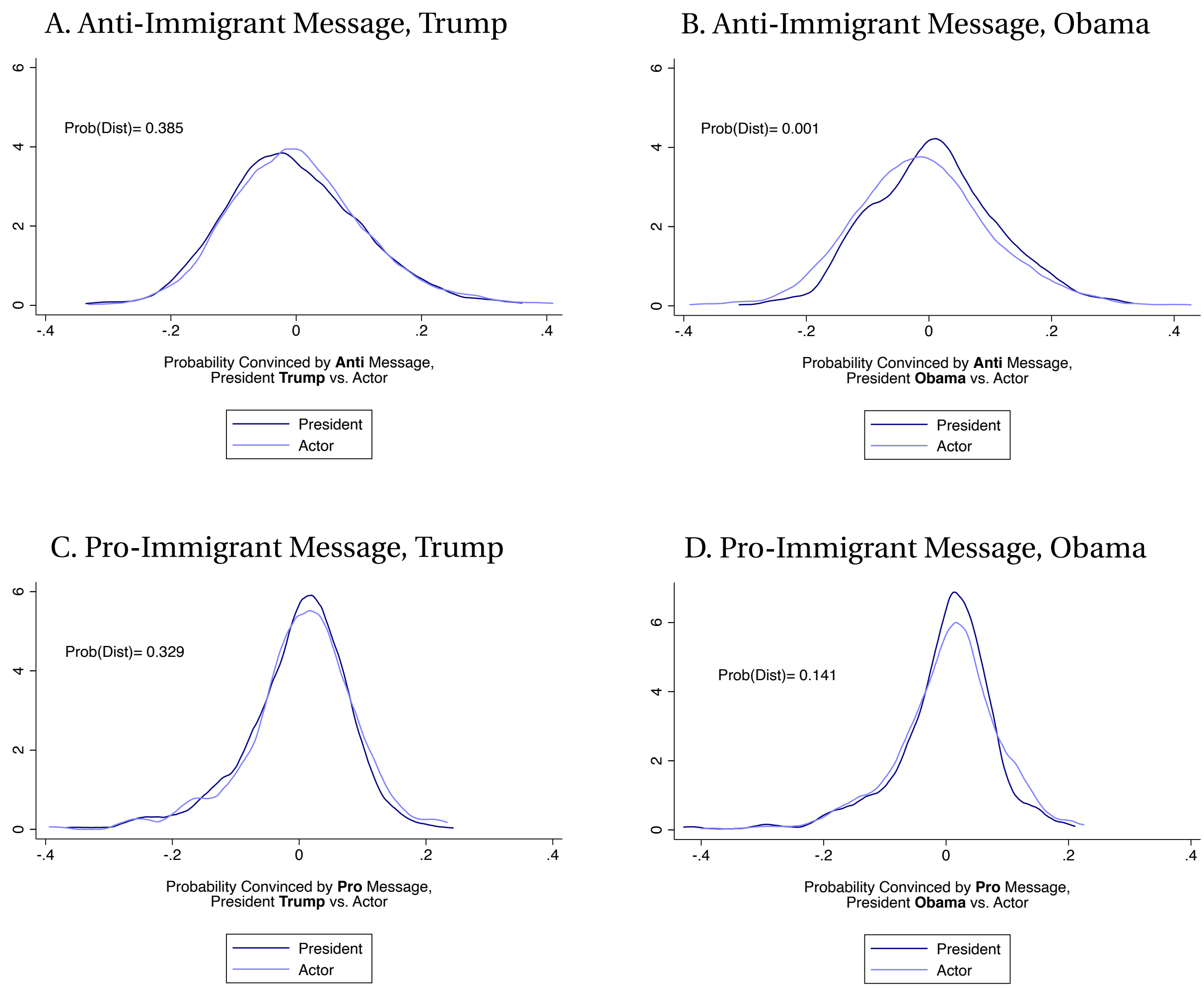


D. Pro-Immigrant Message, Obama



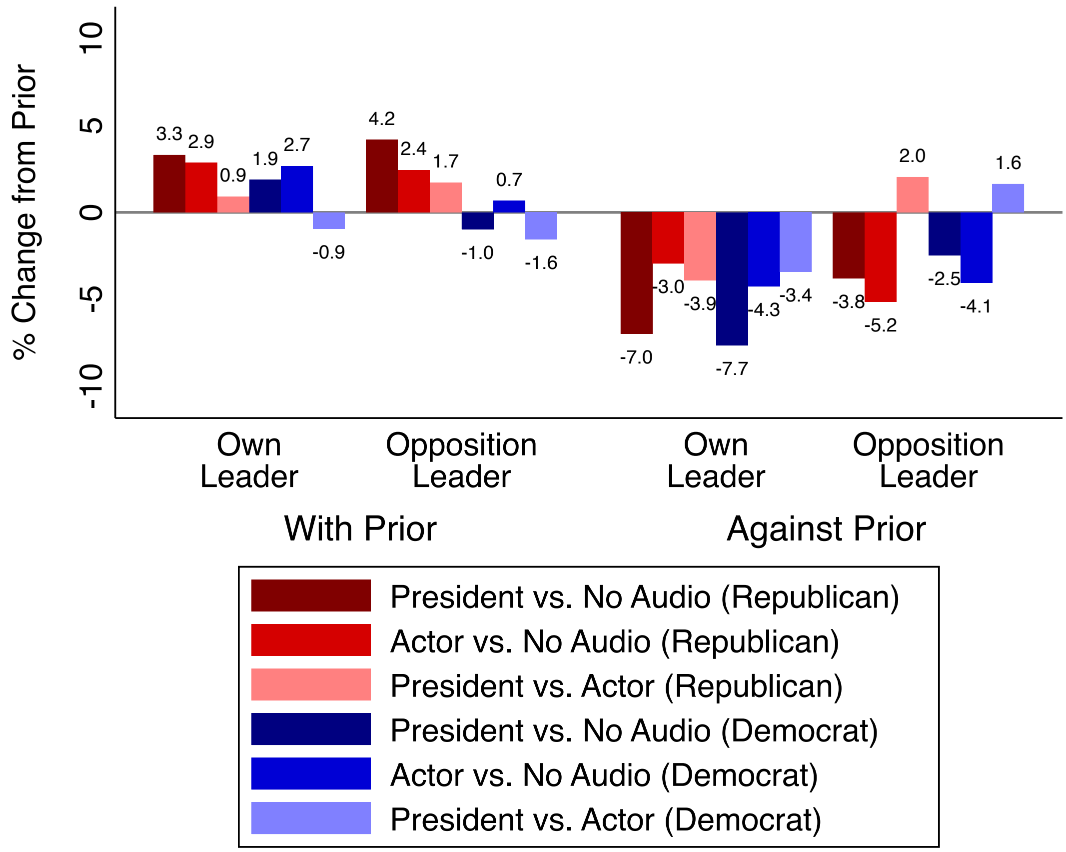
Notes: This figure plots the distribution of the outcomes for the president treatments, relative to the replicate versions recorded by the actor. Outcomes are measured as $ln(P(Convinced)+1)$, where the outcome is the immigrant index associated with the direction of the message, ex. anti-immigrant (pro-immigrant) index for an anti-immigrant (pro-immigrant) speech. Outcomes are first residualized by the set of control variables from the double lasso procedure described in Appendix A4 before being plotted. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A6: Difference in Probability of Being Convinced, Democrats



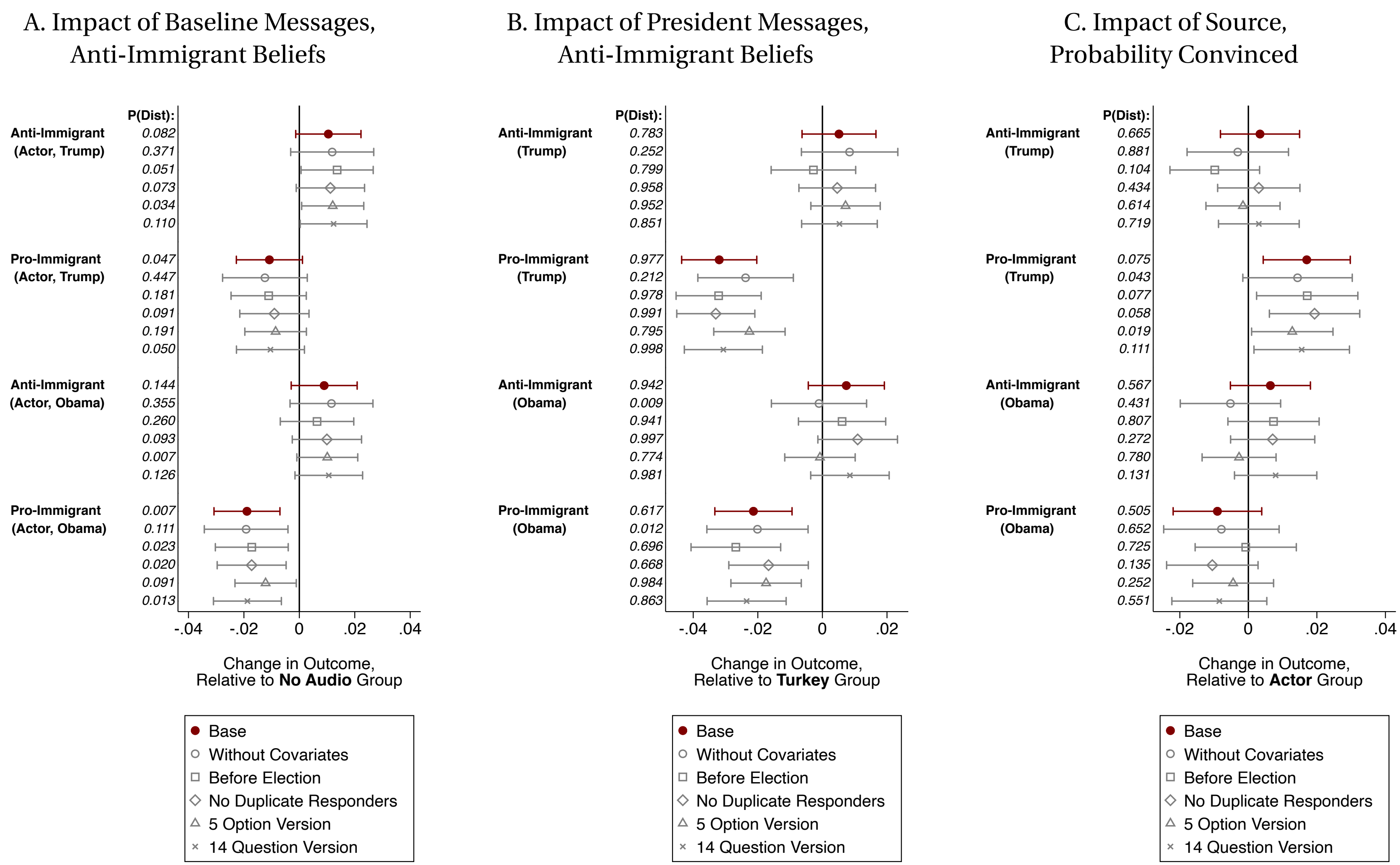
Notes: This figure plots the distribution of the outcomes for the president treatments, relative to the replicate versions recorded by the actor. Outcomes are measured as $ln(P(Convinced)+1)$, where the outcome is the immigrant index associated with the direction of the message, ex. anti-immigrant index for an anti-immigrant speech. Outcomes are first residualized by the set of control variables from the double lasso procedure described in Appendix A4 before being plotted. “P(Dist)” is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A7: Change from Prior in Counterfactual Scenarios, By Party



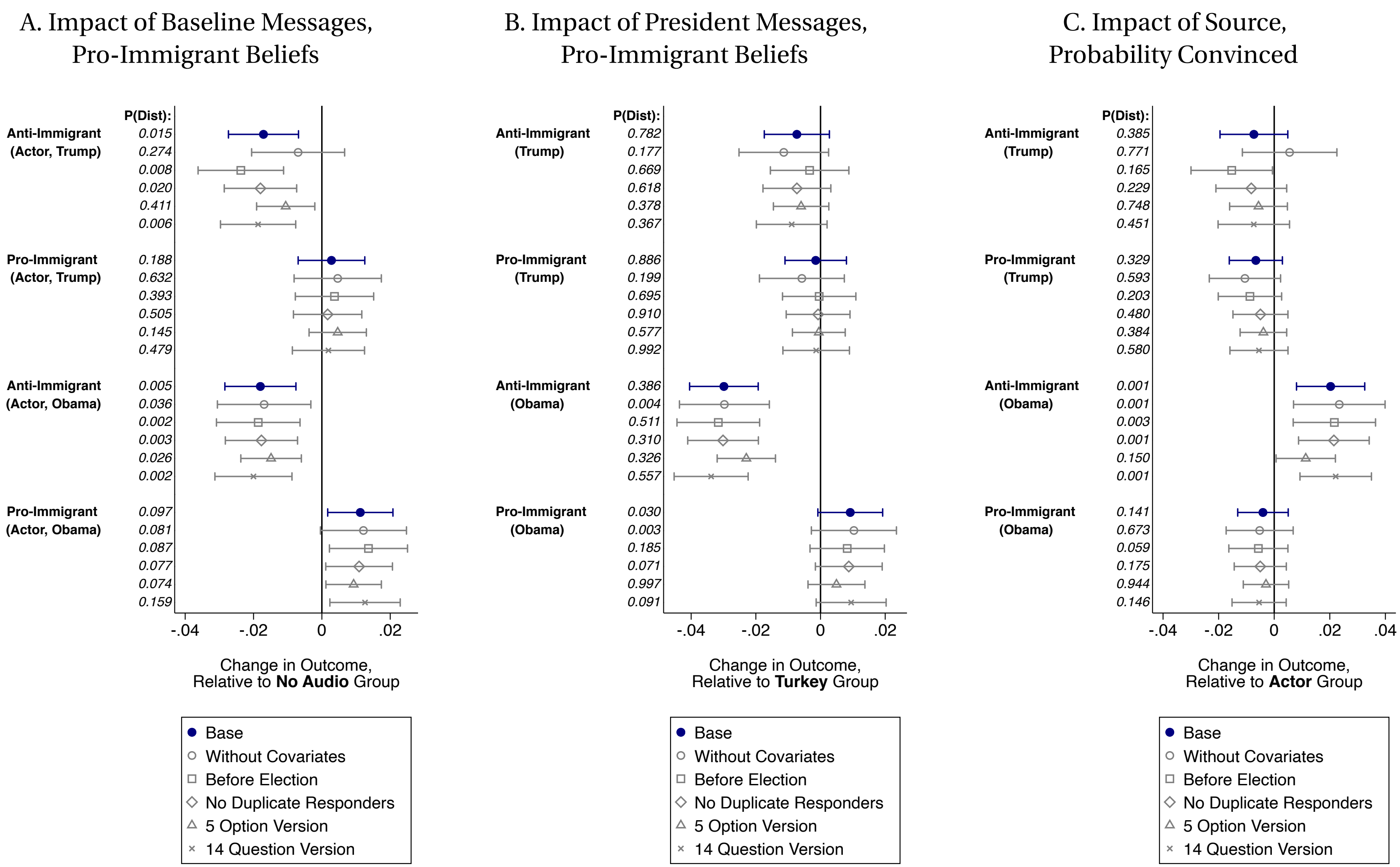
Notes: This plot corresponds to the estimates in Table A2. The estimates represent percent changes in the anti-immigrant index probability for Republicans and changes in the pro-immigrant index probability for Democrats. The control group in the plot above is either the no audio group or the actor version of a speech, as defined in the legend. Specifically, the corresponding regressions are $ln(P(Outcome)+1) = \alpha + \beta_1 Treat + \gamma X + \varepsilon$, where the coefficient β_1 represents a change in the direction of the party prior. The covariates X correspond to those in Appendix A4. A positive estimate is an increase in polarization. "Own Leaders" are treatments that are Trump for Republicans and Obama for Democrats, while "Opposition Leaders" are the converse. "With Prior" messages are anti-immigrant speeches for Republicans and pro-immigrant speeches for Democrats, and "Against Prior" messages are the converse.

Figure A8: Robustness Specifications, Republicans



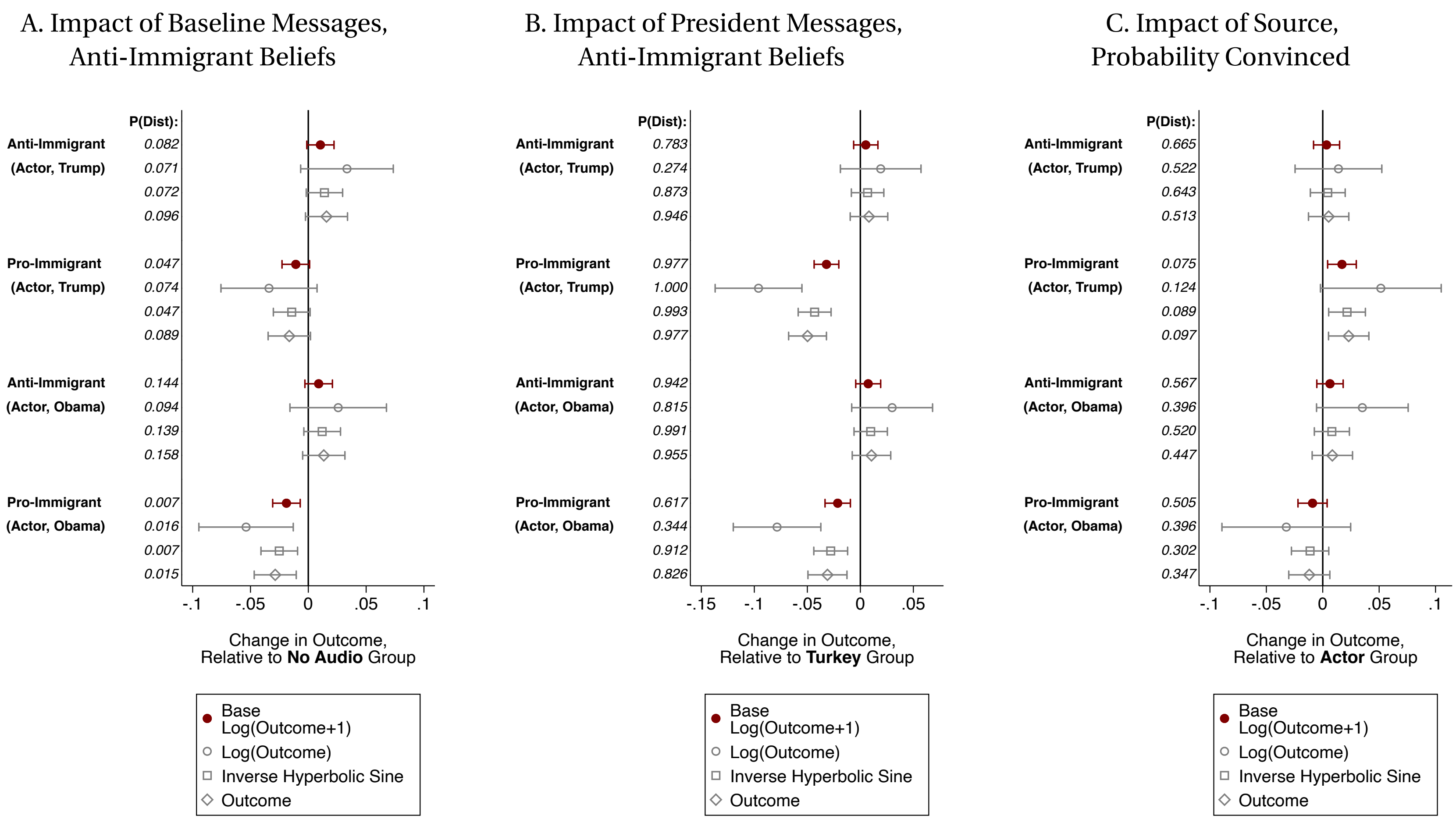
Notes: These figures replicate the baseline specification for alternative sample restrictions and specifications. All outcomes are measured using $ln(P(Outcome)+1)$, where $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. "Without Covariates" is a specification that drops the covariate controls described in Appendix A4. "No Duplicate Responders" excludes any individuals who attempted the survey more than once; the baseline sample includes first attempts of the survey for these individuals. "5 Option Version" constructs the index using the full set of 5 options for questions that have 5 options (e.g. Strongly Agree, Agree, Neutral, Disagree, Strongly Agree), rather than the baseline version that collapses all answers to have 3 options to match 3 option questions (e.g. Agree, Neutral, Disagree). "14 Question Version" constructs the index to match the pre-registered version, which excludes questions on the societal contribution of legal immigrants or immigrants from English-speaking countries. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A9: Robustness Specifications, Democrats



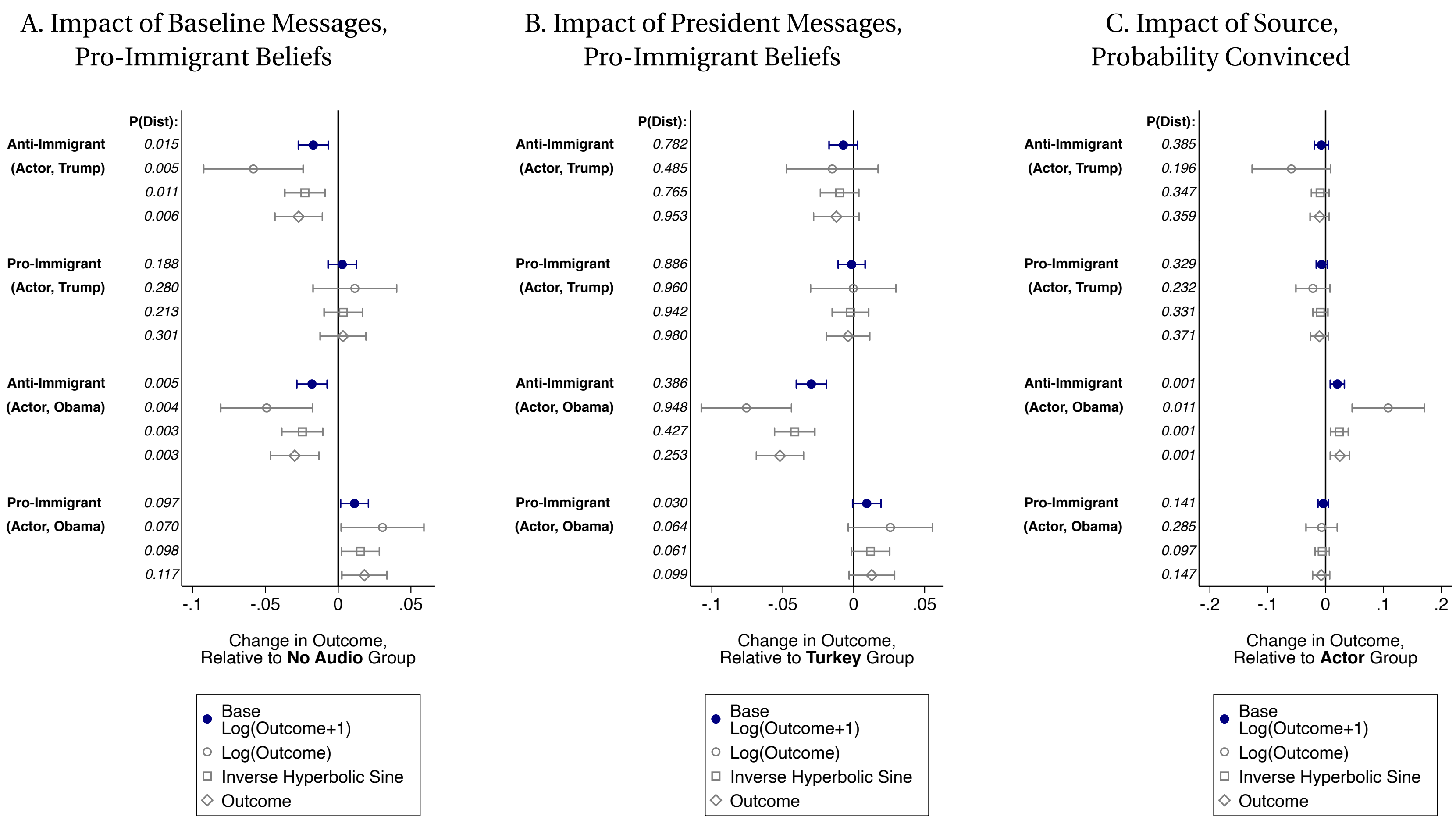
Notes: These figures replicate the baseline specification for alternative sample restrictions and specifications. All outcomes are measured using $ln(P(Outcome)+1)$, where $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. "Without Covariates" is a specification that drops the covariate controls described in Appendix A4. "Without Independents" drops individuals who were recruited as either Democrats or Republicans and then later stated that they were Independents within the survey. "No Duplicate Responders" excludes any individuals who attempted the survey more than once; the baseline sample includes first attempts of the survey for these individuals. "5 Option Version" constructs the index using the full set of 5 options for questions that have 5 options (e.g. Strongly Agree, Agree, Neutral, Disagree, Strongly Agree), rather than the baseline version that collapses all answers to have 3 options to match 3 option questions (e.g. Agree, Neutral, Disagree). "14 Question Version" constructs the index to match the pre-registered version, which excludes questions on the societal contribution of legal immigrants or immigrants from English-speaking countries. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A10: Robustness of Log Specification, Republicans



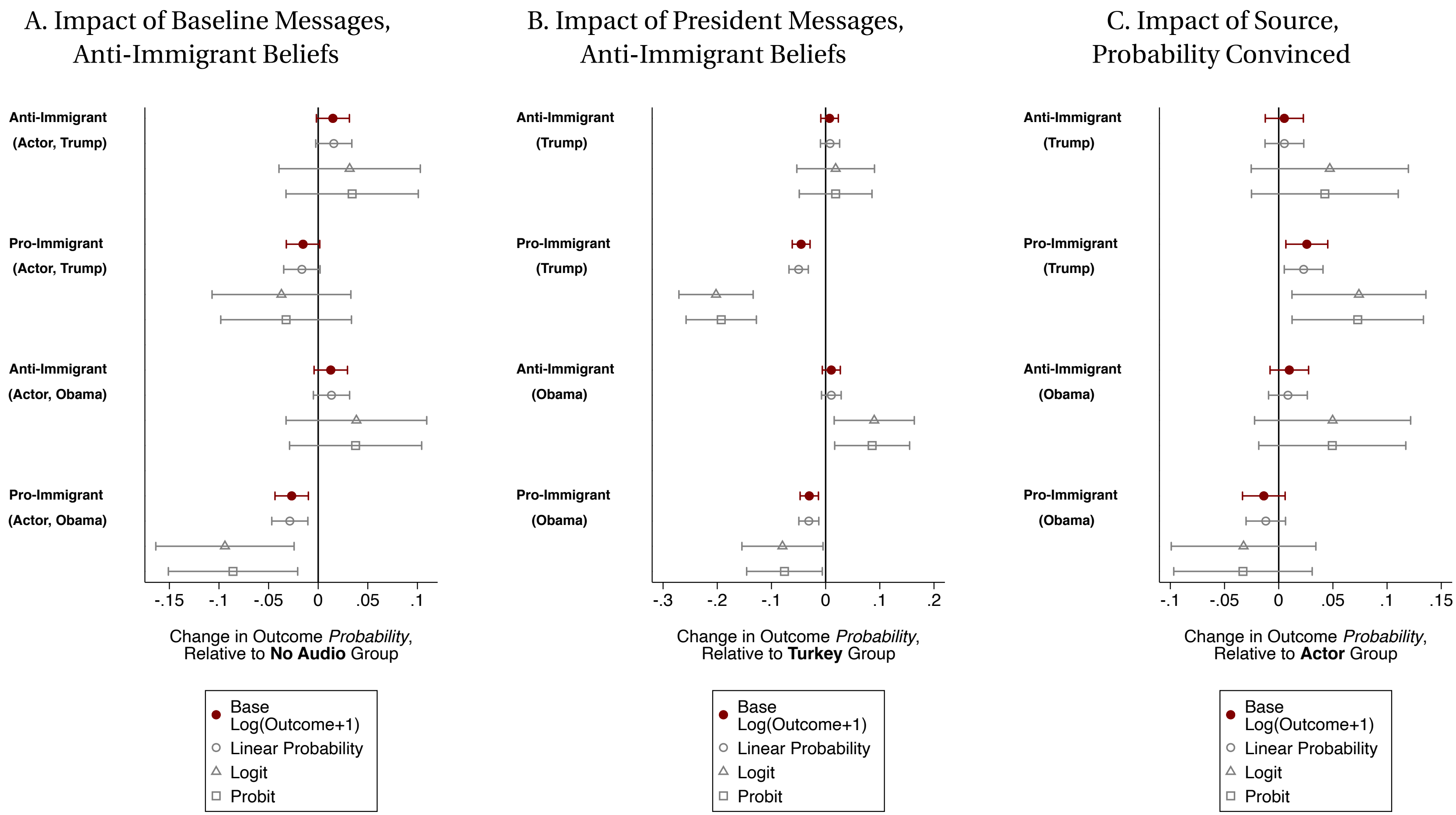
Notes: These figures replicate the baseline specification for alternative sample restrictions and specifications. The specifications compare the baseline transformation, $ln(P(Outcome)+1)$, with the alternative transformation, $ln(P(Outcome))$, where any observations with $P(Outcome) = 0$ are dropped. The figure also includes the unlogged outcome $P(Outcome)$ as the dependent variable. The standardized version of the unlogged outcome is not included in this plot given the difference in scale of this outcome; however, the results using this transformation are identical in significance and direction as the unlogged outcome, given that standardization is a linear transformation. $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A11: Robustness of Log Specification, Democrats



Notes: These figures replicate the baseline specification for alternative sample restrictions and specifications. The specifications compare the baseline transformation, $ln(P(Outcome)+1)$, with the alternative transformation, $ln(P(Outcome))$, where any observations with $P(Outcome) = 0$ are dropped. The figure also includes the unlogged outcome $P(Outcome)$ as the dependent variable. The standardized version of the unlogged outcome is not included in this plot given the difference in scale of this outcome; however, the results using this transformation are identical in significance and direction as the unlogged outcome, given that standardization is a linear transformation. $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A12: Change in Probability: Log, Linear Probability, Logit & Probit, Republicans



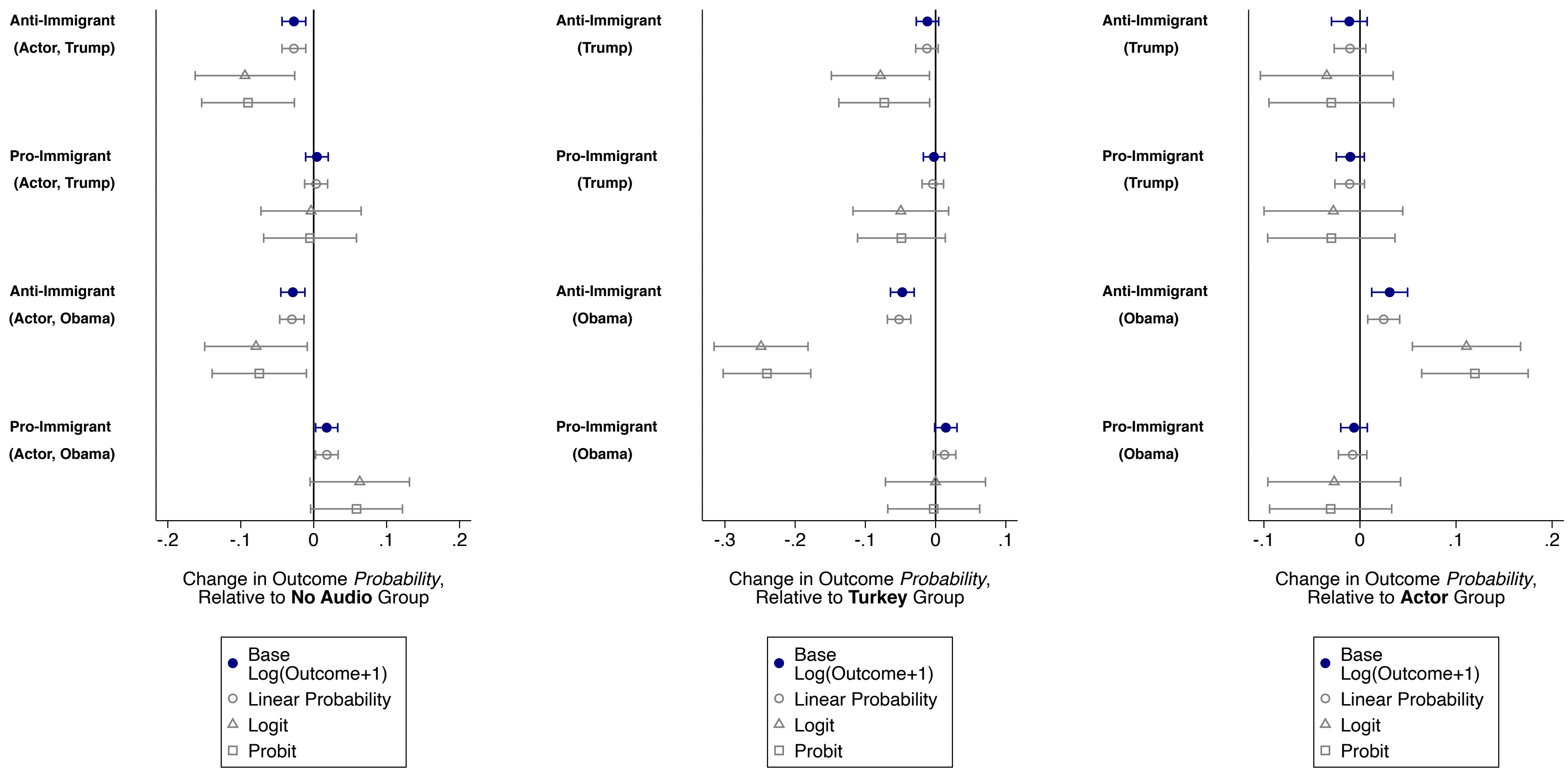
Notes: These figures replicate the baseline specification for alternative sample restrictions and specifications. Unlike the preceding tables, all coefficients from specifications are converted to corresponding changes in underlying probability indexes using the delta method, or estimates of marginal effects (at means) from logit or probit models. The estimates are drawn from the baseline transformation, $ln(P(Outcome)+1)$, a linear probability model with $P(Outcome)$ as the dependent variable, a logit model, and a probit model. To estimate the probit and logit models, the continuous $P(Outcome)$ measures are discretized at the median of the corresponding party no audio control group. $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. All regressions are adjusted for covariates in Appendix A4.

Figure A13: Change in Probability: Log, Linear Probability, Logit & Probit, Democrats

A. Impact of Baseline Messages, Pro-Immigrant Beliefs

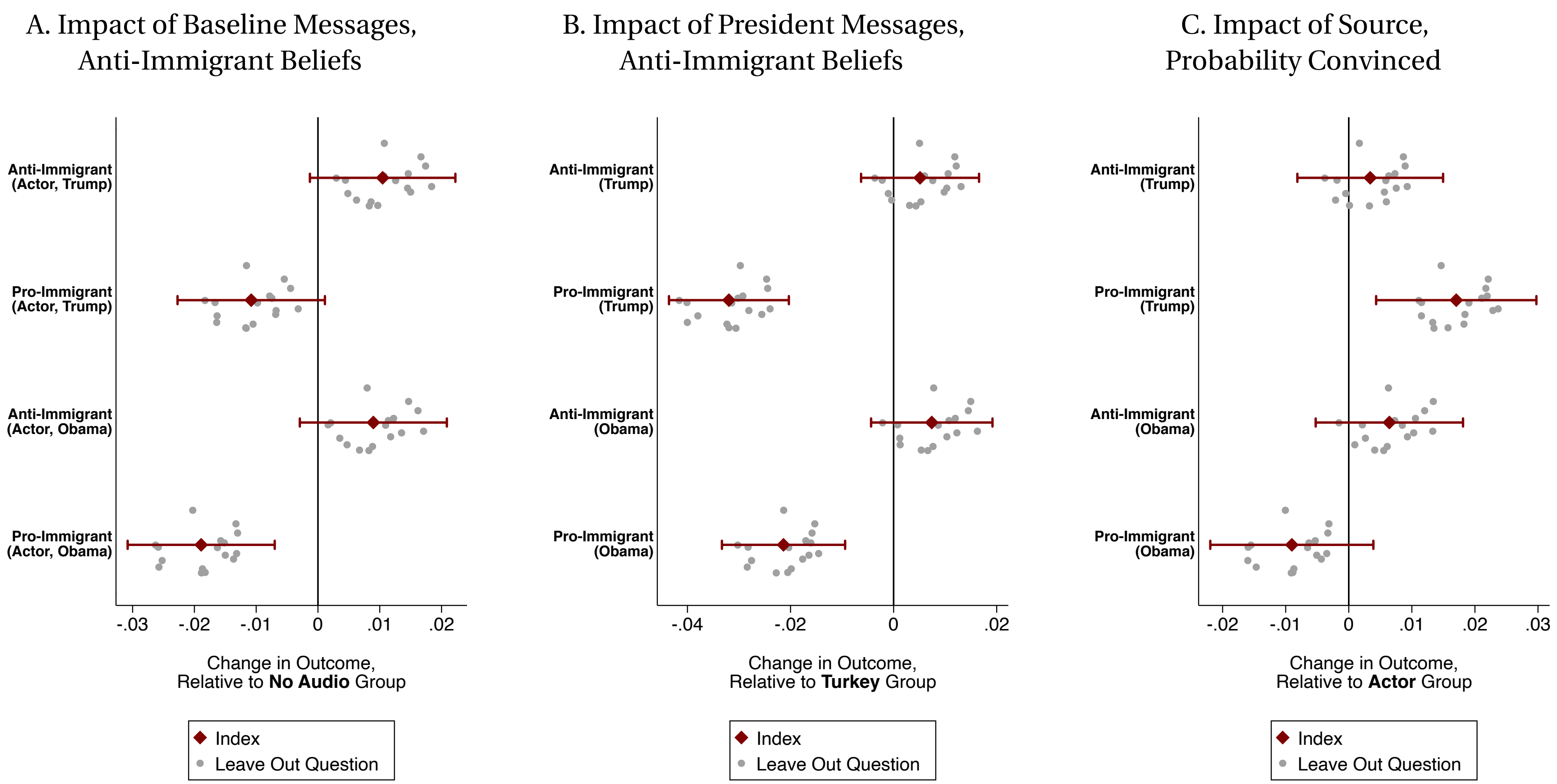
B. Impact of President Messages, Pro-Immigrant Beliefs

C. Impact of Source, Probability Convinced



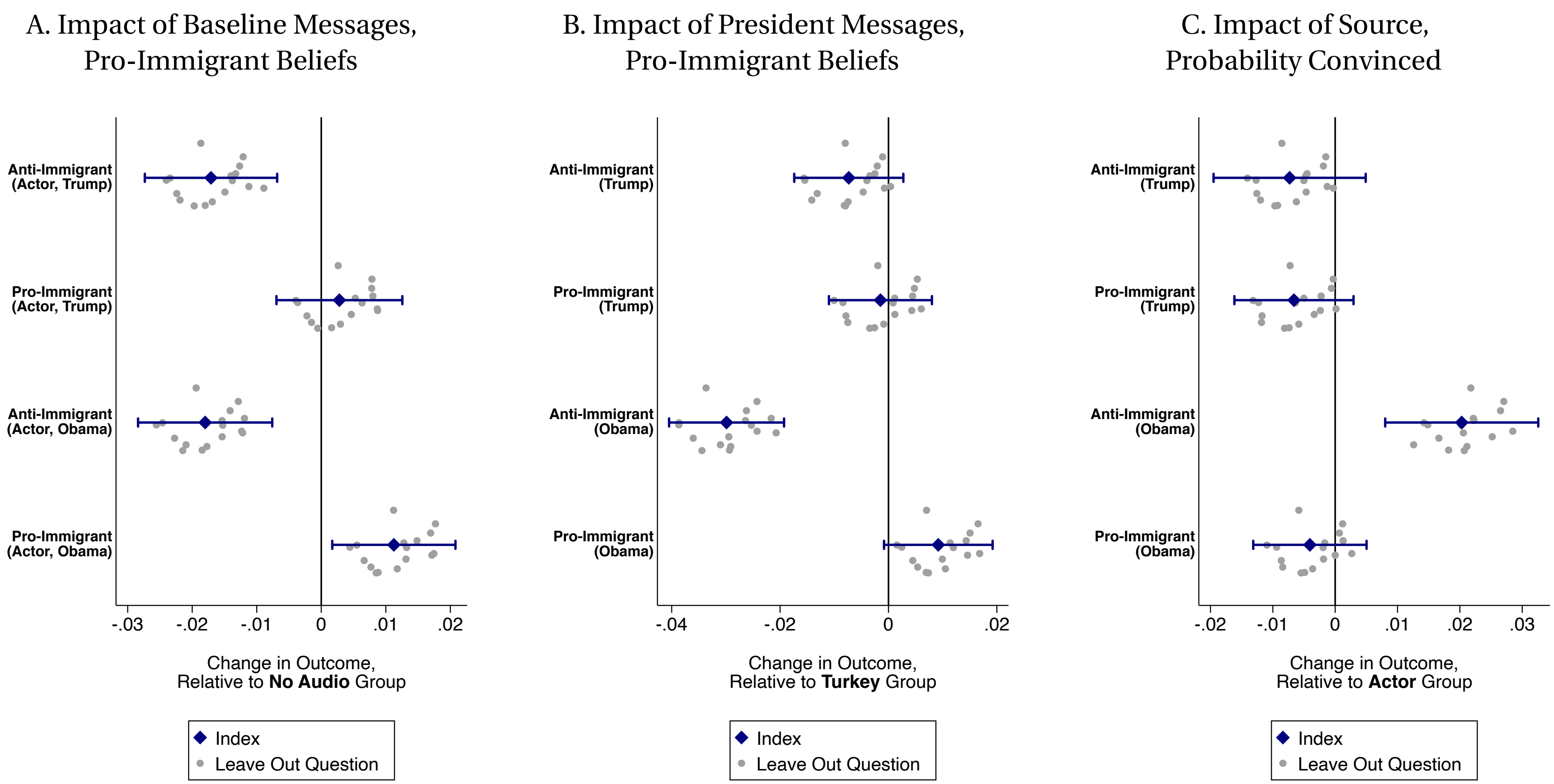
Notes: These figures replicate the baseline specification for alternative sample restrictions and specifications. Unlike the preceding tables, all coefficients from specifications are converted to corresponding changes in underlying probability indexes using the delta method, or estimates of marginal effects (at means) from logit or probit models. The estimates are drawn from the baseline transformation, $ln(P(Outcome)+1)$, a linear probability model with $P(Outcome)$ as the dependent variable, a logit model, and a probit model. To estimate the probit and logit models, the continuous $P(Outcome)$ measures are discretized at the median of the corresponding party no audio control group. $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. All regressions are adjusted for covariates in Appendix A4.

Figure A14: Leave-Out Question Distribution, Republicans



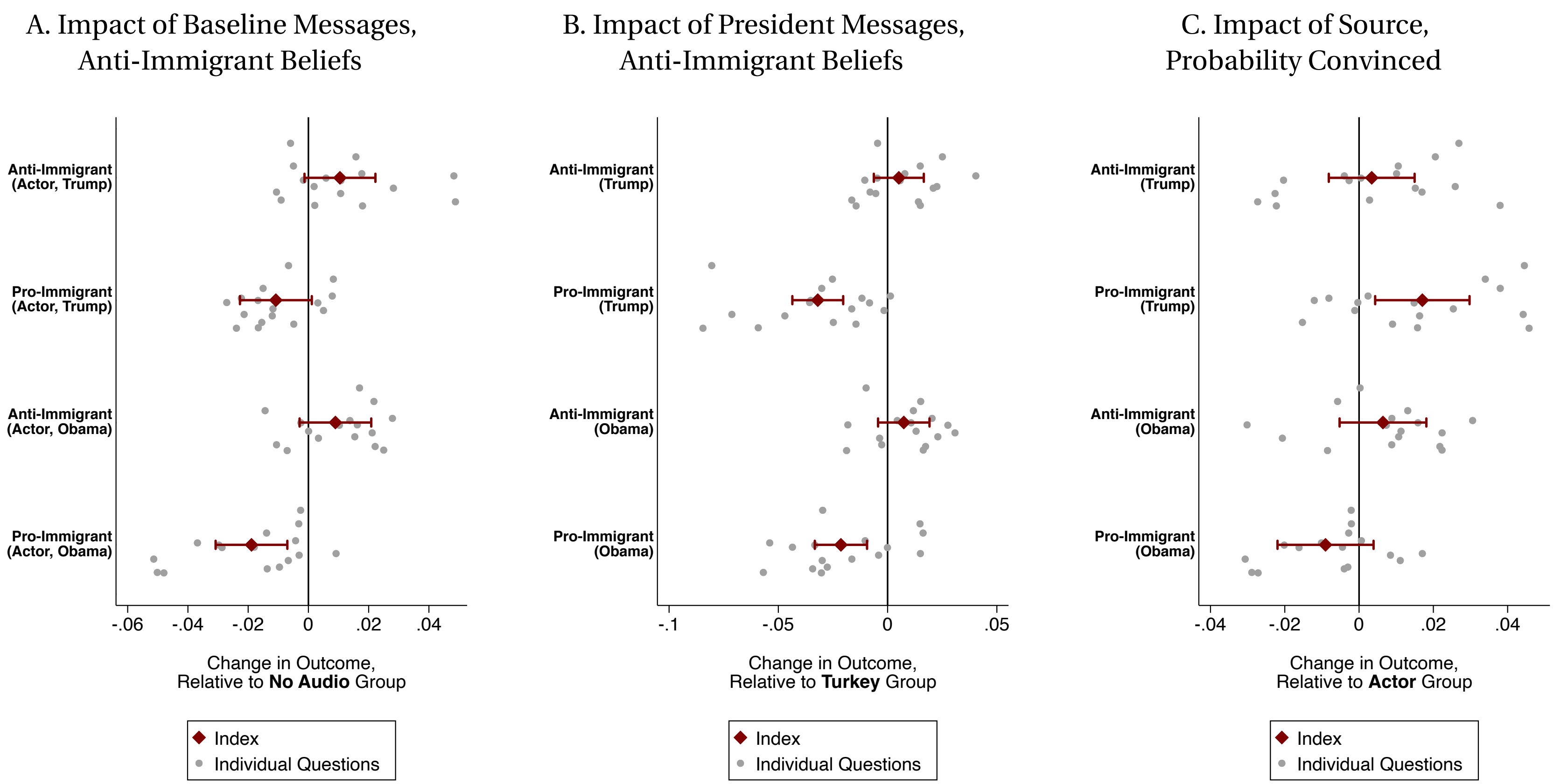
Notes: This figure overlays the baseline estimates of each test using the full index on top of 16 separate versions of each outcome that each leave out a single question from the index. All outcomes are measured using $ln(P(Outcome)+1)$ and coefficients represent changes in this transformed outcome. $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. All regressions are adjusted for covariates in Appendix A4. “P(Dist)” is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A15: Leave-Out Question Distribution, Democrats



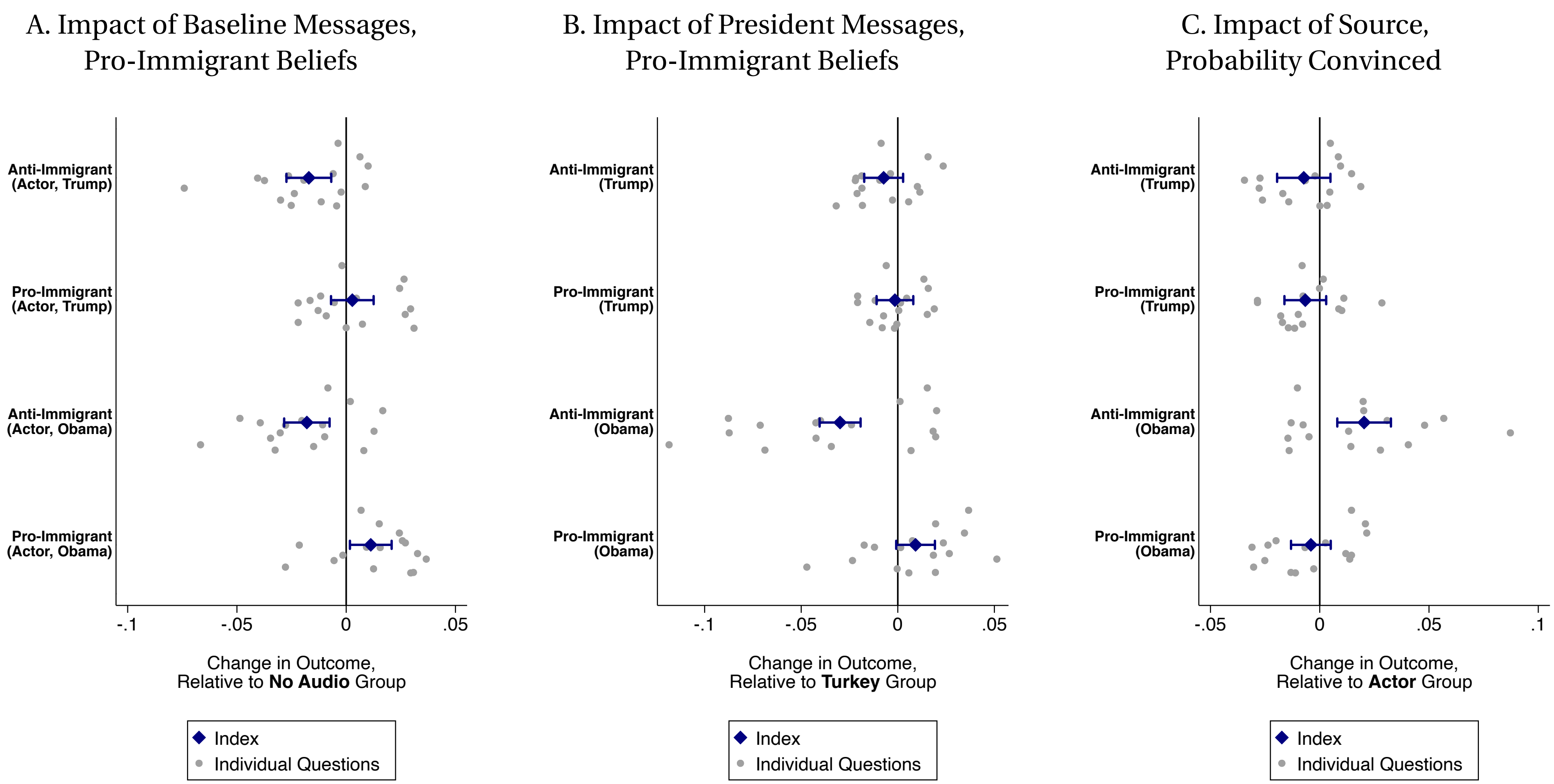
Notes: This figure overlays the baseline estimates of each test using the full index on top of 16 separate versions of each outcome that each leave out a single question from the index. All outcomes are measured using $ln(P(Outcome)+1)$ and coefficients represent changes in this transformed outcome. $P(Outcome)$ is either the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A16: By Question Distribution, Republicans



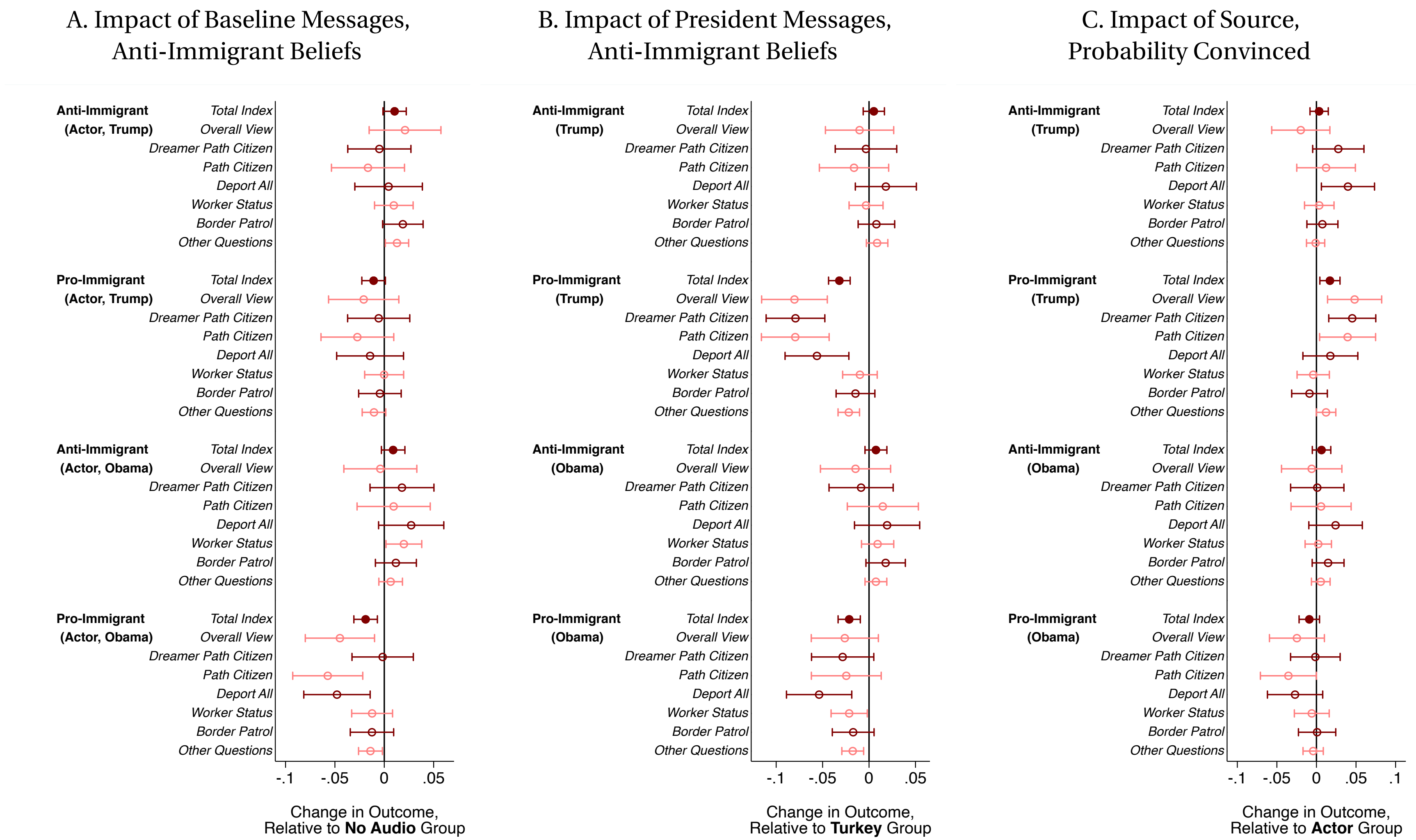
Notes: This figure overlays the baseline estimates of each test using the full index on top of the outcome for each of the 16 questions that comprises the index. All outcomes are measured using $ln(P(Outcome)+1)$ and coefficients represent changes in this transformed outcome. $P(Outcome)$ corresponds to the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced, or the individual question response version that enters each index. All regressions are adjusted for covariates in Appendix A4. “P(Dist)” is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A17: By Question Distribution, Democrats



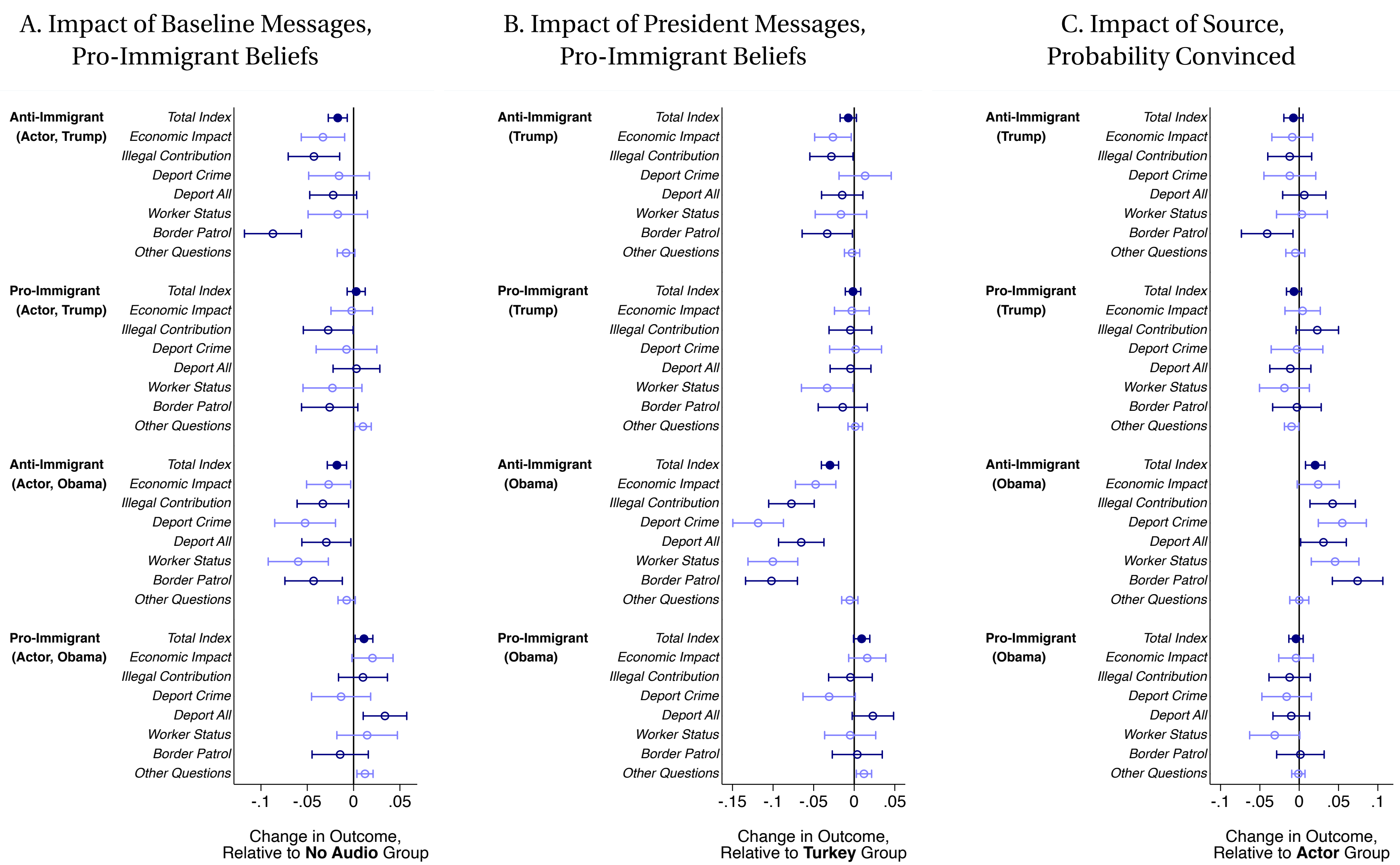
Notes: This figure overlays the baseline estimates of each test using the full index on top of the outcome for each of the 16 questions that comprises the index. All outcomes are measured using $ln(P(Outcome)+1)$ and coefficients represent changes in this transformed outcome. $P(Outcome)$ corresponds to the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced, or the individual question response version that enters each index. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A18: By Question Effects, Republicans



Notes: This figure overlays the baseline estimates of each test using the total index on top, followed by the 6 questions that show the most movement for each party, and an outcome that averages the remaining 10 questions of the index. All outcomes are measured using $ln(P(Outcome)+1)$ and coefficients represent changes in this transformed outcome. $P(Outcome)$ corresponds to the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced, or the individual question response version that enters each index. All regressions are adjusted for covariates in Appendix A4.

Figure A19: By Question Effects, Democrats



Notes: This figure overlays the baseline estimates of each test using the total index on top, followed by the 6 questions that show the most movement for each party, and an outcome that averages the remaining 10 questions of the index. All outcomes are measured using $ln(P(Outcome)+1)$ and coefficients represent changes in this transformed outcome. $P(Outcome)$ corresponds to the anti-immigrant index probability, the pro-immigrant index probability, or the probability convinced, or the individual question response version that enters each index. All regressions are adjusted for covariates in Appendix A4.

Figure A20: By Question Effects (All Questions), Republicans
(Expanded version of Figure A18)

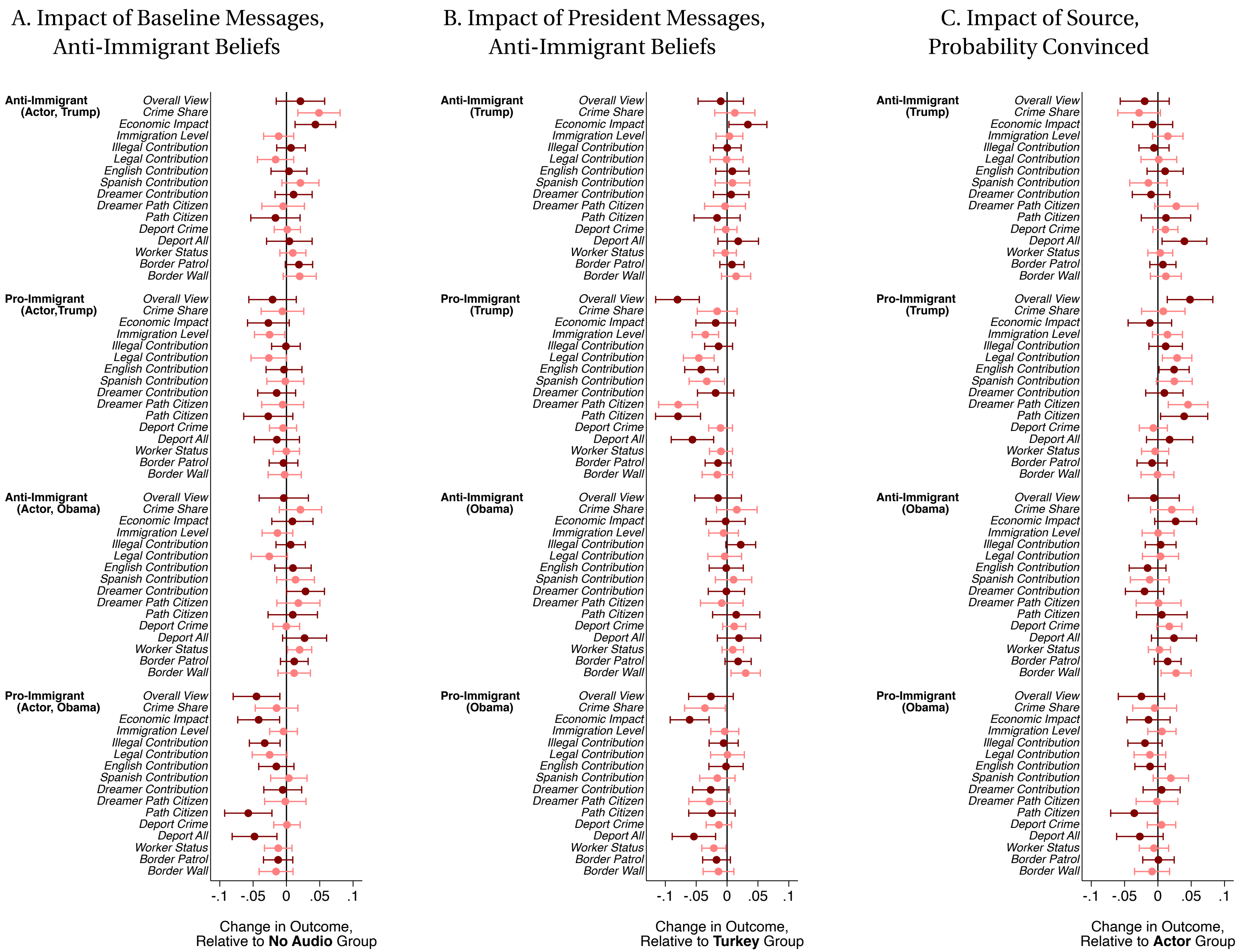


Figure A21: By Question Effects (All Questions), Democrats
(Expanded version of Figure A19)

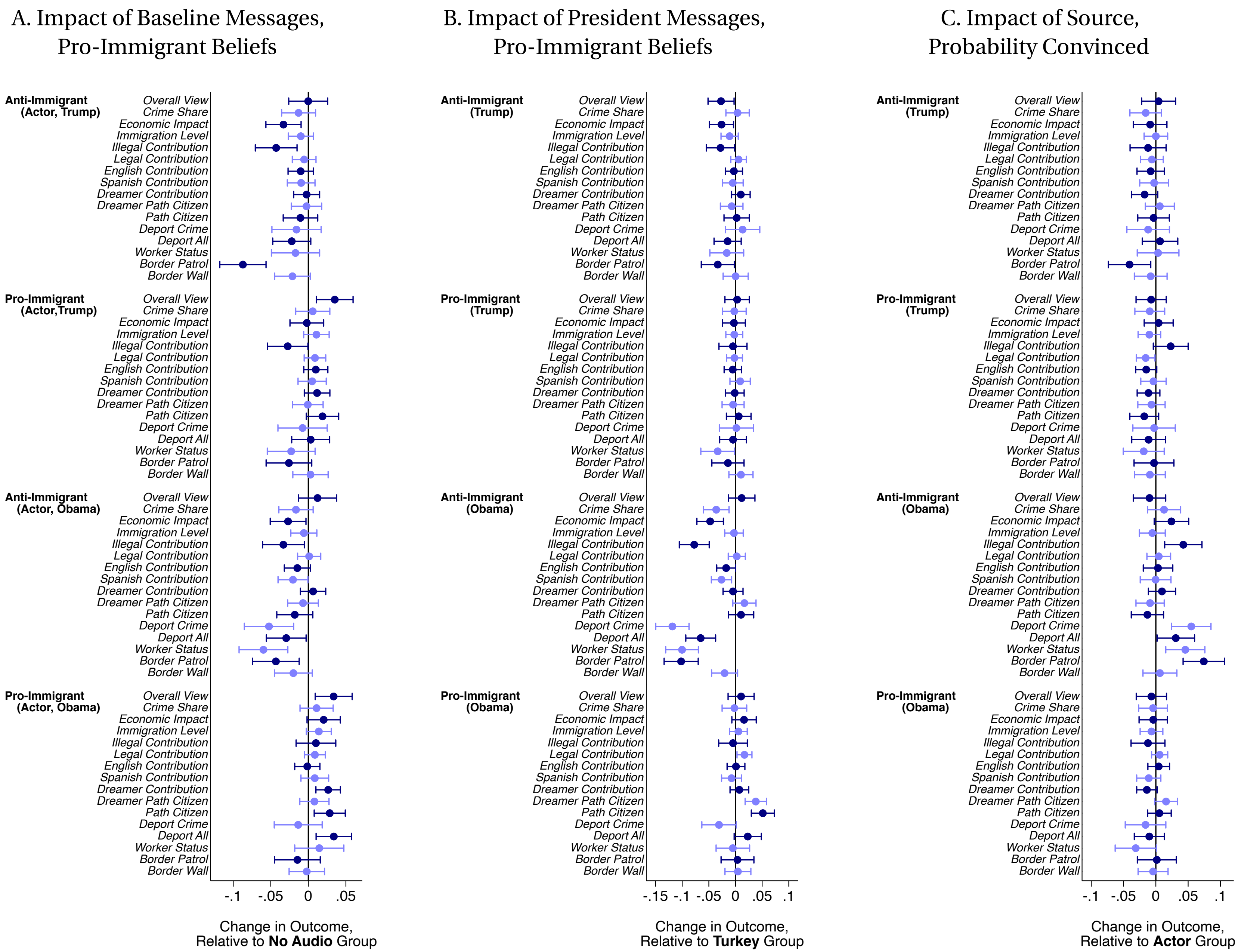
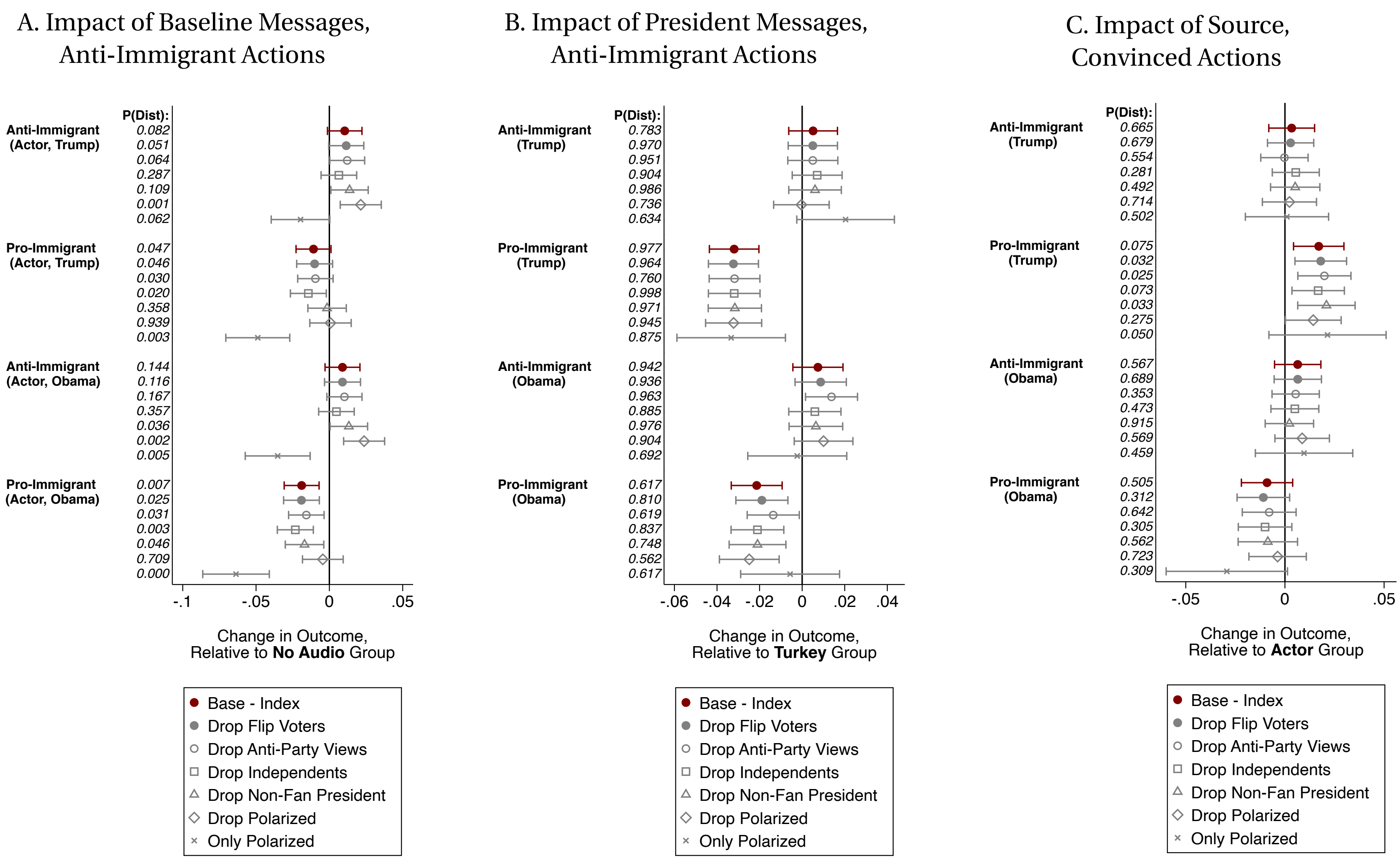
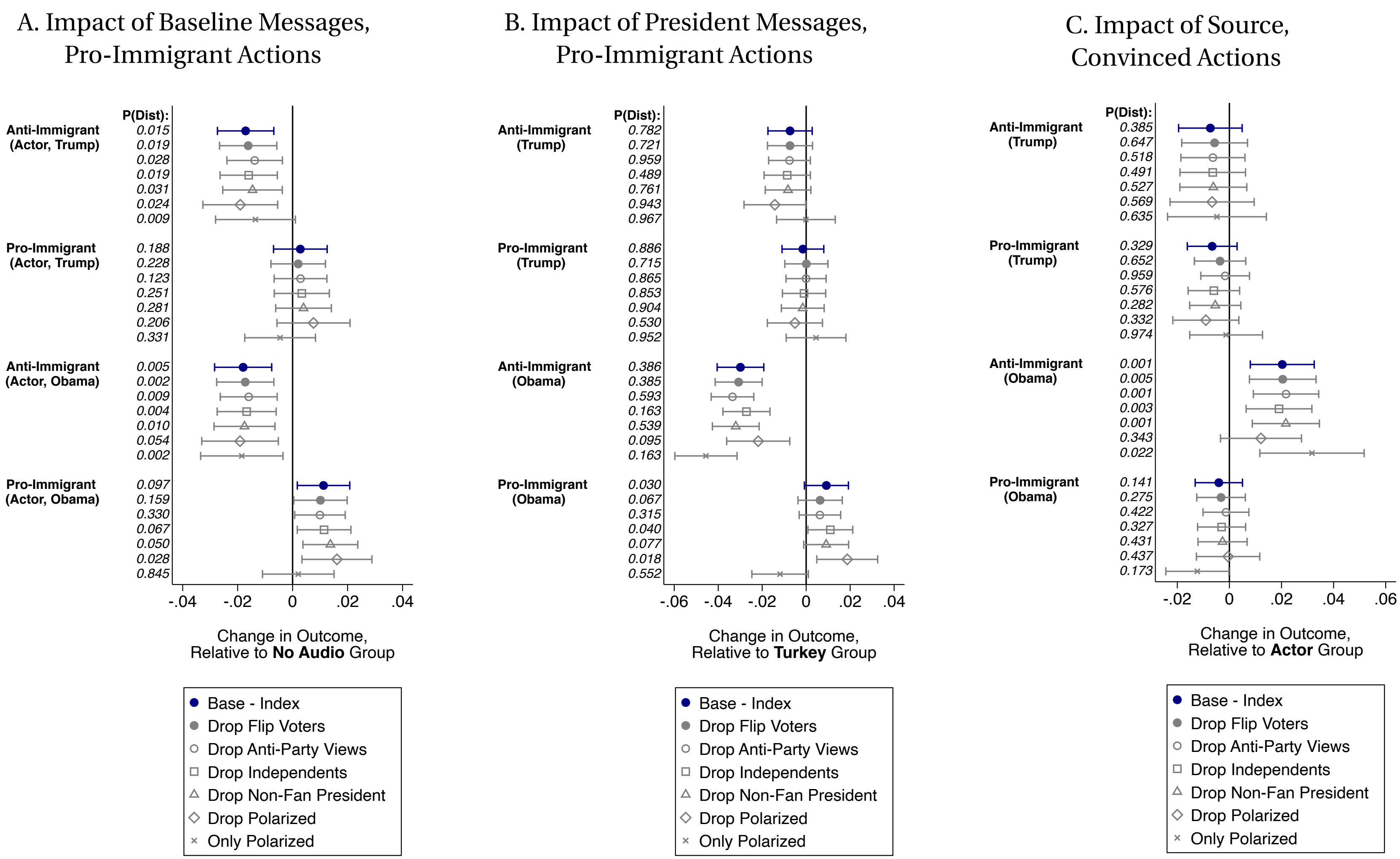


Figure A22: Heterogeneity: Moderates, Republicans



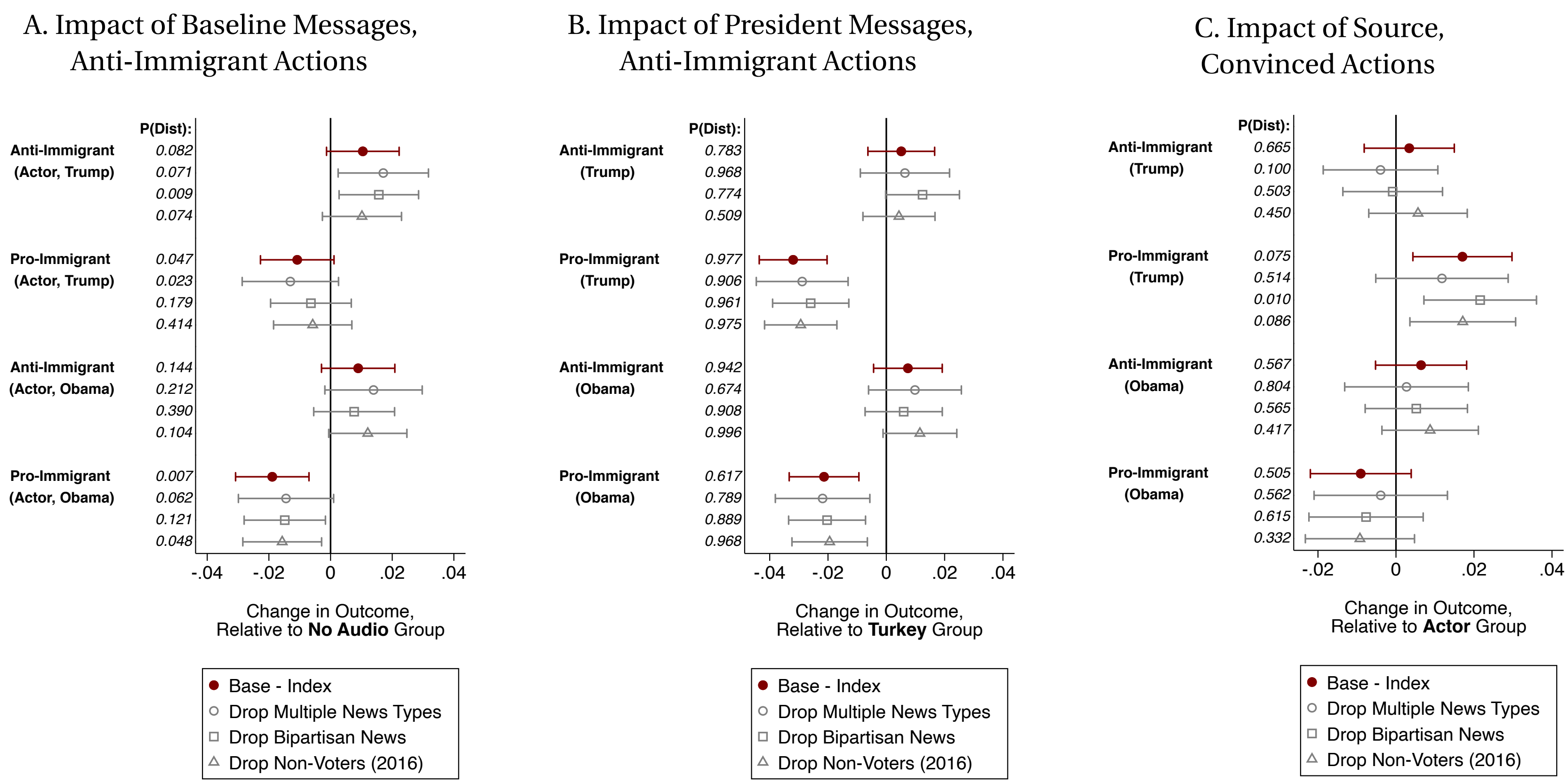
Notes: This plot shows the estimates for the baseline outcomes as compared to samples that remove either moderates or extremists. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. "Drop Flip Voters" removes Republicans who voted for Hillary Clinton in 2016 and Democrats who voted for Donald Trump in 2016. "Drop Anti-Party Views" removes Republicans (Democrats) who answered a pre-treatment question saying that they think immigration should be expanded (restricted). "Drop Independents" drops individuals who were recruited as either Democrats or Republicans and then later stated that they were Independents within the survey. "Drop Non-Fan President" excludes individuals who are not fans of the president from their party. "Drop Polarized" excludes those who are predicted to have extreme immigration views using a probit predictive model with pre-treatment characteristics estimated on the control group and extrapolated to the rest of the sample. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A23: Heterogeneity: Moderates, Democrats



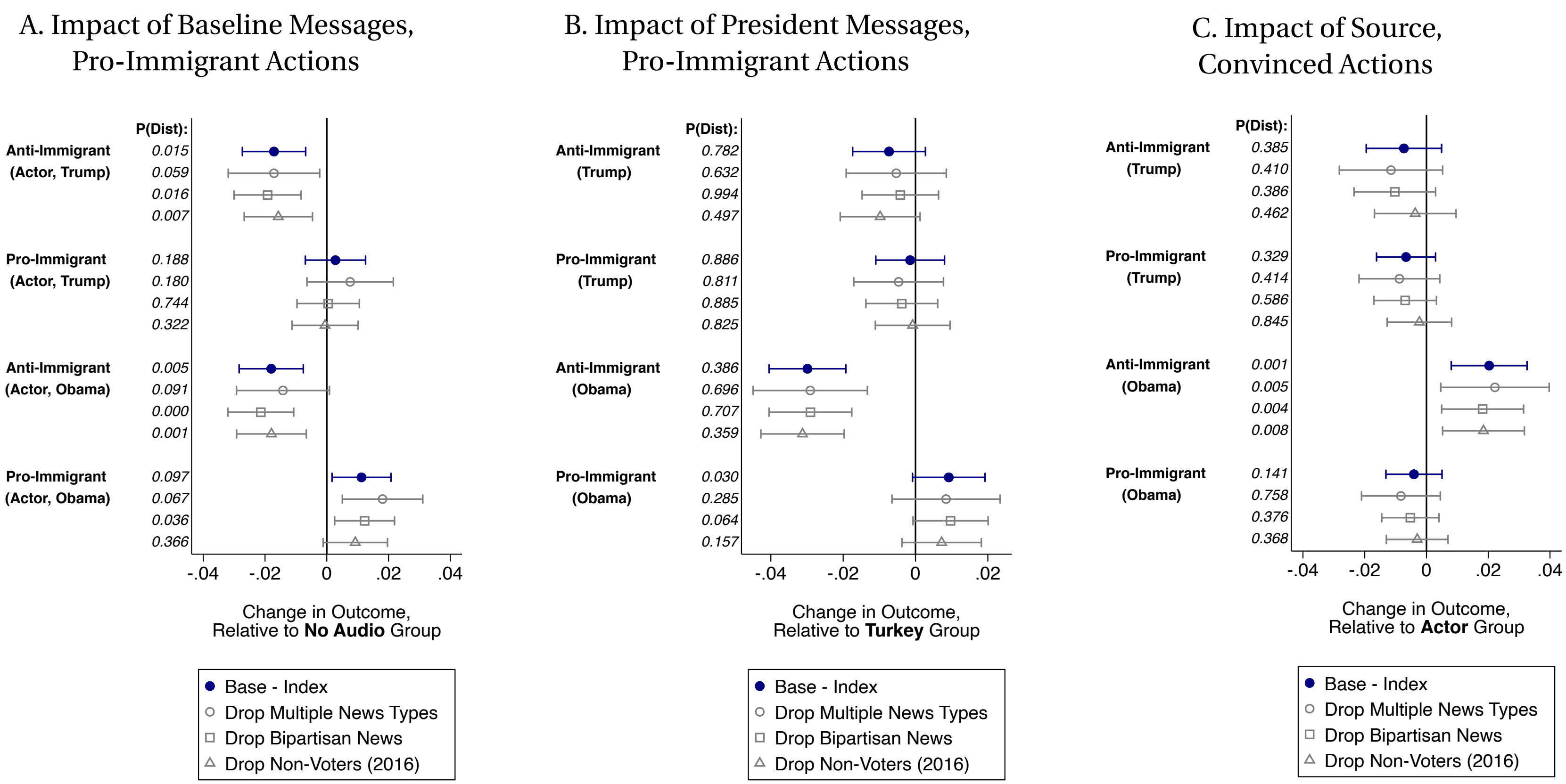
Notes: This plot shows the estimates for the baseline outcomes as compared to samples that remove either moderates or extremists. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. “Drop Flip Voters” removes Republicans who voted for Hillary Clinton in 2016 and Democrats who voted for Donald Trump in 2016. “Drop Anti-Party Views” removes Republicans (Democrats) who answered a pre-treatment question saying that they think immigration should be expanded (restricted). “Drop Independents” drops individuals who were recruited as either Democrats or Republicans and then later stated that they were Independents within the survey. “Drop Non-Fan President” excludes individuals who are not fans of the president from their party. “Drop Polarized” excludes those who are predicted to have extreme immigration views using a probit predictive model with pre-treatment characteristics estimated on the control group and extrapolated to the rest of the sample. All regressions are adjusted for covariates in Appendix A4. “P(Dist)” is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A24: Heterogeneity: Informed/Engaged, Republicans



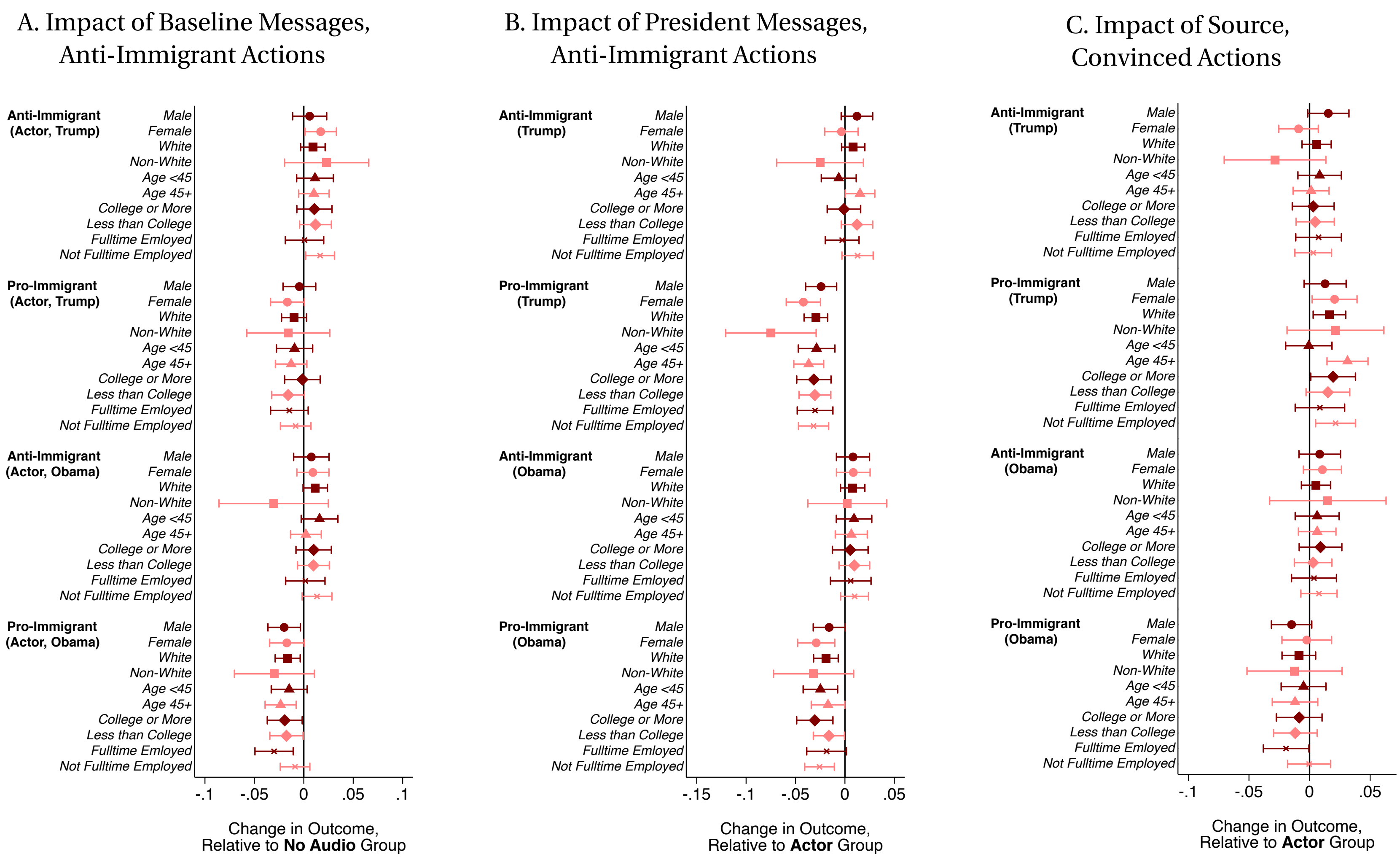
Notes: This plot shows the estimates for the baseline outcomes as compared to samples that remove engaged/informed or non-engaged/non-informed participants. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. "Drop Multiple News Types" removes participants who consume news at least daily through more than one mode: Newspaper, TV, Twitter and/or Facebook. "Drop Bi-Partisan News" removes participants who consume news from both a right-leaning and left-leaning news outlet at least weekly. "Drop Non-Voters (2016)" removes individuals who did not vote in the 2016 presidential election. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A25: Heterogeneity: Informed/Engaged, Democrats



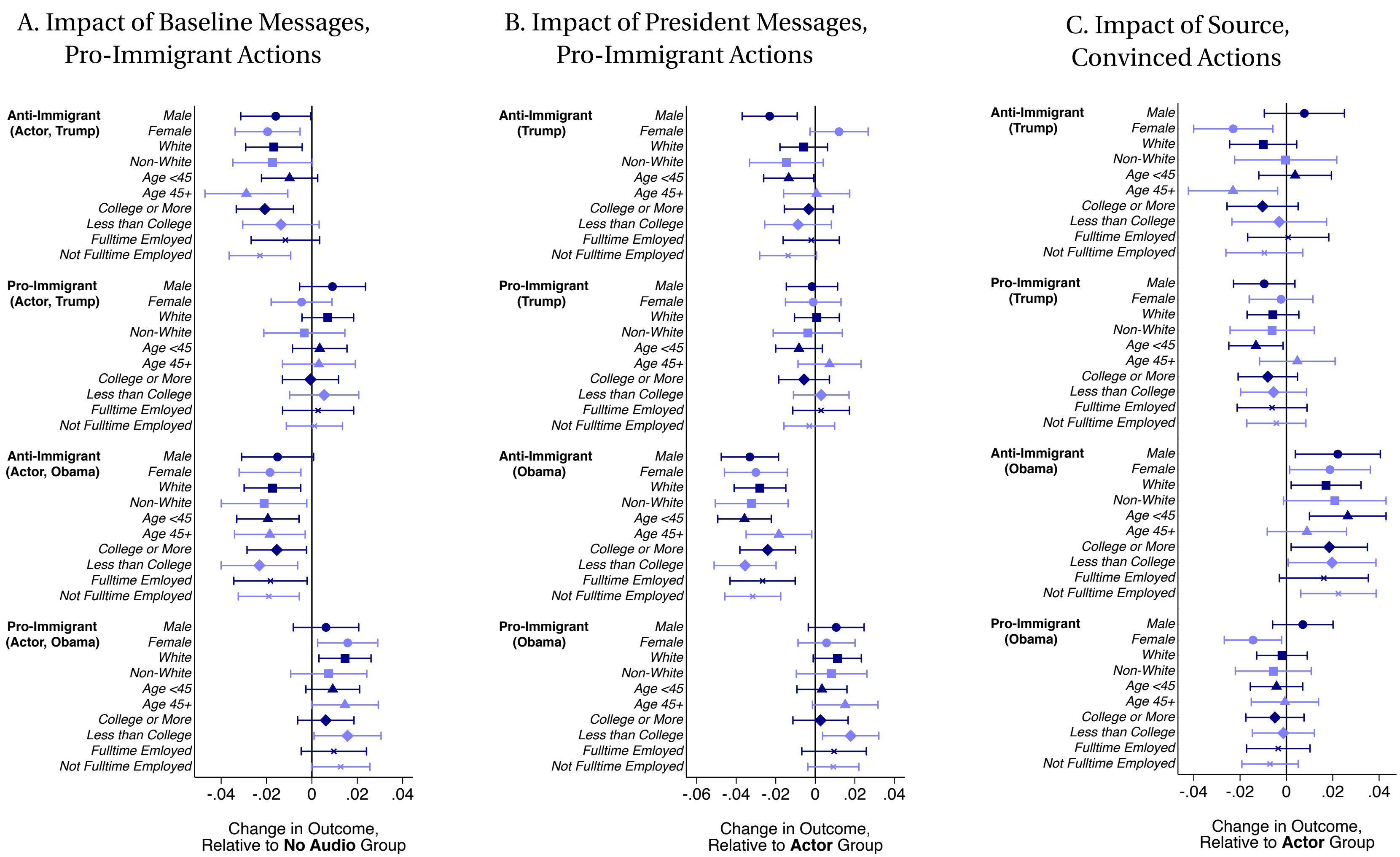
Notes: This plot shows the estimates for the baseline outcomes as compared to samples that remove engaged/informed or non-engaged/non-informed participants. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. “Drop Multiple News Types” removes participants who consume news at least daily through more than one mode: Newspaper, TV, Twitter and/or Facebook. “Drop Bi-Partisan News” removes participants who consume news from both a right-leaning and left-leaning news outlet at least weekly. “Drop Non-Voters (2016)” removes individuals who did not vote in the 2016 presidential election. All regressions are adjusted for covariates in Appendix A4. “P(Dist)” is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A26: Heterogeneity: Demographics, Republicans



Notes: This plot shows the estimates for demographic sub-groups of the data. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. All regressions are adjusted for covariates in Appendix A4.

Figure A27: Heterogeneity: Demographics, Democrats



Notes: This plot shows the estimates for demographic sub-groups of the data. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. All regressions are adjusted for covariates in Appendix A4.

A3 Bayesian Framework: Additional Details

This section provides detailed derivations for results presented in Section 2.

Derivations for Equation (2). Adding 1 to both sides of Equation (1), we get

$$1+\mathbb{P}(\omega|m) = 1+\mathbb{P}(\omega)\times\Theta(m|\omega) = (1+\mathbb{P}(\omega))(1+\frac{\mathbb{P}(\omega)}{1+\mathbb{P}(\omega)}(\Theta(m|\omega)-1)) \tag{15}$$

Taking logs of both sides of the equation above gives us Equation (2).

Derivations for Remark 1. Differentiating $\Delta(m|\omega)$ with respect to $\Theta(m|\omega)$, we have

$$\frac{\partial\Delta(m|\omega)}{\partial\Theta(m|\omega)} = \frac{\mathbb{P}(\omega)}{1+\mathbb{P}(\omega)\Theta(m|\omega)} \geq 0 \tag{16}$$

Since the derivative is positive, we see that $\Delta(m|\omega)$ is increasing in $\Theta(m|\omega)$. Finally, note that if $\Theta(m|\omega)=1$ then $\Delta(m|\omega)=\ln(1)=0$.

Derivations for Equation (7). Adding 1 to both sides of Equation (4), we get

$$1+\mathbb{P}(\omega|s) = 1+\mathbb{P}(\omega)\times\Theta(s|\omega) = (1+\mathbb{P}(\omega))(1+\frac{\mathbb{P}(\omega)}{1+\mathbb{P}(\omega)}(\Theta(m|\omega)-1)) \tag{17}$$

Taking logs of both sides of the equation above gives us Equation (7).

Derivations for Remark 3. Using Equation (8) we have

$$\Theta(s|\omega,m)-1 = \frac{\mathbb{P}(s|\omega,m)-\mathbb{P}(\omega|m)}{\mathbb{P}(\omega|m)} \tag{18}$$

Now, note that

$$\mathbb{P}(\omega|m) = \mathbb{P}(\omega,s|m)+\mathbb{P}(\omega,\neg s|m) = \mathbb{P}(s|m)\mathbb{P}(\omega|s,m)+\mathbb{P}(\neg s|m)\mathbb{P}(\omega|\neg s,m) \tag{19}$$

Finally, plugging this into the numerator in Equation (18) and using $\mathbb{P}(\neg s|m) = 1-\mathbb{P}(s|m)$ gives us the expression of interest:

$$\Theta(s|\omega,m)-1 = \frac{(1-\mathbb{P}(s|m))(\mathbb{P}(\omega|s,m)-\mathbb{P}(\omega|\neg s,m))}{\mathbb{P}(\omega|m)} \tag{20}$$

Derivations for Equation (11). Adding 1 to both sides of Equation (8), we get

$$1+\mathbb{P}(\omega|s,m) = 1+\mathbb{P}(\omega|m)\times\Theta(s|\omega,m) = (1+\mathbb{P}(\omega|m))(1+\frac{\mathbb{P}(\omega|m)}{1+\mathbb{P}(\omega|m)}(\Theta(s|\omega,m)-1)) \tag{21}$$

Taking logs of both sides of the equation above gives us Equation (11).

A4 Selecting of Control Variables: Double Selection Lasso

Approach

To select the control variables for our baseline specification we use the post double selection lasso (PDS) methodology of Belloni et al., (2012, 2014, 2015, 2016), using the STATA package by Ahrens et al., (2020). The procedure consists of first including all potential control variables from baseline and converting categorical variables into sets of indicators. Our set of potential controls is quite rich given the large number of survey questions we ask to participants prior to the treatment. These questions can be viewed in Supplemental Materials Appendix S1.

Next we keep the controls selected from the LASSO minimization of a regression which predicts treatment assignment. Then we repeat this exercise for a regression that predicts the outcome variable and take the union of controls selected from both procedures as our baseline controls. We do this procedure separately for Democrats and Republicans, given that our randomization was stratified by party and that the demographic characteristics of party groups differ in meaningful ways.

Control Variables in the Model

From the first step, none of the control variables are selected under the lasso procedure as important predictors of treatment status. From the second step, several variables are selected which are predictive of the outcome variable.

Republican Sample

The variables selected for Republicans are whether a participant is Hispanic, is in the age group 55 to 64, voted for Hillary Clinton in 2016, voted for a third party candidate in 2016, had the view that the level of immigration in the U.S. should decrease (pre-treatment), has the view that gun control regulations should be less strict, has the view that abortion should be illegal in all cases, believes taxes are too high, believes that there should be no government intervention in the healthcare system, is an occasional twitter user, is an occasional Buzzfeed reader, is a fan of Donald Trump, is a fan of Lebron James, is a fan of Taylor Swift, is a fan of Bill Gates, is a fan of Barack Obama, has no opinion towards Barack Obama, and views the most important policy issue to their vote as healthcare.

Democrat Sample

The variables selected for Democrats are whether a participant is Black, is Hispanic, is in the age group 35 to 44, is in the age group 45 to 54, has a high school degree as their highest level of education, voted for Hillary Clinton in 2016, had the view that the level of immigration in the U.S. should decrease (pre-treatment), had the view that the level of immigration in the U.S. should stay the same (pre-treatment), has the view that abortion should be illegal in all cases, has the view that abortion should be legal in some cases, believes taxes are too high, believes that there should be no government intervention in the healthcare system, feels neutral about government intervention in the healthcare system, is a daily, weekly, or occasionally reader of the New York Times (separate indicators), is a daily or occasional viewer of TV as a news source (separate indicators), is a daily newspaper reader, is a daily, weekly, or occasional viewer of Fox News (separate indicators), is a daily or weekly Breitbart News reader (separate indicators), is an occasional Buzzfeed reader, is not a Lebron James fan, is a Barack Obama fan, is a Donald Trump fan, is not a fan of Donald Trump, lives in the Western U.S., region response west, and views the most important policy issue to their vote as taxes.

While these variables improve the precision of estimates, it is important to note that the models are robust to excluding all controls, as is shown in Appendix Figures A8 and A9.

A5 Out of Sample Outcomes Measured in Survey

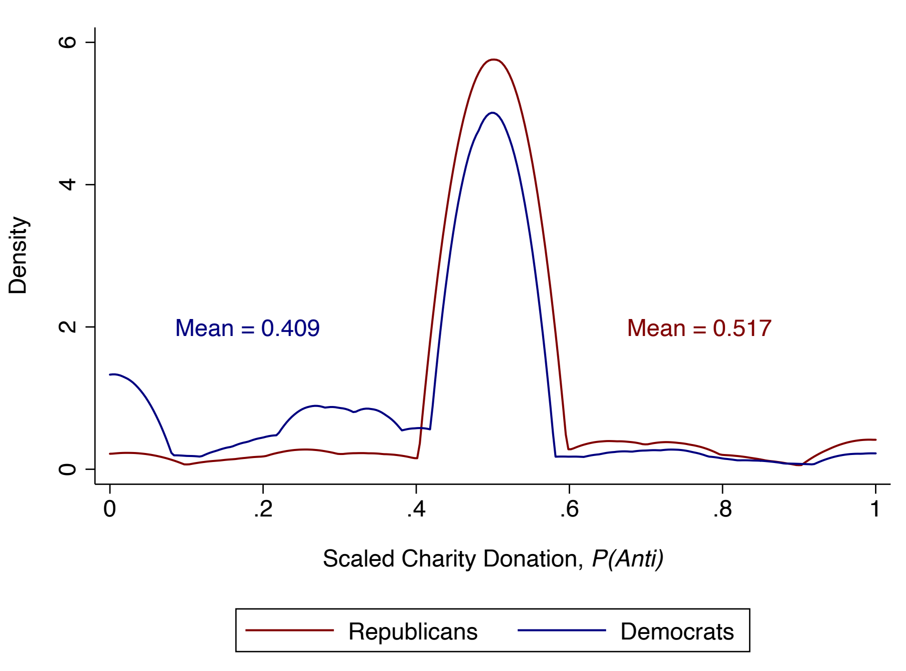
In addition to collecting information on immigrant views, we also attempted to measure two out-of-sample outcomes as a way to capture how actions might respond to a change in immigration beliefs. First, we asked participants about their intended vote in the 2020 presidential election, where we interpret a vote for Donald Trump as an anti-immigrant outcome and a vote for Joseph Biden as a pro-immigrant outcome. Second, we asked whether the participant would be interested in donating money to an anti-immigrant (Federation for American Immigration Reform (FAIR)) or a pro-immigrant charity (American Civil Liberties Union (ACLU) Immigrants' Rights Project), out of potential lottery earnings, if the participant wins a lottery that was conducted as part of the survey. The lottery offered a potential prize of $25 payable directly to the winner. If the winner chooses to donate to one of the two charities, she could choose the amount to be donated of the earnings. Any donation would be doubled before donation, such that the maximum donation amount was $50.

Figure A29 and A30 shows the results for these two outcomes as compared to the baseline immigration index outcomes. Unfortunately, the results here are noisy, imprecise and cannot reject the null hypothesis of no effect for either outcome. In cases where there is a significant change in the coefficient (or mean difference across groups), there is no corresponding significant difference in the distributions using the Kolmogorov-Smirnov test, and vice versa.

Multiple factors may contribute to the lack of effects we find for these outcomes. First, we may fail to find effects for these outcomes could be due to lack of precision and limited power. In addition, our ability to measure the voting outcome is limited by the fact that many participants may have voted prior to taking the survey, through early or on-time voting at polling places. While we intended to complete the study prior to the election on November 3, 2020, our survey recruitment period took longer than we had anticipated and continued past the presidential election to to November 10th. We need to exclude the data collected on or after the election for the voting outcome, which also limits the sample size for this test.

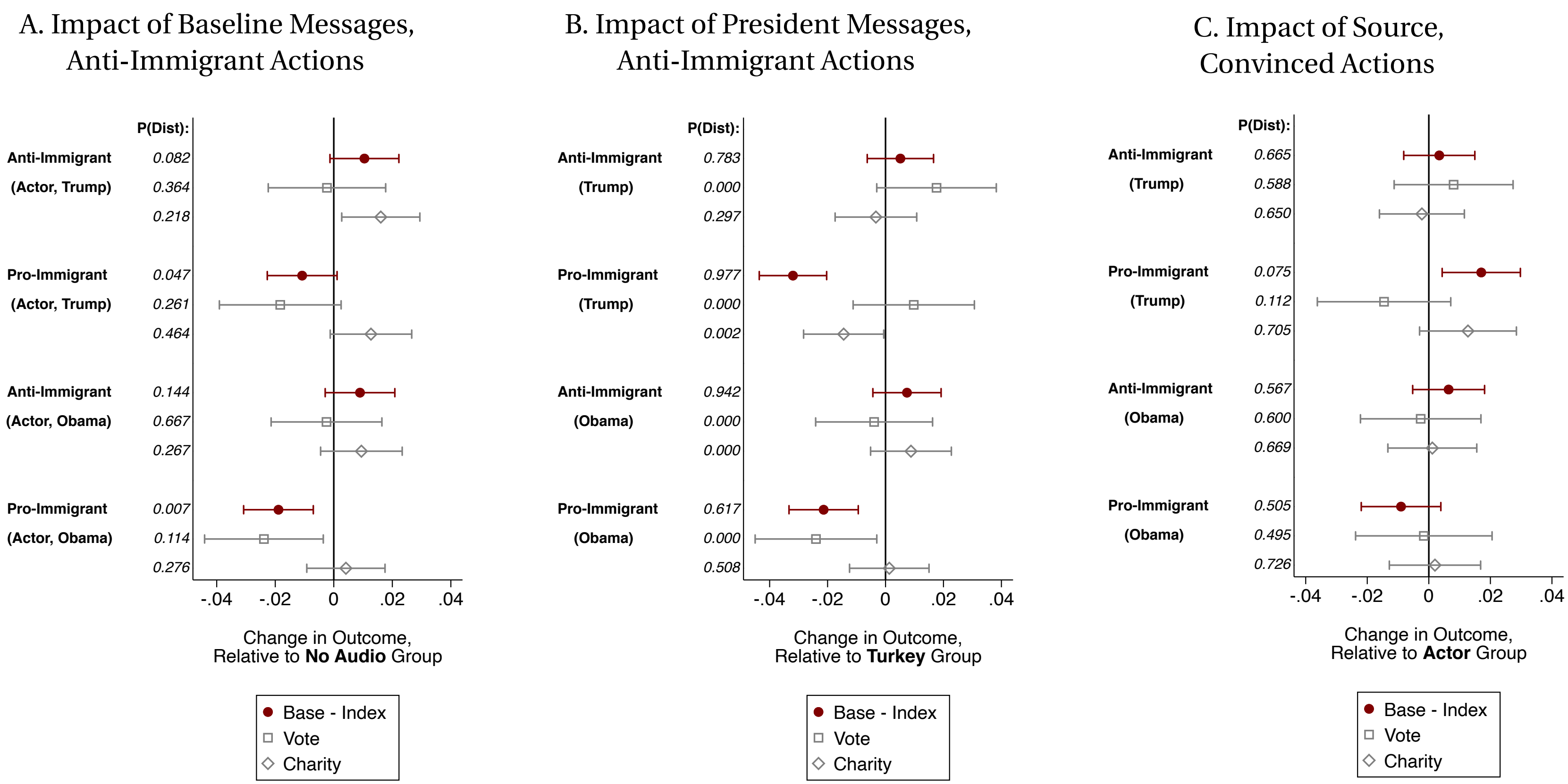
For the charity donation outcome, our total variation in the outcome is quite limited. Figure A28 plots the charity donation outcome in the control group of the study, where this outcome is normalized to be between 0 and 1. Here, 0 corresponds to a choice to donate the entire potential prize to the ACLU (pro-immigrant charity), 1 corresponds to the choice to donate the entire potential prize to FAIR (anti-immigrant charity), and 0.5 corresponds to the choice not to donate to either charity. Nearly all survey participants chose to award themselves the money (if they were to win). One reason that we may not find frequent responses to donate to one of the charities is that the prize may not be big enough, a constraint that we faced given our research budget. Effectively, the lack of variation we see in this outcome contributes to our limited ability to observe meaningful or significant changes in the charity donation choice after different treatments. It is not surprising that we find noisy and imprecise effects for this outcome given that nearly all participants chose to give themselves the prize, regardless of treatment status.

Figure A28: Charity Donation Outcome in No Audio Control Group



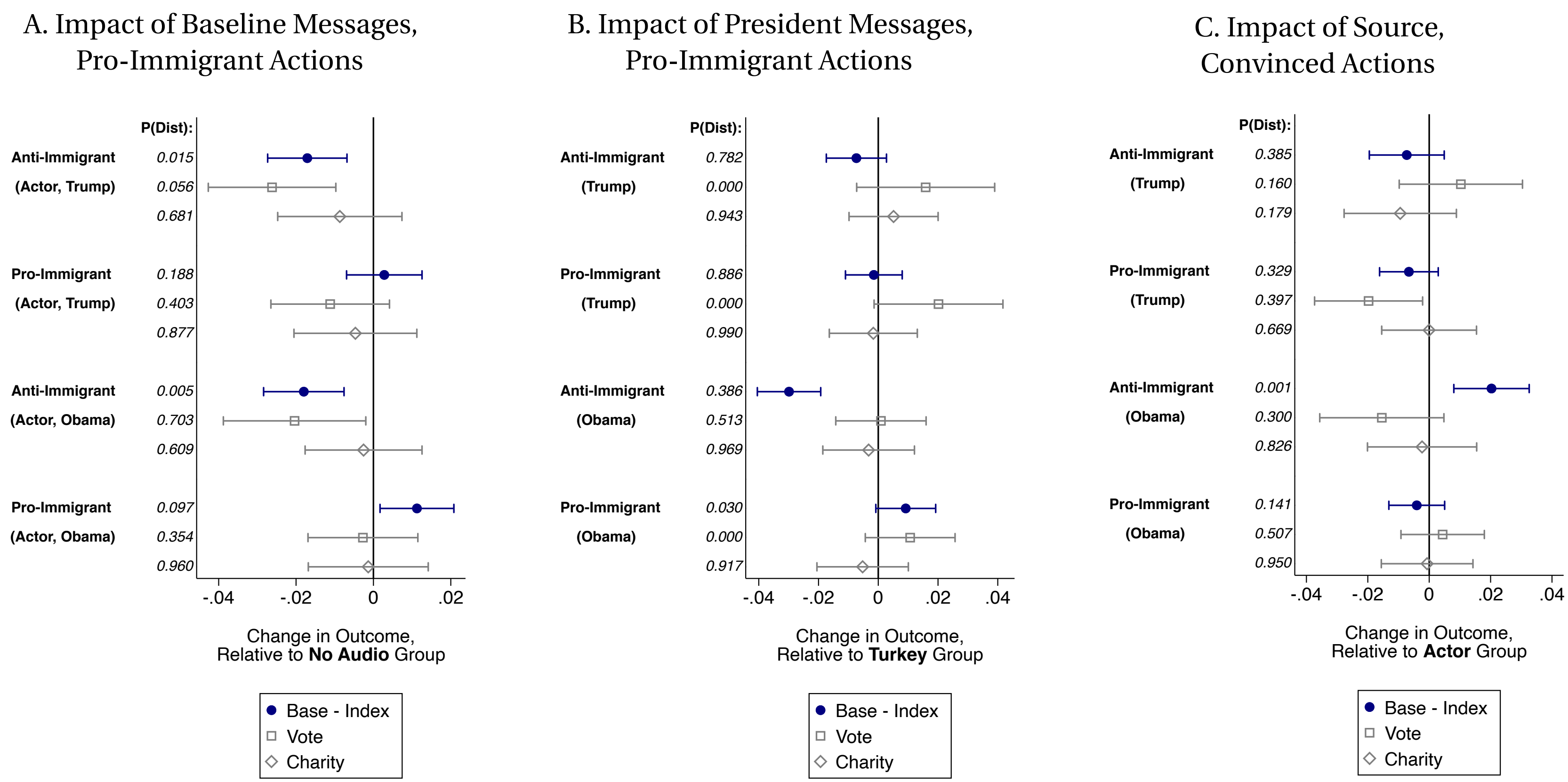
Notes: This plot shows the baseline distribution of the charity donation outcome in the no audio control group. The charity outcome is measured from a question where participants are told that they have been automatically entered into a lottery for $25, of which they can donate a portion to either an anti-immigrant charity, the Federation for American Immigration Reform (FAIR), or a pro-immigrant charity, the American Civil Liberties Union (ACLU) Immigrants' Rights Project, at a rate that will double any dollar that is donated. This outcome re-scales the donation choice based on these dollar donations to have a scale of $[0,1]$. A value of 0.5 corresponds to the participant electing not to make any donation.

Figure A29: Out of Sample Outcomes, Republicans



Notes: This plot shows the estimates for the baseline outcomes as compared to two out of sample outcomes. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. Vote is whether the participant stated that they intended to vote for Joseph Biden or Donald Trump in 2020, where a Biden vote is coded as pro-immigrant and a Trump vote is coded as anti-immigrant. This outcome is measured as $ln(1[Vote]+1)$. The charity outcome is measured from a question where participants are told that they have been automatically entered into a lottery for \$25, of which they can donate a portion to either an anti-immigrant charity, the Federation for American Immigration Reform (FAIR), or a pro-immigrant charity, the American Civil Liberties Union (ACLU) Immigrants' Rights Project, at a rate that will double any dollar that is donated. This outcome re-scales the donation choice based on these dollar donations to have a scale of $[0,1]$. For example, a fully anti-immigrant choice would be donating all of the lottery winnings to the anti-immigrant charity and a fully pro-immigrant choice would be donating all lottery winnings to the pro-immigrant charity. Again, this outcome is entered in the regression as $ln(Charity+1)$. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.

Figure A30: Out of Sample Outcomes, Democrats



Notes: This plot shows the estimates for the baseline outcomes as compared to two out of sample outcomes. The outcome is $ln(P(Outcome+1)$, where $P(Outcome)$ is either anti-immigrant or pro-immigrant index probability or probability convinced corresponds to the baseline outcome. Vote is whether the participant stated that they intended to vote for Joseph Biden or Donald Trump in 2020, where a Biden vote is coded as pro-immigrant and a Trump vote is coded as anti-immigrant. This outcome is measured as $ln(1[Vote]+1)$. The charity outcome is measured from a question where participants are told that they have been automatically entered into a lottery for $25, of which they can donate a portion to either an anti-immigrant charity, FAIR, or a pro-immigrant charity, ACLU, at a rate that will double any dollar that is donated. This outcome re-scales the donation choice based on these dollar donations to have a scale of $[0,1]$. For example, a fully anti-immigrant choice would be donating all of the lottery winnings to the anti-immigrant charity and a fully pro-immigrant choice would be donating all lottery winnings to the pro-immigrant charity. Again, this outcome is entered in the regression as $ln(Charity+1)$. All regressions are adjusted for covariates in Appendix A4. "P(Dist)" is the p-value from a Kolmogorov-Smirnov test of equality of distributions between the treatment and control groups for each estimate.