The Price of Intolerance: Asymmetries in Willingness to Suppress Opposing Views in the United States

Ritwik Banerjee[†] Lata Gangadharan[‡] Marie Claire Villeval[§]

August 12, 2026

Preliminary draft. Please do not circulate.

Abstract

Intolerance can undermine social cohesion by breeding division and animosity toward those holding different views. We study intolerance as a behavioral preference, measured by individuals' willingness to incur costs to prevent the spread of disapproved views. Using large-scale incentivized experiments with U.S. participants, we elicit this willingness across ideologically charged and neutral issues. A structural framework quantifies the intensity of intolerance and its sensitivity to ideological distance and audience size, while also examining personal and social norms. We show that this measure predicts consequential political behavior and document substantial heterogeneity across the ideological spectrum.

Keywords: Intolerance, Personal Norm, Ideology, Social Norm

JEL codes: C91, C93, D90

[*] *Acknowledgments*: We thank seminar participants in the University of Lund, TU Munich, University of Lyon, Ashoka University, the European University Institute, the Behavior and Experiment Economics Workshop at SUSTECH in Shenzen, the Workshop on Recent Advances in Behavioral Economics in Lyon, the 2026 SDU International Cooperation Forum on Behavioral and Experimental Economics at Shandong University. The paper has benefited greatly from comments from U. Gneezy, G. Genicot, Y. Chen, D. Rustagi, D. Mukherjee, A. Chandrasekhar, and D. Ray. This research has been approved by the IRB of Monash University (Approval # 41462). This research has been pre-registered at AEA Social Science Registry (pre-registration # AEARCTR-0013848). This research has benefited from the financial support of Monash University, and the LABEX CORTEX (ANR-11-LABX-0042) of Université Claude Bernard Lyon 1, within the Programme "Investissements d'Avenir" (2019-ANR-LABX-02) operated by the French National Research Agency (ANR).

[†]Indian Institute of Management Bangalore, India. Email: ritwikbanerjee@iimb.ac.in.

[‡]Department of Economics, Monash University, Clayton, VIC 3800 Australia. Email: lata.gangadharan@monash.edu.

[§]CNRS, Université Lumiére Lyon 2, Université Jean-Monnet Saint-Etienne, emlyon business school, GATE, 69007, Lyon, France; IZA@LISER, Luxembourg. Email: villeval@gate.cnrs.fr.

"*The golden rule of conduct, therefore, is mutual toleration, seeing we will never all think alike and that we shall always see Truth in fragment and from different angles of vision.*"

—Mahatma Gandhi, 1926

"*If all mankind minus one, were of one opinion, and only one person were of contrary opinion, mankind would be no more justified in silencing that one person, than he, if he had the power, would be justified in silencing mankind.*"

—John Stuart Mill, 1859

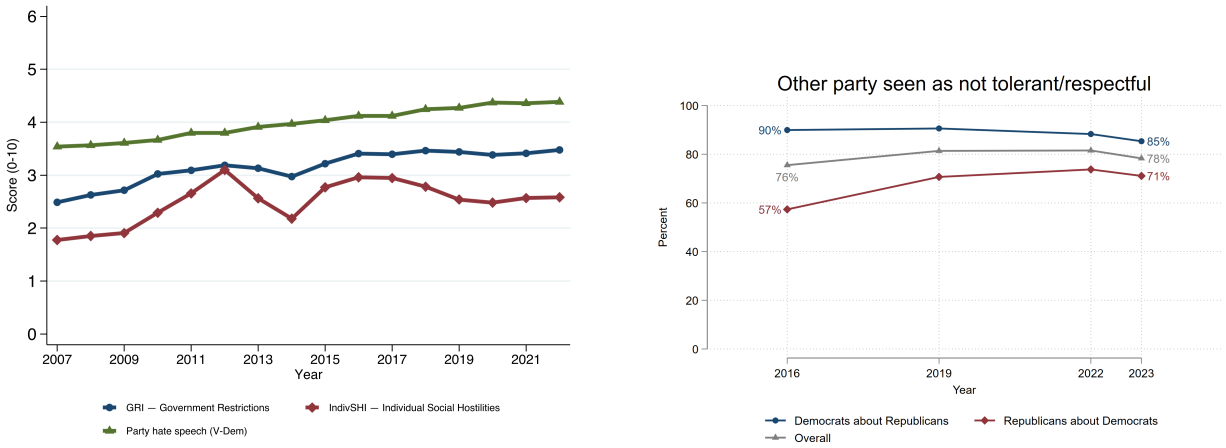
1 Introduction

Tolerance has long been regarded as a foundational virtue of open societies. European Enlightenment thinkers such as Locke and Spinoza grounded their defense of tolerance in arguments for protecting freedom of thought. Others, including Voltaire, Kant, and later John Stuart Mill, championed the use of reason and the right to argue, criticize, and dissent.[1] Similar ideas can be found in Eastern traditions. Gandhi insisted that mutual toleration is the only viable rule of conduct among people who will never think alike, particularly in a diverse society, while Confucius emphasized that harmony should take precedence over uniformity. Yet, there is growing concern across much of the world that the erosion of this value is now well underway.[2]

The recent rise of intolerance is reflected in broad cross-country trends. Panel (a) of Figure 1 shows that measures of political, governmental, and religious intolerance have all increased worldwide since 2007. While the magnitude and timing differ across indices, the common upward trend suggests that declining tolerance is neither limited to a particular country nor limited to a single sphere of public life. Consistent with this pattern, recent cross-country evidence documents growing divergence in tolerance and openness across nations (Jackson and Medvedev, 2024). Within countries, perceptions of intolerance are also on the rise. Panel (b) presents evidence from the United States, where large majorities of both Democrats and Republicans view the opposite party as intolerant. Perceptions of Democrats' intolerance among Republicans have risen sharply over time, whereas perceptions of Republicans' intolerance among Democrats remain persistently high—a salient feature of contemporary political polarization.

[1] Elizabeth Hall writes in *Friends of Voltaire*, "I disapprove of what you say, but I will defend to the death your right to say it,' was his [Voltaire's] attitude now." The quote is sometimes mistakenly attributed to Voltaire himself.

[2] Numerous episodes come to mind: the 1989 fatwa against Salman Rushdie over *The Satanic Verses* and the 2022 attack on him; the 2015 assault on *Charlie Hebdo* and the threats against *Jyllands-Posten* over their depictions of the Prophet Muhammad; the formal apparatus of blasphemy laws across South Asia and the Middle East; the expansion of campus speech codes in the United States; and the persecution of critical voices in India, Hungary, Turkey, and Russia. These developments coincide with the global rise of ethnonationalist forces and the deepening of partisan identities.



(a) Religious and political intolerance, 2007–2022

(b) Partisan perceptions of out-party intolerance, United States

Figure 1: Trends in intolerance: cross-country indices and US partisan perceptions

Notes: **Panel (a)** plots yearly cross-country averages of three intolerance indices, normalized to a common 0–10 scale. *Government intolerance* is measured using Pew's Government Restrictions Index, *religious intolerance* using a custom sub-index of Pew Social Hostilities items, and *political intolerance* using V-Dem's party-hate-speech measure. **Panel (b)** plots the share of U.S. partisans describing the out-party as *not* tolerant or respectful in Pew American Trends Panel surveys (2016-2023), separated by party and for the cross-party average.

Despite growing concern about intolerance, little is known about its behavioral foundations. To what extent are individuals willing to incur costs to suppress views they oppose? What determines the intensity of this willingness? Is it mainly driven by ideological divergences or by the fact that individual views can reach a massive audience on social media? We address these questions by developing a behavioral measure of intolerance and using it to quantify its prevalence, identify its underlying mechanisms, and examine how it differs across the political spectrum.

We contribute to three literatures. First, we introduce the first incentivized revealed-preference measure of intolerance, defined as individuals' willingness to incur monetary costs to prevent the dissemination of views they oppose. Existing work studies the political supply of hatred (Glaeser, 2005), tolerance in network link formation (Genicot, 2022), and willingness to engage with opposing views (Belot and Briscese, 2026); we instead quantify the demand to suppress such views. Second, we contribute to the economics of political and affective polarization (Gentzkow et al., 2019; Iyengar et al., 2019; Bursztyn et al., 2020; Boxell et al., 2024). Rather than measuring partisan affect or willingness to express controversial views, we measure the willingness to prevent the spread of opposing views and how it influences cooperation. Third, we complement studies on moral universalism (Enke, 2020; Enke et al., 2023, 2025). Instead of studying how altruism declines with social distance, we measure how intolerance varies with ideological distance.

Political scientists define tolerance as "the degree to which we accept things of which we disapprove" (Crick, 1971), or as "the belief that all citizens have the right to express their political

views, regardless of how dangerous or repugnant those views may be . . . " (Nie et al., 1996; Sullivan et al., 1979, 1982). These definitions imply that individuals must genuinely disapprove of a view, and they must recognize others' right to express it. Intolerance can therefore consist of a preference to prevent the expression—and thus the spread—of disapproved views.[3] This definition underpins our empirical measure. An alternative conceptualization is offered by Genicot (2022), who models tolerance as the maximum ideological distance an individual is willing to accept when forming social ties. Its central insight is that intolerance grows with the ideological distance between an opposing view and one's own ideal point. In our study, personal norms of appropriateness are used as a proxy for individuals' ideal points, allowing us to examine whether distance from ideal points is the key behavioral mechanism shaping intolerance.

If individuals have a preference to suppress views they dislike, they should be willing to incur real costs to do so. Building on this premise, we develop a simple theoretical model of intolerance in which willingness to pay to prevent the dissemination of opposing views depends on ideological distance, audience size, and baseline intolerance. The model delivers predictions that we can test experimentally, and it is structurally estimated using the experimental data. It also allows us to investigate whether this measure of intolerance predicts economically relevant behaviors, including cooperation, and whether it differs systematically between liberals and conservatives.

To measure this preference, we conduct two pre-registered experiments (Banerjee et al., 2026) with politically diverse U.S. participants recruited from Prolific. Participants evaluate a series of vignettes describing choices on salient political issues, including immigration, LGBTQ rights, abortion, and social security, as well as ideologically neutral issues. We first elicit participants' *personal norms* and their perceptions of prevailing *social norms*. We then measure their willingness to pay to prevent a vignette from being disseminated to other participants through an incentive-compatible mechanism. By randomly exposing liberals and conservatives to either ideologically congruent or incongruent views, our between-subject design identifies intolerance toward opposing viewpoints. Experiment 1 was conducted on October 30-31, 2024, immediately before the U.S. presidential election and includes a representative panel of 1009 participants.

Experiment 2 is a quasi-replication that extends the design in three directions. First, a *spread treatment* randomly varies the size of the audience exposed, by default, to the message (one vs. 100 persons), allowing us to decompose the willingness to pay into a component driven by the mere exposure to an opposing view and a component driven by concern about its dissemination. Second, participants play a *prisoner's dilemma game*, enabling us to examine whether the counterpart's intolerance predicts cooperation. Third, we elicit the motives underlying intolerance. Experiment

[3] A weaker conception of tolerance grants others only the right to *hold* rather than *express* objectionable views. Under that conception, intolerance would involve preventing beliefs themselves. Because such a notion is difficult to operationalize, we adopt the more tractable definition based on suppressing expression.

2 was conducted in January 2026 with 1007 participants from another representative Prolific sample.

We also assess the external validity of our measure by randomly exposing participants to authentic Democratic and Republican fundraising posters, and measuring their willingness to block their dissemination. Finally, we use our incentivized measure of intolerance to validate scalable hypothetical measures based on both a survey and a staircase procedure.

Experiment 1 yields three main findings. First, intolerance exists and is widespread. 61% of participants exhibit a strictly positive willingness-to-pay to block ideologically salient views, and they are willing to forgo, on average, 21.9% of their endowment to that end. Second, ideological distance matters. Participants exposed to views incongruent with their own ideology pay $1.42 more to block them than those exposed to congruent views. Third, personal and perceived social norms are both associated with the willingness to pay to block the spread of a view, but personal norms are the main predictor of behavior, consistent with the findings of Bašić and Verrina (2024).

The findings broadly replicate in Experiment 2, which provides three additional insights. First, concern about dissemination is negligible, as shown by a structural estimation which decomposes willingness to pay into a baseline sensitivity to disagreement, a curvature term capturing sensitivity to ideological distance, and sensitivity to dissemination. Ideological distance and its curvature are the primary determinants of willingness to pay. Second, intolerant behavior generalizes in a consequential environment. The willingness to pay to prevent the dissemination of experimental vignettes predicts blocking authentic fundraising posters from the opposing political party. Intolerance has also economically meaningful consequences, as participants cooperate 71% less with intolerant partners in a prisoner's dilemma.

Third, structural estimates show that liberals and conservatives display similar baseline aversion to disagreement, but willingness to suppress opposing views rises much more steeply with ideological distance among liberals than conservatives. Context-free staircase and survey-based measures also show that liberals score higher than conservatives on intolerance, consistent with the asymmetry documented in the Pew data presented in panel (b) of Figure 1.

The paper proceeds as follows. Section 2 describes our contributions to the literature. Section 3 describes the experimental design. Section 4 presents the theoretical framework and derives testable predictions. Sections 5 and 6 report results from the two experiments, Section 7 presents the structural estimation and predictive validity of our approach, and Section 8 compares liberal's and conversatives' levels of intolerance. Finally, Section 9 concludes.

2 Literature Review

Our paper contributes to three literatures. First, it relates to the emerging economics of intolerance. In particular, Genicot (2022) characterizes how tolerance shapes network formation, and Belot and Briscese (2026) experimentally study individuals' willingness to *listen* to opposing views.

They find that two-thirds of respondents voluntarily expose themselves to ideologically contrary perspectives when given an opportunity to opt out, and emphasizing common ground raises both engagement and openness to reconsidering one's views. We measure the opposite margin. Rather than asking whether individuals are willing to consume opposite views, we ask whether they are willing to incur private costs to prevent others from being exposed to them. Together, the two approaches characterize both the demand for and the supply of cross-ideological exposure.

Second, our paper speaks to the literature on political polarization. In its classical economic formulation, polarization is a *relational* property of a population—a distribution of attributes that collapses into a small number of internally homogeneous, mutually antagonistic clusters (Esteban and Ray, 1994; Duclos et al., 2004)—and has been most directly linked to inter-group conflict (Esteban and Ray, 2011). A parallel strand documents *positional* polarization, that is, the growing divergence in the average policy positions of partisan groups (Gentzkow et al., 2019). A more recent strand distinguishes ideological polarization from *affective* polarization—the gap between in-party warmth and out-party animus (Iyengar et al., 2019; Boxell et al., 2024). This affective gap has been tied to media and platform exposure (Martin and Yurukoglu, 2017; Levy, 2021; Allcott et al., 2020) and to identity salience (Bonomi et al., 2021). It therefore provides a natural environment for intolerance to arise.

We contribute to this literature in several ways. We draw a sharp conceptual distinction between polarization, which describes the distribution of views in a population, and intolerance, which is an individual's behavioral disposition toward an out-group. A polarized society can still be tolerant if rival clusters respect one another's civic standing (e.g., Boxell et al., 2024). Rather than leading to open conflict, the distance between clusters may manifest through the costly suppression of opposing views. Moreover, our measure of intolerance is the mirror image of Bursztyn et al. (2020), who elicit individuals' willingness to *express* socially stigmatized views. Whereas they study the private cost of expression, we study the private costs of *suppressing* others' expression. Both approaches rest on the distinction between personal and social norms, but they capture opposite sides of ideological conflict. Finally, our approach supplies a behavioral micro-foundation for the affective gap that the polarization literature has so far documented at the aggregate level.

Third, our paper is closely related to the literature on *moral universalism*. Building on Tabellini (2008), Enke (2020) and Enke et al. (2023, 2025) conceptualize universalism as the *slope* of altruism and trust over social distance: universalists extend the same generosity and trust to socially distant strangers as to kin, whereas communitarians' altruism falls steeply with distance. This slope predicts voting behavior (Enke, 2020), the correlation structure of policy views (Enke et al., 2023), and cross-country variation in attitudes (Enke et al., 2025). We study an analogous object in the domain of speech. Instead of the slope of altruism with respect to social distance, we estimate the slope of intolerance with respect to *ideological distance*—the distance between an individual's

ideal point and an opposing view. This analogy provides a common framework for thinking about moral universalism across both social and epistemic domains.

Together, these literatures leave open the central question we address: whether individuals are willing to incur private costs to suppress opposing views, and what determines that willingness.

3 Experimental Design

Studying intolerance requires an experimental environment with two key features: (i) participants must be exposed to a view they may disapprove of, and (ii) they can choose whether to prevent the dissemination of those views. We implement these features in two randomized experiments. An independent validation study has been used to verify that our experimental stimuli were perceived as intended. We describe each in turn.

3.1 Experiment 1

Vignettes. Central to our experimental design is our use of six *vignettes* shown in Appendix D. These vignettes present participants with a series of hypothetical situations that describe a fictitious person's dilemma about a decision that needs to be made. The dilemma resolves either conservatively or liberally. Consider, for example, the vignette that describes the dilemma faced by a voter in an abortion referendum:

> *Maria sat in her room, the leaflet that had been distributed glaring back at her. The upcoming referendum on abortion was fast approaching, and she had to choose a side: pro-life or pro-choice. Her mind raced –what did she really believe in? Pro-life echoed her concern that human life begins at conception, but pro-choice resonated with the unflinching belief in personal freedom and autonomy. She knew she would have to take a stand and vote, but she was conflicted.*

This dilemma has two possible resolutions: a liberal resolution ("*After much deliberation, she decided to vote in favor of Choice.*"—Treatment Liberal, L henceforth) or a conservative resolution ("*After much deliberation, she decided to vote in favor of Life.*"—Treatment Conservative, C henceforth), randomly assigned across participants. Combined with participants' own ideological orientation, this yields four possible variations: liberals or conservatives exposed to either a congruent or an incongruent resolution. Thus, in a between-subject design, participants are randomly exposed to views that either align with or contradict their ideological position.

Our experiment involves six such vignettes. Four are ideologically salient, describing situations on issues that are known to generate polarized views in the American context, while two are

ideologically neutral. The ideologically salient vignettes describe decisions related to (i) hiring a homosexual teacher at a school, (ii) voting at an abortion referendum, (iii) signing a petition on social security benefits, and (iv) signing an anti-immigration petition. The two ideologically neutral vignettes allow us to isolate the role of descriptive norms from ideological disagreement. If participants are willing to prevent the spread of a vignette even when the issue is ideologically neutral, this would suggest that descriptive norms alone can motivate the willingness to pay. The neutral vignettes concern (v) whether to buy trousers or a shirt, and (vi) whether to buy a blue or a beige shirt. The two neutral vignettes differ in one important respect. The choice between trousers and a shirt is unlikely to evoke either descriptive or injunctive norms. In contrast, the choice between shirt colors may involve descriptive norms (e.g., one color being more popular than another), but it is unlikely to involve injunctive norms, as selecting a particular color rarely attracts social approval or disapproval.[4]

Validation of the vignettes. To ensure that the ideologically salient stimuli accurately represent liberal and conservative positions on each issue, we conducted an independent online validation study on August 28, 2024, before the main experiment. Participants were randomly assigned to the vignettes used in either the Treatment Conservative or the Treatment Liberal. They were first rated the ideological position of the decision maker and the associated view on a 0-10 scale in response to this question: *"In social or political matters, people often talk of 'the left' and 'the right'. How will you rate <decision maker's> decision and the view expressed in this post on this scale?"*. Next, using an incentivized coordination game, they reported their second-order beliefs by predicting the most frequent response to the same question among ten randomly 10 randomly selected participants. Correct predictions were rewarded with $1.50. These responses allow us to verify that the liberal and conservative resolutions are perceived as intended, both in participants' own ideological assessments and in their second-order beliefs. Table B.2 in the Appendix confirms that the two treatments differ significantly along both dimensions.

Timeline. Experiment 1 was conducted online from October 30-31, 2024. The top panel of Figure 2 displays the timeline of the experiment that we now describe in detail.

Personal norm elicitation (part 1). Each participant is shown the six vignettes of a given treatment (either C or L) in a random order. After each vignette, participants report their personal norm regarding the character's decision in the vignette on a five-point appropriateness scale. Specifically, they indicate whether, in their own and irrespective of others' opinions, the decision is

[4]Blue and beige were chosen because blue is consistently the most popular color in the US ($https://yougov.com/articles/12336-why-blue-worlds-favorite-color-1?utm_source=chatgpt.com$; accessed on August 4, 2026), whereas beige is relatively neutral.

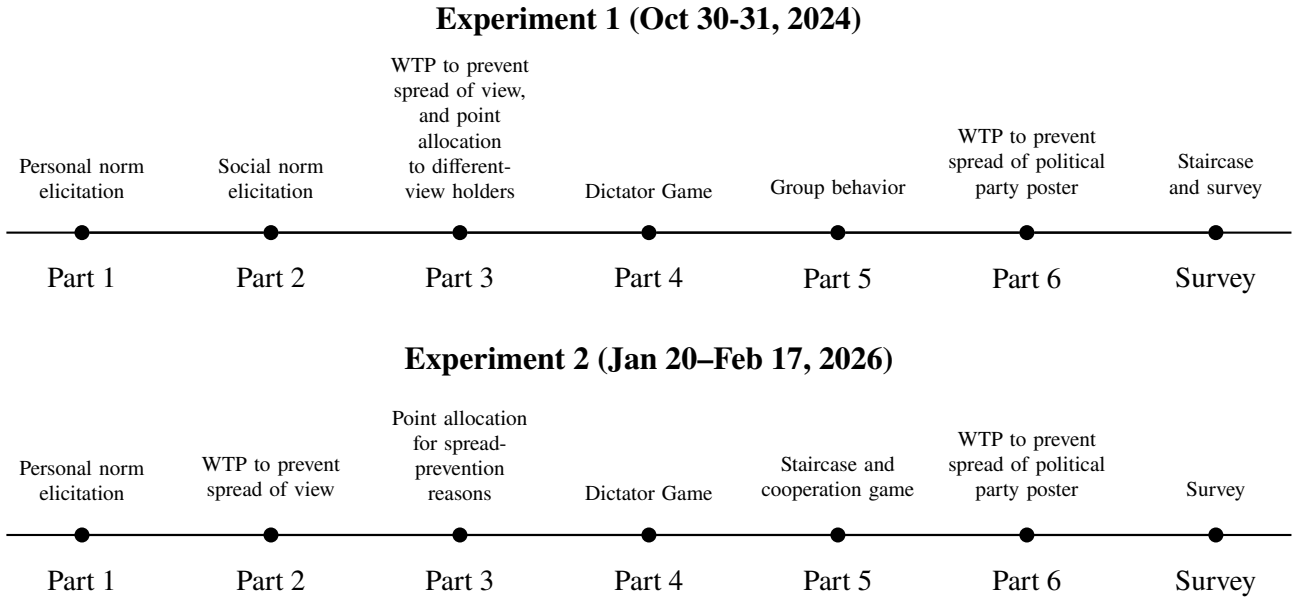


Figure 2: Timeline of Experiment 1 and Experiment 2

either very appropriate, appropriate, neither appropriate nor inappropriate, inappropriate, or very inappropriate. This measure is not incentivized.

Social norm elicitation (part 2). Next, participants are shown the same vignettes in a random order. Following each vignette, we elicit second-order beliefs, using the coordination game of Krupka and Weber (2013). Participants guess which response on the appropriateness scale described above is most frequently chosen by other participants in general, among liberals, and among conservatives. If this part is selected for payment at the end, we randomly select one vignette and pay a $4 bonus if the participant's guess is correct. These incentivized guesses provide measures of the perceived social norms associated with the character's decision.

Willingness-to-Pay elicitation (part 3). This part measures our main outcome: intolerance. Participants are endowed with $10. They are shown each of the same six vignettes along with a view always consistent with the character's decision, presented in a format resembling a Facebook post, like those used in the validation pre-experiment with different participants (see Figure B.1).[5] They are told that, by default, the post will be circulated among 100 other people. They can prevent its spread by paying any amount between 0 and $10; any amount not spent is kept as earnings. We elicit the willingness to pay to prevent the spread of the post using a Becker-DeGroot-Marschak

[5]For example, the views added to the abortion vignette were "*In my view, abortions should be allowed because it is a human right, a personal choice, and a health issue. No one should force or judge women's decisions.*" (Treatment L), and "*In my view, abortion or taking away life should not be allowed because it is a moral crime and a health risk. Everyone should protect and respect life from conception.*" (Treatment C).

(BDM) mechanism (Becker et al., 1964). A random number is drawn, and if the participant's stated willingness to pay is at least as high as that number, they pay the drawn amount, and the post is not circulated; otherwise, the post is circulated, and they pay nothing. The stated willingness to pay provides our measure of intolerance—the desire to prevent the spread of a view.

A willingness to pay of zero is potentially ambiguous: it may reflect tolerance or simply a reluctance to sacrifice one's own earnings. To distinguish between these motives, participants reporting a zero willingness to pay for an ideologically salient vignette are asked whether they would prevent the dissemination of the view in exchange for \$0.5. If they decline, the offer is increased to \$1. Acceptance of either offer indicates that the initial zero willingness to pay reflected self-interest rather than tolerance.

Dictator Game (part 4). Next, participants play a Dictator Game to help identify the motives underlying intolerance. A positive willingness-to-pay to suppress a view may reflect a genuine preference to prevent its dissemination, but it may also be driven by more general social preferences, such as selfishness or spite. To disentangle these mechanisms, participants start with 50 cents and are endowed with \$10. They all decide as a sender on how much to transfer to an anonymous receiver. Roles (sender or receiver) are randomly determined for payment at the end of the experiment. The amount transferred provides a measure of prosociality that we relate to participants' willingness to pay to prevent the spread of posts. To further capture anti-social preferences, participants who send zero are asked if they would choose to reduce the receiver's payoff by 50 cents.

Groupism (part 5). Finally, we measure groupism following Bauer et al. (2023). Participants are randomly assigned to one of two groups, Circle or Triangle, and allocate \$6 between two *other* individuals. They make allocation decisions in three scenarios: (i) both recipients belong to the participants' own group; (ii) both belong to the other group; and (iii) one recipient belongs to the participant's group and the other to the out-group. Including all three scenarios provides a more complete elicitation of allocation preferences and makes the purpose of the task less transparent, reducing the salience of group favoritism. Our measure of groupism is based on the third scenario, as it directly captures the extent to which participants favor an in-group member over an otherwise comparable out-group member.

Field behavior (part 6). To predict whether our experimental measure of intolerance predicts behavior in a more naturalistic setting, participants are shown a political fundraising poster containing information on how to donate to a political party in the U.S.. By default, the poster would be circulated to 100 other persons, but participants could prevent its dissemination by paying money. As in part 3, participants are endowed with \$10 and willingness to pay is elicited using a BDM

mechanism. Participants are randomly assigned to view either a poster supporting fundraising for the Democratic Party or the Republican Party. Thus, liberals and conservatives are each exposed to either an ideologically congruent or incongruent poster, allowing us to estimate intolerance toward politically aligned or opposing messages. The experiment was conducted three days before the U.S. presidential election. At that time, political fundraising messages were highly salient and directly relevant, making the decision to prevent their spread both symbolically and materially consequential.

Sample. The validation pre-experiment was conducted with a representative sample in terms of age, sex, and political affiliation of 297 participants recruited in the United States through Academic Prolific's representative sample option.[6] Experiment 1 was conducted with a separate sample of 1009 citizens from the U.S., also recruited through Prolific's representative sample option. Descriptive summary statistics are reported in Appendix Table B.1 for the validation pre-experiment and in Appendix Table B.3 for the full sample in Experiment 1.

On average, Experiment 1 lasted 44 minutes, and participants earned $4.84 in addition to a $5 fixed earnings. Variable earnings were determined by randomly selecting one of parts 2-6 for payment.

3.2 Experiment 2

Experiment 2 serves as a quasi-replication and extension of Experiment 1. Conducted between January 20 and February 17, 2026, it retains the core design used to measure intolerance, allowing us to assess the robustness of our main results in a context different from that of the presidential election. It extends the analysis in three important ways. First, it varies the size of the potential audience. Second, it investigates the reasons underlying individuals' willingness to pay to prevent the spread of opposing views. Third, it examines the economic consequences of intolerance through an incentivized social dilemma game. The timeline of Experiment 2 is presented in the bottom panel of Figure 2.

Design modifications. Experiment 2 employs a 2×2 between-subjects design. On the first dimension, as in Experiment 1, participants are randomly assigned to either Treatment Conservative (C), which presents vignettes with conservative resolutions, or Treatment Liberal (L), which presents vignettes with liberal resolutions. On the second dimension, we vary the default reach of the Facebook-style posts: in the *Spread-1* treatment, posts are shared with one person by default, whereas in the *Spread-100* treatment, posts are shared with 100 people by default. This variation

[6] For choosing the colors in one neutral vignette, we conducted a separate experiment for which the sample size was 173.

allows us to test whether the potential audience size affects individuals' willingness to pay to block the dissemination of opposing views. The experiment includes four ideologically salient vignettes. Three are carried over from Experiment 1, while one is new.[7] Unlike Experiment 1, we do not include ideologically neutral vignettes, as their primary purpose was to provide a control condition.

To investigate the motives underlying intolerance, after participants report their personal norms in part 1 and willingness to pay to block the dissemination of the posts in part 2, part 3 elicits two additional measures for each vignette. First, participants rate, on a 0-10 scale, how likely they believe the post is to influence those who see it. Second, they report the reasons underlying their decision to pay or not to pay to block the post. Specifically, participants allocate a total of 100 points across up to six possible motivations. For participants with a positive willingness to pay, these motivations include moral repugnance toward the view, strategic concerns about the potential harm it could cause to themselves, their family, businesses, or society at large, the perception that the post itself is intolerant, and the belief that it conveys false information. Participants with a zero willingness to pay are presented with the corresponding absence of these concerns (e.g., no ethical concerns, no perceived harm, no intolerance). This procedure allows us to decompose intolerance into its ethical and consequentialist components.[8]

To investigate the economic consequences of intolerance, part 5 introduces an incentivized Prisoner's Dilemma game.[9] Participants are informed that they will be matched with another participant who is likely to hold different social or political views. Before making their cooperation decision, they are told whether their partner had previously reported a relatively high willingness to pay in the abstract context presented in the next subsection, or no or only a relatively low willingness to pay, to prevent the dissemination of a view that the participant supports. Each participant is sequentially matched with two partners: one classified as intolerant and one classified as tolerant. This within-subject design allows us to test whether individuals are less willing to cooperate with someone who has displayed intolerance toward their own views.

As in Experiment 1, part 6 elicits participants' willingness to pay, using a BDM mechanism, out of a $10 endowment to prevent the dissemination of a real political fundraising poster supporting

[7] We replaced the vignette concerning the recruitment of a homosexual teacher with one describing a manager deciding whether to distribute free coffee mugs supporting LGBTQ rights to employees (see Appendix XXX). We realized that the original vignette could be interpreted as depicting illegal discriminatory behavior. A new validation experiment tested only this new vignette with a separate sample of 273 participants recruited through Prolific and confirmed its conservative or liberal resolution across the two treatments; see Appendix Table C.2.

[8] Unlike in Experiment 1, we do not elicit social norms in Experiment 2. Experiment 1 shows that personal norms have greater explanatory power than social norms, as discussed below, allowing us to omit the latter and devote the additional time to eliciting the motives underlying participants' willingness to pay. For the same reason, while we retain the Dictator Game in part 4, we omit the groupism task included in Experiment 1 to keep the overall duration of the experiment manageable.

[9] If both players cooperate, each earns $4. If both defects, each earns $2. If one defects while the other cooperates, the defector earns $3 and the cooperator earns $0.

either the Democratic or the Republican Party in the U.S. (see Appendix Figure C.3). Random assignment determines which poster participants see. Depending on the treatment, the poster is, by default, disseminated to either 1 or 100 other Prolific participants recruited at the end of the experiment.

Sample. Experiment 2 was conducted with a new representative sample from the United States, recruited through Academic Prolific, excluding participants from Experiment 1, and matching the target population on age, sex, and political affiliation. Consistent with the pre-registered exclusion criteria, the experimental platform screened out individuals who (i) indicated a political view of 5 on the 0-10 ideological scale, (ii) identified as neither a Democrat nor a Republican, or (iii) reported a political view on the ideological scale inconsistent with the party for which they voted. The final sample comprised 1007 participants, of whom 503 classified themselves as liberals and 504 as conservatives. Descriptive summary statistics are reported by treatment in the Appendix Table C.1 for the validation experiment and in Appendix Tables C.3 and C.4 for Experiment 2.

On average, Experiment 2 lasted 43 minutes, and participants earned $4.92, in addition to a $5 fixed earnings. Variable earnings were determined by randomly selecting one of parts 2 and 4-6 for payment.

3.3 Context-free measures of intolerance

Our two controlled experiments provide an incentivized, context-specific measure of intolerance. A secondary objective of our study is to develop and validate context-free, portable measures that predict this incentivized measure. To this end, we implemented two complementary approaches in the two experiments.

Survey-based measures. We elicit four survey-based measures to capture different aspects of intolerance. Participants have to indicate on an 11-point scale: (1) whether it is important to be open to views opposite to one's own or fine to ignore such views; (2) whether friends or relatives think of the participant as someone who shows very little or very high tolerance; (3) whether it is important to let people spread views opposite to one's own or whether such views should be prevented; and (4) how likely the participant is to be friends with someone who holds a different view on an important matter. These questions were complemented with questions on participants' perceptions of intolerance in society.

Staircase measure. In addition to these survey questions, we use a staircase method to construct a context-free measure of intolerance that ranks people according to their tolerance level. Participants first read the following text in Experiment 1: *People have different views on issues such as rights*

of LGBTQ+ individuals, or political matters, immigration policies, and so on. Imagine a situation where you disagree with a certain view, and the view of the person will be circulated to 100 other people. But you can prevent the spread of that view (with which you disagree), by some monetary transaction. We will present you with four questions, aimed at understanding how much you will be willing to pay to prevent the spread of that view. In Experiment 2, the default for the dissemination is either 1 or 100 persons. Participants are then asked whether they would prefer to receive $10 and be given the opportunity to prevent the spread of a view (arm A), or receive $10 without that opportunity (arm B). The subsequent binary choices are contingent on the previous answer, progressively narrowing the participant's hypothetical willingness to pay. This staircase procedure classifies participants into one of 16 levels of intolerance. The complete staircase design is shown in the Appendix Figure **??**.

3.4 Final questionnaire and ideology classification

Our experimental design relies on distinguishing whether participants are exposed to views that are ideologically congruent or incongruent with their own position. To classify participants' ideological position, we elicit three complementary measures. First, participants answer a battery of ten policy questions covering abortion, transgender rights, health care, immigration, welfare spending, military spending, affirmative action, gun control, taxes, and equal pay, with responses recorded on a 7-point Likert scale from "strongly oppose" to "strongly favor".[10] Second, participants place themselves on an eleven-point left-right scale, where the left corresponds to a liberal position and the right to a conservative position. Finally, participants report their voting preference between the Democratic and Republican parties. Although each measure has limitations, they allow us to perform numerous robustness checks to verify that our results are not driven by any particular definition of 'ideological position'.

4 Theoretical Framework and Predictions

We develop a formal model of ideological intolerance that generates the hypotheses tested experimentally. The model captures three key features: (i) individuals hold ideological views; (ii) exposure to differing views generates psychological disutility; and (iii) individuals can pay to suppress their dissemination. It predicts how intolerance varies with ideological distance, the scale of dissemination, and the nature of the underlying issue.

[10]This approach to estimating ideal points was pioneered by Poole and Rosenthal (1985) for members of the U.S. Congress. Since then, the battery of questions has been widely applied to estimate citizens' ideal point in several contexts (Clinton et al., 2004; Hill and Tausanovitch, 2015; Ansolabehere et al., 2008).

4.1 The Model

Consider a population of individuals indexed by $j \in N$. Each individual holds a view $v_j \in R$ on a given issue, interpreted as their *personal norm*—their belief about what constitutes appropriate behavior. Views lie on an ideological spectrum where lower values correspond to more liberal positions and higher values to more conservative positions.

Individual j derives baseline utility $g(v_j) > 0$ from holding this view. Following Akerlof and Kranton (2000), we interpret $g(v_j)$ as *identity utility*: individuals derive psychological value from having a coherent value system that defines their identity. The function $g(\cdot)$ captures the idea that views are not merely instrumental beliefs but constitutive of an individual's self-concept.

Disutility from opposing views. When individual j is exposed to another individual i's view v_i, she experiences disutility whenever $v_i \neq v_j$. This disutility has two sources:

- *Exposure:* The mere act of encountering a different view creates psychological discomfort, captured by $h(|v_i - v_j|)$, where $h : R_+ \to R_+$ is increasing and satisfies $h(0) = 0$.
- *Spread:* Beyond direct exposure, individual j may dislike the prospect that v_i is disseminated to n_i other people. This additional distutility scales with both ideological distance and audience size, and is given by $h(|v_i - v_j|) \cdot \psi(n_i)$, where $\psi : N \to R_+$ is increasing in n.

Combining these components, total disutility is:

$$D_j(v_i, n_i) = \alpha_j \cdot h(|v_i - v_j|) \cdot [1 + \psi(n_i)] \tag{1}$$

where $\alpha_j > 0$ is individual j's *intolerance preference parameter*, capturing her baseline sensitivity to ideological disagreement.

Assumption 1 (Properties of disutility components)**.** *The functions $h(\cdot)$ and $\psi(\cdot)$ satisfy:*

(i) $h(0) = 0$ *and* $h(d) > 0$ *for all* $d > 0$ *(disagreement induces disutility);*

(ii) $h'(d) > 0$ *(disutility increases in ideological distance);*

(iii) $\psi(n) \geq 0$ *for all* $n \geq 1$ *with* $\psi'(n) > 0$ *(concern about dissemination is non-negative and increasing in audience size)*

Note that even when $n_i = 1$ (the view reaches only one other person), disutility is positive as long as $\psi(1) \geq 0$. The term $[1 + \psi(n_i)]$ ensures that exposure alone generates disutility, while spread amplifies it.

A natural functional form is:

$$h(d) = d^{\gamma}, \quad \psi(n) = \beta \cdot (n-1) \tag{2}$$

where $\gamma > 0$ captures the curvature of disutility with respect to ideological distance ($\gamma = 1$ linear, $\gamma > 1$ convex, $\gamma < 1$ concave), and $\beta \geq 0$ captures sensitivity to audience size. When $\beta = 0$, only exposure matters; when $\beta > 0$, spread amplifies disutility.

Substituting into Equation (1), the complete functional form for disutility is:

$$D_j(v_i, n_i) = \alpha_j \cdot |v_i - v_j|^{\gamma} \cdot [1 + \beta(n_i - 1)] \tag{3}$$

Willingness to pay to prevent dissemination. Individual j has the option to block the dissemination of view v_i. Blocking is a binary decision: either j blocks, preventing any dissemination of v_i, or she does not block.

- If j does not block, she experiences disutility $D_j(v_i, n_i)$ from the spread of an opposing view.
- If j blocks, she avoids the disutility from spread but incurs the monetary cost c.

Individual j blocks if and only if the cost of blocking is less than the disutility avoided:

$$\text{Block} \iff c \leq D_j(v_i, n_i) \tag{4}$$

Her willingness to pay (WTP) is the maximum c she would accept—the threshold at which she is indifferent between blocking and not blocking.[11] Setting $c^* = D_j(v_i, n_i)$ gives the indifference point.

Definition 1 (Willingness to Pay for Blocking)**.** *Individual j's willingness to pay to block view v_i from spreading to n_i people is:*

$$WTP_j(v_i, n_i) = D_j(v_i, n_i) = \alpha_j \cdot h(|v_i - v_j|) \cdot [1 + \psi(n_i)] \tag{5}$$

Ideologically salient vs. ideologically neutral issues. Not all disagreements generate intolerance. We distinguish between two types of issues. (i) Ideologically neutral issues: Differences in views reflect personal preferences with no broader significance (e.g., preferred shirt color). (ii) Ideologically salient issues: Differences in views carry moral, political, or social weight (e.g.,

[11] While the blocking decision is binary, WTP is continuous: it measures the monetary value j assigns to preventing the spread of an opposing view. In the experiment, we elicit this WTP directly by asking participants how much they are willing to pay to block each view.

immigration policy) and may generate psychological disutility. We capture this distinction by allowing the intolerance parameter to be issue-specific:

$$\alpha_j(\theta) = \begin{cases} \alpha_j > 0 & \text{if } \theta \text{ is ideologically salient} \\ 0 & \text{if } \theta \text{ is ideologically neutral} \end{cases} \tag{6}$$

When $\alpha_j(\theta) = 0$, Equation (5) implies $\text{WTP}_j = 0$ regardless of the ideological distance $|v_i - v_j|$. Disagreement on ideologically neutral issues generates no psychological disutility, and hence no willingness to pay to block.

4.2 Predictions

We now derive the model's formal predictions that form the main hypotheses tested in our experiments. All formal proofs are available in Appendix A.

Proposition 1 (Existence)**.** *If $\alpha_j > 0$ and individuals are exposed to views where $v_i \neq v_j$, then $WTP_j(v_i, n_i) > 0$.*

This proposition motivates **Hypothesis 1 (Existence)**: Individuals exhibit positive willingness to pay to block opposing views.[12]

Next, consider a liberal individual j^L with view $v_j^L < \bar{v}$ (below the median) and a conservative individual j^C with view $v_j^C > \bar{v}$ (above the median). Let v^L denote a liberal view and v^C a conservative view, where $v^L < \bar{v} < v^C$.

Proposition 2 (Ideological congruence)**.** *For a liberal individual j^L:*

$$WTP_{j^L}(v^C, n) > WTP_{j^L}(v^L, n) \quad \forall n \geq 1 \tag{7}$$

Symmetrically, for a conservative individual j^C:

$$WTP_{j^C}(v^L, n) > WTP_{j^C}(v^C, n) \quad \forall n \geq 1 \tag{8}$$

Hypothesis 2 (Liberal Incongruence): Liberal participants exposed to conservative views exhibit higher intolerance than liberal participants exposed to liberal views.

Hypothesis 3 (Conservative Incongruence): Conservative participants exposed to liberal views exhibit higher intolerance than conservative participants exposed to conservative views.

[12] Since α_j can vary across individuals, the willingness to pay may vary across individuals exposed to identical views, with individuals with higher α_j exhibiting also higher WTP across all views.

Proposition 3 (Ideological Distance)**.** *For any individual j and two views v_i and v_k with $|v_k - v_j| > |v_i - v_j|$:*

$$WTP_j(v_k, n) > WTP_j(v_i, n) \quad \forall n \geq 1 \tag{9}$$

Hypothesis 4 (Ideological Distance): Greater deviation of a view from one's personal norm leads to greater intolerance.

Proposition 4 (Spread)**.** *For any individual j, view $v_i \neq v_j$, and audience sizes $n' > n$:*

$$WTP_j(v_i, n') > WTP_j(v_i, n) \tag{10}$$

To interpret the treatment comparison, consider the linear specification $\psi(n) = \beta(n-1)$. The difference between treatments,

$$\text{WTP}_j(v_i, 100) - \text{WTP}_j(v_i, 1) = 99\beta \cdot \alpha_j \cdot h(|v_i - v_j|) \tag{11}$$

isolates the marginal value of preventing dissemination. If $\text{WTP}(100) \approx \text{WTP}(1)$, then $\beta \approx 0$ and individuals care only about the existence of opposing views, not their spread. If $\text{WTP}(100) \gg \text{WTP}(1)$, then $\beta > 0$ and concern about spread is substantial.

Hypothesis 5 (Spread): Individuals exhibit a higher willingness to pay to block views when those views would spread, by default, to a higher number of people.

Proposition 5 (Salience)**.** *For ideologically salient issues θ^S and ideologically neutral issues θ^N:*

$$WTP_j(v_i, n; \theta^S) > 0 = WTP_j(v_i, n; \theta^N) \tag{12}$$

Hypothesis 6 (Salience): Individuals exhibit a higher willingness to pay to prevent the dissemination of views on ideologically salient issues than on ideologically neutral issues, where intolerance is predicted to be nonexistent even when views differ.[13]

4.3 Additional Behavioral Predictions

We conclude with preregistered hypotheses that are not directly implied by the model.

[13] Our pre-registration distinguished four related hypotheses: 1) A greater deviation of a view from the personal norm leads to greater intolerance in the ideologically salient vignettes. 2) People do not exhibit intolerance when a view differs from their personal norms in the neutral vignette with a descriptive norm. 3) People do not exhibit intolerance when a view differs from their personal norms in the neutral vignette without a descriptive norm. 4) Intolerance is higher for ideologically salient than for neutral vignettes. For expositional simplicity, we combine these into a single hypothesis.

Our framework is agnostic about the underlying motive of intolerance. We distinguish between two broad motives: First, *moral repugnance*: individuals may derive disutility from views they perceive as morally wrong, offensive, intolerant, or violating fundamental values. Second, *instrumental opposition*: individuals seek to prevent the dissemination of views they believe may harm themselves, their family, their community, or society at large. Both motives may contribute to disutility from exposure and from dissemination. Moral repugnance may generate a desire to block a view even when it reaches only one person, whereas strategic concerns are likely to become more pronounced as dissemination increases. Our framework can also accommodate other motives, such as perceiving a view to be factually incorrect.

Hypothesis 7 (Sources of intolerance): Moral repugnance and strategic concerns about the consequences of a view are the primary motives underlying individuals' willingness to pay to prevent its dissemination.

Our framework can also accommodate differences in intolerance between liberals and conservatives, but does not predict their direction. Let $\bar{\alpha}^L$ and $\bar{\alpha}^C$ denote the average intolerance parameters for liberals and conservatives, respectively. Then, $\bar{\alpha}^L \neq \bar{\alpha}^C$.

Hypothesis 8 (Asymmetric Intolerance): Average intolerance differs across ideological groups.

The model does not make a direct prediction about cooperation. However, if intolerance signals hostility toward individuals holding opposing views, then participants should be less willing to cooperate with intolerant partners. Specifically, if individual j learns that a partner i, who holds opposing views, has previously exhibited intolerance toward views similar to j's, then j's propensity to cooperate with i should decrease.

Hypothesis 9 (Cooperation): Individuals cooperate less with partners holding opposing views when these partners have previously exhibited intolerance toward views they support.

Assuming that α captures a stable preference parameter, intolerance should generalize across contexts involving the suppression of opposing views.

Hypothesis 10 (Predictive Validity): Experimental incentivized measures of intolerance predict willingness to block consequential real-world content (e.g., political posters in our experiments).

Finally, we expect our unincenvized measures of intolerance to capture the same underlying preference for intolerance than our experimental incentivized measures.

Hypothesis 11 (Unincentivized Measures): The survey-based and staircase measures of intolerance positively correlate with the incentivized measures of intolerance.

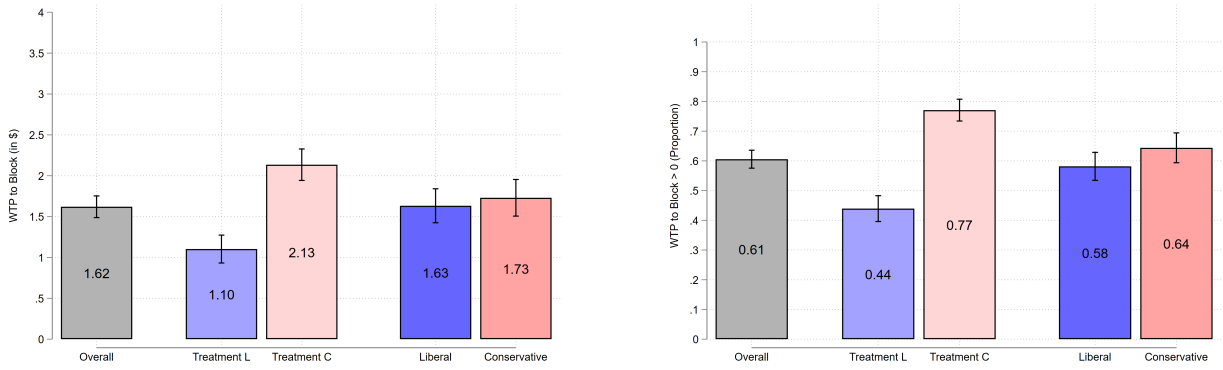
Hypotheses 1-4, 6, 8, and 10-11 can be tested with both experiments, while hypotheses 5, 7, and 9 can only be tested with Experiment 2.

5 Results - Experiment 1

The total sample comprises 1009 participants. Our analyses focus on respondents who can be unambiguously classified as liberal or conservative according to our operational definition of political ideology.[14] Unless otherwise specified, we use the self-reported left–right as our primary measure of ideology. Participants selecting the midpoint (5) are classified as centrists and excluded from the main analyses. The resulting analysis sample consists of 423 liberals and 351 conservatives.

5.1 Willingness to Pay to Block Views - Descriptive Results

We start by presenting descriptive statistics on participants' willingness to pay (WTP) to prevent the dissemination of views. Figure 3(a) reports the mean WTP, overall, by treatment, and separately for Liberals and Conservatives. On average, participants are willing to forgo $\$\mathbf{1.62}$, or 16% of their \$10 endowment, to prevent the dissemination of a view. Mean WTP is \$1.10 in Treatment Liberal and \$2.13 in Treatment Conservative, a statistically significant difference (t-test, $p < 0.01$). In contrast, we find no significant difference in mean WTP between liberal and conservative participants. Figure 3(b) reports the percentage of participants with a strictly positive WTP. Overall, **60%** of participants are willing to pay a positive amount to block a view. This percentage is 43% in Treatment L and 78% in Treatment C, a statistically significant difference (Fisher's exact test, $p = xx$). In contrast, the percentages do not differ significantly between liberal and conservative participants (Fisher's exact test, $p = xx$). Overall, descriptive statistics support Hypothesis 1.



(a) WTP to block (in $)

(b) Proportion of participants with positive WTP

Figure 3: Willingness to Pay (WTP)

The averages reported above pool ideologically salient and ideologically neutral vignettes. Participants, however, respond very differently to the two types of vignettes. For ideologically salient vignettes, participants are willing to pay \$2.19 on average, compared with only \$0.68 for

[14] The correlations of the different ideal points are shown in Appendix Figure B.2.

ideologically neutral vignettes (t-test, $p < 0.01$). The same pattern appears on the extensive margin: 60.77% of them are willing to pay to block an ideologically salient vignette, whereas only 18.72% do so for ideologically neutral vignettes (Fisher's exact test, $p < 0.01$). This supports Hypothesis 1.

5.2 Congruent vs. Incongruent Treatments

We define participants as being in a congruent treatment when the assigned vignettes are consistent with their ideology (i.e., Conservatives assigned to Treatment C and Liberals assigned to Treatment L). Conversely, participants are in an incongruent treatment when exposed to vignettes that are inconsistent with their ideology. Figure 4 compares the mean WTP to block views in the congruent and incongruent treatments, separately for each vignette. The corresponding linear regression estimates are reported in Table B.4 in the Appendix, with Romano-Wolf corrections for multiple hypothesis testing. For ideologically salient vignettes, participants are willing to pay \$1.42 more to block views in the incongruent treatment than in the congruent treatment (R-Wolf $p < 0.01$). In contrast, treatment congruence has no significant effect for the ideologically neutral vignettes.[15]

The congruence effect partly arises at the extensive margin, as shown in Figure B.4 and the corresponding regression Table B.6 in the Appendix. The proportion of participants with a strictly positive WTP is significantly greater in an incongruent than in a congruent treatment. Overall, this analysis of the congruence effects supports Hypothesis 6.

5.3 Treatment Liberal vs. Treatment Conservative

While the previous section established that participants are willing to pay more to block ideologically incongruent views, we now examine whether this pattern differs between liberals and conservatives. Figure 5 reports the mean WTP to block views in Treatment Liberal and Treatment Conservative, separately for liberals and conservatives. Liberals exhibit a substantially higher WTP when exposed to conservative than to liberal views, yielding a positive and statistically significant treatment difference for all four ideologically salient vignettes. By contrast, for conservatives, treatment differences are either null or positive. Neither group exhibits a significant treatment effect for the two ideologically neutral vignettes (Color and Shopping). Similar patterns appear when considering the extensive margin (see Appendix Figure B.5).

The corresponding regressions reported in Table 1 confirm that exposure to Treatment C significantly increases WTP among liberals, but has no significant effect among conservatives.[16]

[15] Since WTP is censored at 0, we report Tobit specifications in Table B.5 in the Appendix. Results are qualitatively unchanged: incongruent treatment increases WTP by between \$2.5 to \$4.4 relative to congruent treatment.

[16] The results for liberals are robust to controlling for Lasso-selected covariates (Table B.7), Tobit estimation (Table B.8), and extensive-margin specifications (Table B.9). For conservatives, the estimated effects are sometimes negative, as expected, but generally less precise and not robust across specifications.



Figure 4: Mean Willingness to Pay to Prevent the Dissemination of Views, by Treatment Congruence–Experiment 1

Notes: The bars show the mean WTP in $ for congruent and incongruent treatments, separately for liberals and conservatives. The red dots indicate the mean difference (congruent–incongruent), with 95% confidence intervals.



Figure 5: Mean Willingness to Pay to Prevent the Dissemination of Views among Liberals and Conservatives, by Treatment–Experiment 1

Notes: The bars show the mean WTP in $ for Treatment Liberal (TL) and Treatment Conservative (TC) vignettes, separately for liberals and conservatives. The red dots indicate the mean difference (TC–TL), with 95% confidence intervals.

Table 1: Willingness to Pay among Liberals and Conservatives–Experiment 1

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Liberals	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	3.046***	2.289***	1.962***	3.561***	-0.283	-0.118	2.717***	-0.203
	(0.324)	(0.304)	(0.322)	(0.338)	(0.183)	(0.190)	(0.263)	(0.180)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.199]]	[[0.508]]	[[0.003]]	[[0.322]]
Constant	1.931	3.227***	2.227	7.657***	0.231	0.287	3.831***	0.252
	(1.614)	(1.054)	(1.674)	(1.742)	(0.668)	(0.656)	(1.159)	(0.637)
Observations	423	423	423	423	423	423	423	423
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Conservatives	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	-0.440	0.329	-0.634	0.624*	0.168	-0.002	-0.021	0.065
	(0.373)	(0.340)	(0.389)	(0.373)	(0.260)	(0.259)	(0.283)	(0.246)
	[[0.628]]	[[0.711]]	[[0.382]]	[[0.382]]	[[0.807]]	[[0.993]]	[[0.990]]	[[0.967]]
Constant	6.368***	5.886***	4.307*	-0.427	1.386	1.529	3.775**	1.935
	(2.126)	(1.805)	(2.587)	(2.208)	(1.501)	(1.498)	(1.910)	(1.373)
Observations	351	351	351	351	351	351	351	351
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear estimates of the WTP in $ to block views, separately for each vignette, with standard errors in parentheses. "Loaded" aggregates the four ideologically salient vignettes, and "neutral" aggregates the two ideologically neutral vignettes. Panel A is for liberal participants, and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

As a robustness check of treatment effects, we replicate the analysis using two alternative measures of ideology instead of using the self-reported left-right ideological scale: party affiliation (Democrat vs. Republican: Figure B.6a), and participants' expressed views on the vignette issue (Figure B.6b). Both figures closely replicate our main findings. Regardless of how ideology is measured, liberals exhibit positive TC–TL differences for all four ideologically salient vignettes, while conservatives display a more contrasted behavior, and treatment effects remain negligible for the ideologically neutral vignettes. These results indicate that our findings are not driven by the particular measure of ideological affiliation. Whether ideology is defined by left-right placement, party affiliation, or expressed views, the same asymmetry holds: participants are willing to pay more to block ideologically incongruent views than congruent views.

Overall, these findings support Hypothesis 2 but not Hypothesis 3. We return to possible explanations for the asymmetry in the structural analysis (Section 7).

5.4 Personal and Social Norms

Hypothesis 4 predicts that intolerance increases with the distance between a view and an individual's personal norm. However, social norms may also influence blocking decisions. We therefore first examine the elicited personal and social norms before assessing their relationship with willingness to pay to prevent the dissemination of views.

Norms across treatments. Figure 6 reports personal and social norms across all vignettes by treatment congruence. In congruent treatment (top panel), both norms are positive, indicating approval of aligned views. For ideologically salient vignettes, personal norms range from approximately 0.7 to 1.5, and social norms from 0.2 to 0.9. In incongruent treatment (bottom panel), both fall sharply, indicating disapproval. In contrast, norms remain positive for the ideologically neutral vignettes in both conditions, suggesting that the treatment effect operates through ideological content rather than other vignette characteristics. Appendix Table B.10 presents the corresponding regression analysis. For ideologically salient vignettes, incongruent treatment significantly reduces both personal and social norms. On average, personal norms decrease by 1.10 units and social norms by 0.46 units (both R-Wolf $p < 0.01$). The effects are largest for the Abortion vignette and smallest for Social Security. In contrast, incongruent treatment has no significant effect on either norms for ideologically neutral vignettes, confirming that the treatment operates through ideological content.

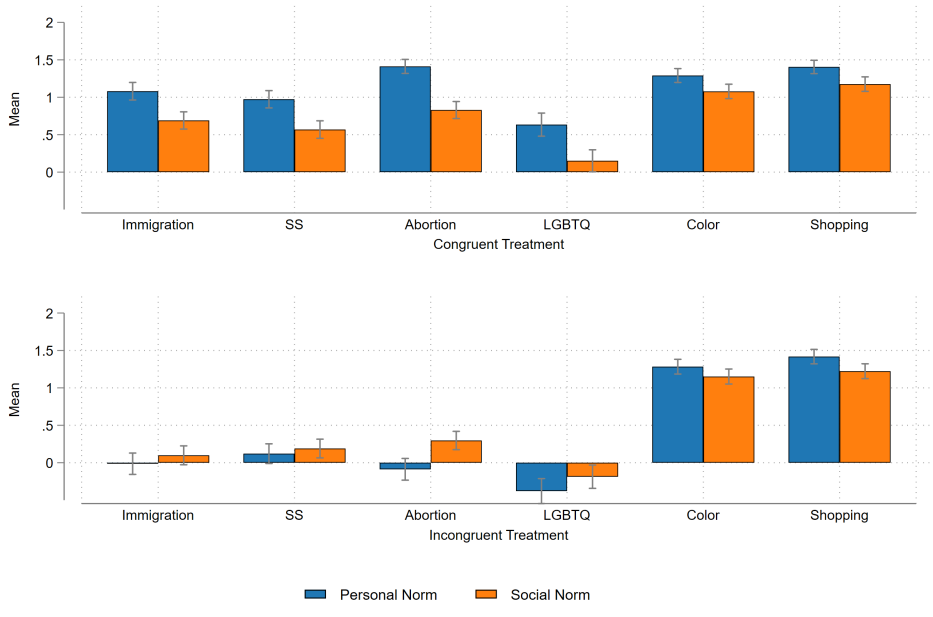


Figure 6: Personal and Social Norms in Congruent and Incongruent Treatments.

Notes. The figure displays average personal norms and social norms, as measured through a five-point appropriateness scale, separately for congruent and incongruent treatments and by vignette. SS for Social Security.

Appendix Figure B.7 disaggregates personal and social norms by political orientation and treatment (C or L). Liberals (top row) assign substantially lower personal and social norms to conservative than to liberal vignettes, whereas conservatives (bottom row) exhibit the mirror pattern. In contrast, ideologically neutral vignettes elicit similar positive norms across groups and treatments. Appendix Table B.11 presents the corresponding regression analysis for personal norms. For liberals, Treatment C significantly reduces personal norms by 2.00 units on average across ideologically salient vignettes (R-Wolf, $p < 0.01$). For conservatives, the average effect is close to zero because treatment effects vary across issues: Treatment C increases personal norms for Abortion but decreases them for LGBTQ. Appendix Table B.12 shows a similar pattern for social norms.

Norms as a mechanism? Having established that personal and social norms vary systematically with treatment assignment, we next examine whether they mediate the treatment effect on WTP. Appendix Tables B.13 with personal norms and B.14 with social norms present linear regressions of WTP showing that both norms are significantly and negatively associated with WTP across most specifications, but the relationship is substantially stronger and more consistent for personal norms.

Following Bašić and Verrina (2024), we estimate and compare three models of WTP for each vignette, separately by participant's ideology and treatment: one including only personal norms, one including only social norms, and one including both. Table 2 reports the corresponding log-likelihoods, likelihood ratio (LR) tests, and information criteria (AIC, BIC).

The results reveal a striking asymmetry. Adding personal norms to a model containing only social norms substantially improves model fit across nearly all specifications. Across panels, the pooled LR statistics range from 6.10 to 25.53, with p-values below 0.02. By contrast, adding social norms to a model that already includes personal norms yields virtually no improvement: the pooled LR statistics range from 0.009 to 0.54, with the lowest p-values at 0.46. The AIC and BIC comparisons lead to the same conclusion: the personal-norm-only model matches or outperforms the combined model in most comparisons, whereas the social-norm-only model is consistently the worst-fitting specification. Unlike Bašić and Verrina (2024), who found that personal and social norms complement each other in predicting prosocial behavior, our results show that personal norms largely subsume the explanatory power of social norms: once individuals' own evaluations of a view are taken into account, perceptions of others' views add little predictive power for intolerance. These results suggest that personal norms are the primary channel through which treatment assignment affects WTP.[17]

[17]These findings are confirmed by formal mediation analyses reported in Appendix Tables B.15 and B.16. For liberals, personal norms mediate a large share of the treatment effect (39-75%), whereas social norms mediate only 4-27%. For conservatives, mediation is less consistent as treatment effects vary across vignettes, but when significant mediation occurs (e.g., Abortion, LGBTQ), personal norms again account for a larger share than social norms.

Table 2: Model Comparison of the Willingness to Pay to Prevent the Dissemination of Views–Experiment 1

Dependent variable: WTP	Immigration			Social Security			Abortion			LGBTQ			Loaded		
	Personal	Social	Personal + Social	Personal	Social	Personal + Social	Personal	Social	Personal + Social	Personal	Social	Personal + Social	Personal	Social	Personal + Social
Liberals — Treatment L															
Adding Personal Norm			0.286			0.005			0.000			0.000			0.000
Adding Social Norm			0.956			0.259			0.488			0.238			0.922
Log-likelihood	-476.217	-476.793	-476.216	-498.747	-502.085	-498.100	-503.663	-528.104	-503.418	-509.792	-516.475	-509.085	-449.196	-461.956	-449.191
Likelihood test PV vs. PV + SV			0.003 (0.956)			1.294 (0.255)			0.489 (0.484)			1.412 (0.235)			0.010 (0.922)
Likelihood test SV vs. PV + SV			1.154 (0.283)			7.970 (0.005)			49.371 (0.000)			14.779 (0.000)			25.530 (0.000)
AIC	956.435	957.585	958.432	1001.494	1008.170	1002.200	1011.326	1060.208	1012.837	1023.583	1036.950	1024.171	902.392	927.912	904.382
BIC	963.276	964.427	968.693	1008.336	1015.012	1012.462	1018.167	1067.049	1023.098	1030.424	1043.791	1034.432	909.233	934.753	914.644
Liberals — Treatment C															
Adding Personal Norm			0.003			0.001			0.004			0.000			0.001
Adding Social Norm			0.357			0.272			0.424			0.657			0.920
Log-likelihood	-526.152	-530.063	-525.720	-516.874	-521.450	-516.260	-526.132	-529.977	-525.807	-533.124	-543.639	-533.023	-489.385	-495.527	-489.380
Likelihood test PV vs. PV + SV			0.864 (0.353)			1.228 (0.268)			0.650 (0.420)			0.201 (0.654)			0.010 (0.919)
Likelihood test SV vs. PV + SV			8.686 (0.003)			10.380 (0.001)			8.340 (0.004)			21.233 (0.000)			12.296 (0.000)
AIC	1056.304	1064.126	1057.440	1037.747	1046.899	1038.519	1056.263	1063.954	1057.614	1070.247	1091.279	1072.046	982.770	995.055	984.759
BIC	1062.871	1070.693	1067.290	1044.314	1053.465	1048.369	1062.830	1070.520	1067.463	1076.814	1097.845	1081.896	989.336	1001.621	994.609
Conservatives — Treatment L															
Adding Personal Norm			0.513			0.135			0.000			0.004			0.015
Adding Social Norm			0.258			0.700			0.245			0.014			0.469
Log-likelihood	-434.440	-434.004	-433.786	-402.234	-403.298	-402.158	-436.517	-443.113	-435.826	-413.396	-414.512	-410.313	-378.279	-381.061	-378.010
Likelihood test PV vs. PV + SV			1.309 (0.253)			0.152 (0.697)			1.383 (0.240)			6.166 (0.013)			0.537 (0.464)
Likelihood test SV vs. PV + SV			0.437 (0.509)			2.281 (0.131)			14.574 (0.000)			8.398 (0.004)			6.101 (0.014)
AIC	872.881	872.008	873.571	808.468	810.597	810.316	877.034	890.226	877.651	830.792	833.023	826.625	760.557	766.122	762.021
BIC	879.117	878.244	882.925	814.704	816.833	819.670	883.270	896.462	887.005	837.028	839.259	835.979	766.793	772.358	771.375
Conservatives — Treatment C															
Adding Personal Norm			0.002			0.002			0.008			0.000			0.000
Adding Social Norm			0.273			0.558			0.393			0.059			0.926
Log-likelihood	-451.665	-455.927	-451.053	-454.567	-459.462	-454.392	-472.718	-475.890	-472.346	-480.044	-484.740	-478.220	-417.952	-430.248	-417.947
Likelihood test PV vs. PV + SV			1.223 (0.269)			0.351 (0.554)			0.745 (0.388)			3.648 (0.056)			0.009 (0.925)
Likelihood test SV vs. PV + SV			9.746 (0.002)			10.141 (0.001)			7.089 (0.008)			13.041 (0.000)			24.601 (0.000)
AIC	907.330	915.853	908.107	913.134	922.924	914.783	949.436	955.780	950.691	964.088	973.481	962.440	839.903	864.496	841.895
BIC	913.759	922.283	917.752	919.564	929.354	924.428	955.865	962.210	960.336	970.518	979.911	972.085	846.333	870.926	851.539

Notes: We estimate three models for each vignette: one including only personal norms, one including only social norms, and one including both. The table reports the corresponding log-likelihoods, likelihood ratio (LR) tests, and information criteria (AIC, BIC). The first LR test (PV vs. PV + SV) tests whether adding the social norm to a model that already contains the personal norm significantly improves fit. The second LR test (SV vs. PV + SV) tests whether adding the personal norm to a model containing only the social norm significantly improves fit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.5 Correlations with alternative measures of intolerance

Our experiment provides an incentivized, context-specific measure of intolerance through WTP to prevent the dissemination of views. Because the literature relies primarily on unincentivized survey-based measures, we examine whether our survey measures provide valid proxies for our incentivized measures. We examine the predictive validity of: (1) the staircase method, which elicits hypothetical WTP to block views, and (2) four survey measures of self-reported tolerance.

Table 3 reports the correlations between these measures and the average WTP to block. The staircase measure is the strongest predictor ($p < 0.01$). Among the survey measures, only "Important to let people spread opposite views" and "Important to be open to opposite views" ($p < 0.01$ and < 0.05, respectively) are significant, although both explain substantially less variation than the staircase method. This provides partial support to Hypothesis 11.

Table 3: Correlation between Staircase Measure and Survey Measures with Willingness to Pay to Prevent the Dissemination of Views–Experiment 1

	(1) Staircase	(2) Important to be open to opposite views	(3) Others view you as very tolerant	(4) Important to let people spread opposite views	(5) Friends with opposite views
Average WTP	0.76***	-0.08**	-0.04	-0.18***	-0.07
	(0.09)	(0.04)	(0.04)	(0.04)	(0.05)
Constant	4.43***	5.79***	6.61***	6.54***	6.22***
	(1.24)	(0.53)	(0.52)	(0.55)	(0.64)
Observations	774	774	774	774	774
R-squared	0.13	0.04	0.06	0.05	0.05

Notes: This table reports OLS regressions, with standard errors in parentheses. The independent variable is the willingness to pay to prevent the dissemination of views. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4 examines whether these measures predict WTP to block political fundraising posters in a field-like setting three days before the 2024 U.S. Presidential election. WTP to block the vignette is the strongest predictor of poster blocking, followed by the staircase measure (both $p < 0.01$). Among survey measures, only "Important to let people spread opposite views" is marginally significant ($p < 0.10$). Similar patterns emerge for the extensive margin. This supports Hypothesis 10.

These findings indicate that both behavioral measures of intolerance (WTP to block vignettes and the staircase method) outperform survey-based measures in predicting both lab and field behavior. The staircase method, while unincentivized, captures a behavioral intention similar to the incentivized measure and therefore serves as a reasonable proxy when incentivized elicitation is not feasible.

Table 4: Correlation between Willingness to Pay to Block Political Posters and Experimental Measure of Intolerance– Experiment 1

(A) WTP to block political posters	(1) WTP to block a vignette	(2) Staircase measure	(3) Important to be open to opposite views	(4) Important to let people spread opposite views
Predictor	0.57***	0.11***	-0.07	-0.12*
	(0.08)	(0.03)	(0.07)	(0.07)
Constant	1.07	1.07	2.08*	2.44**
	(1.05)	(1.09)	(1.17)	(1.18)
Observations	771	771	771	771
R-squared	0.14	0.09	0.08	0.08

(B) I(WTP to block posters > 0)	(5) WTP to block a vignette	(6) Staircase measure	(7) Important to be open to opposite views	(8) Important to let people spread opposite views
Predictor	0.06***	0.01***	-0.01**	-0.02***
	(0.01)	(0.00)	(0.01)	(0.01)
Constant	0.11	0.12	0.26**	0.33***
	(0.09)	(0.10)	(0.11)	(0.11)
Observations	774	774	774	774
R-squared	0.19	0.12	0.10	0.11

Notes: This table reports linear regressions in panel (A) and Probit regressions in panel (B), with standard errors in parentheses. In panel A, the dependent variable is the willingness to pay to prevent the dissemination of views. In panel B, the dependent variable is whether the participant had a positive willingness to pay or not. Predictors are indicated in the headings of columns. All regressions include demographic controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Results - Experiment 2

In this section, we begin by demonstrating that Experiment 2 replicates the main findings of Experiment 1. We then turn to its original contributions: the effects of audience reach on willingness to pay to prevent the dissemination of views, and the motives underlying blocking decisions.

6.1 Replication of Experiment 1

Willingness to pay. As shown in Appendix Figure C.4 and Appendix Table C.5, Experiment 2 replicates the main findings of Experiment 1: participants are willing to pay significantly more to block ideologically incongruent than congruent views. This difference is positive and statistically significant for all vignettes. The same pattern holds on the extensive margin, with a substantially

larger share of participants willing to pay a positive amount to block incongruent views (Appendix Figure C.5 and Table C.6).

Figure 7 and Appendix Table C.7 decompose these effects by political orientation. As in Experiment 1, liberals are willing to pay significantly more to block conservative (TC) than liberal views (TL). Similar patterns hold on the extensive margin (Appendix Figure C.6 and Table C.8) and for personal norms: liberals rate liberal vignettes as more appropriate, while conservatives rate conservative vignettes on abortion and LGBTQ as more appropriate (Appendix Figure C.7 and Table C.10).[18] Table C.7 shows that the responses of conservatives are heterogeneous across issues: they are more willing to block liberal than conservative views on abortion and LGBTQ issues, but a fraction also tries to block conservative views on immigration and social security. We return to this point in Section 6.

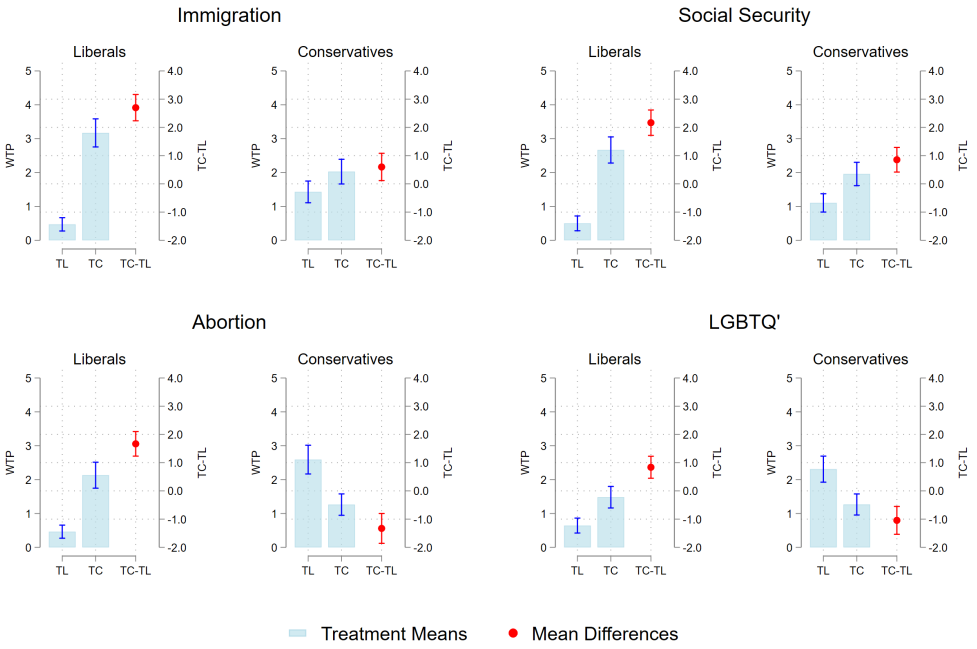


Figure 7: Mean Willingness to Pay to Prevent the Dissemination of Views among Liberals and Conservatives, by Treatment–Experiment 2

Notes: The bars show the mean WTP in $ for Treatment Liberal (TL) and Treatment Conservative (TC) vignettes, separately for liberals and conservatives. The red dots indicate the mean difference (TC–TL), with 95% confidence intervals. This figure is equivalent to Figure 5 in Experiment 1.

External Validity: Blocking Political Posters. The same pattern extends to real political content. Participants could pay to block the dissemination of fundraising posters for either the Republican

[18] Appendix Table C.9 replicates the results of Experiment 1, confirming that ideologically incongruent treatments significantly reduce approval of views expressed in the vignettes. Table C.11 likewise shows that personal norms remain a strong predictor of willingness to pay to block views across most vignettes and for both ideological groups.

or Democratic Party. Figure 8 shows that liberals are willing to pay significantly more to block the Republican poster than the Democratic poster, whereas conservatives exhibit the opposite pattern. Regression Table C.12 in the Appendix confirms that both liberals and conservatives are willing to pay more to block ideologically incongruent fundraising posters.

The regression Table C.13 in the Appendix further supports Hypothesis 10 about the external validity of our measures. Willingness to pay is positively and significantly associated with both the WTP to block vignettes and the staircase measure of intolerance, and negatively associated with the survey measures emphasizing openness to, and the dissemination of, opposing views.

Overall, Experiment 2 confirms the core finding of Experiment 1: individuals exhibit greater intolerance toward views that conflict with their personal norms, this pattern extends to consequential political content, and our experimental measures exhibit strong external validity.

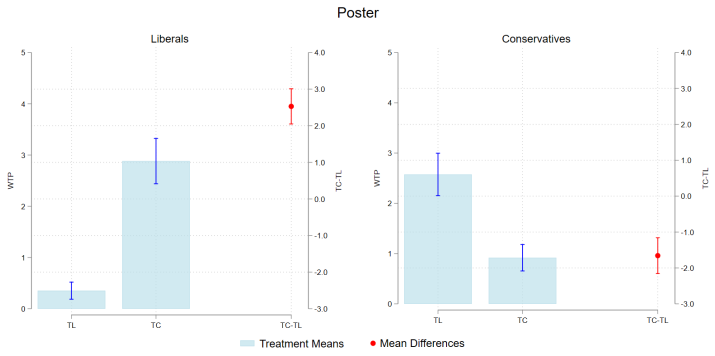


Figure 8: Willingness to Pay to Prevent the Dissemination of Fundraising Political Posters, by Political Orientation–Experiment 2

Notes: The bars show the mean WTP in $ for Treatment Liberal (TL) and Treatment Conservative (TC) posters, separately for liberals and conservatives. The red dots indicate the mean difference (TC–TL), with 95% confidence intervals.

6.2 Effects of Audience Reach

Experiment 2 randomly assigned participants to one of two spread treatments: Spread-1 (T1), in which the view would be disseminated to one person by default, and Spread-100 (T100), in which it would be disseminated to 100 people. This manipulation allows us to test whether the potential audience affects the willingness to pay to prevent the dissemination of a view. As shown in Appendix Figure C.8, mean WTP is equal to or higher under the Spread-100 treatment than under the Spread-1 treatment, for both liberals and conservatives. On average, the Spread-100 treatment increases WTP by approximately $0.3.

Figure 9 reports the estimated spread treatment effect (T100–T1) separately by vignette and political orientation. The estimated coefficients are positive for all vignettes, indicating that greater

reach increases the willingness to pay to prevent the dissemination of views. This supports Hypothesis 5. The effect is generally larger for conservatives than liberals, particularly for the immigration and abortion vignettes. Overall, these results suggest that audience reach amplifies intolerance, supporting Hypothesis 5: individuals are willing to pay more to block views that would reach a larger audience, although the magnitude of this effect remains modest.

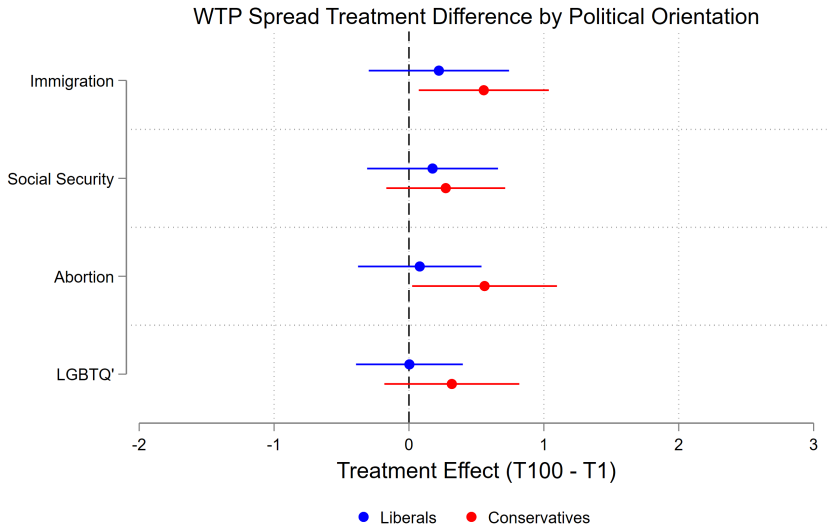


Figure 9: Spread Treatment Effect (T100–T1), by Vignette and Ideological Orientation

Notes: This figure plots the spread treatment effect (T100–T1) for liberals and conservatives separately for each vignette. Positive values indicate higher willingness to pay to prevent the dissemination of views in the T100 treatment than in the T1 treatment.

6.3 Reasons for Intolerance

To better understand the motives underlying intolerance, participants with a positive willingness to pay allocated 100 points across five possible reasons for blocking a view: (1) moral or ethical concerns, (2) false information, (3) harm to oneself, one's family, or society, (4) the view being intolerant, and (5) other reasons. Appendix Figure C.9 reports the average point allocations by ideology and for each vignette separately for the participants exposed to an incongruent treatment–liberals exposed to conservative views (TC) and conservatives exposed to liberal views (TL). Table 5 reports the corresponding linear regression results, using "other reason" as the reference category.

Among liberals exposed to conservative views, the most important motives are moral or ethical concerns, followed by perceived harm, and the view being intolerant, all of which receive significantly more weight than the reference category. In contrast, conservatives exposed to liberal views allocate significantly fewer points to the view being intolerant, while the weight assigned to the other reasons is broadly similar. These results, robust to the inclusion of demographic controls,

support Hypothesis 7: moral repugnance and instrumental concerns are the main reported motives for blocking views.

Table 5: Reasons for Preventing the Dissemination of Views

Dependent variable: Point allocations	No Controls		With Controls	
	Coef.	S.E.	Coef.	S.E.
Moral/Ethical reasons	17.75***	(2.70)	17.75***	(2.71)
False Information	-3.73*	(2.21)	-3.73*	(2.22)
Personal/Family/Societal Harm	10.32***	(2.64)	10.32***	(2.65)
Intolerant Vignette	-2.06	(2.18)	-2.06	(2.19)
Liberal participant	-2.84	(2.26)	-2.84	(2.27)
Moral/Ethical × Liberal	-4.73	(3.38)	-4.73	(3.40)
False Information × Liberal	5.26*	(2.92)	5.26*	(2.93)
Personal/Family/Societal Harm × Liberal	0.83	(3.36)	0.83	(3.38)
Intolerant Vignette × Liberal	12.85***	(3.04)	12.85***	(3.06)
Constant (Other, Conservative)	15.54***	(1.74)	15.54***	(1.75)
Observations	1815		1815	
R^2	0.151		0.151	

Notes: This table reports linear regressions of the number of points allocated to a reason, with standard errors clustered at the individual level in parentheses. The omitted category is *Other reason*. The sample is restricted to participants assigned to incongruent treatments with a positive willingness to pay to prevent the dissemination of a view. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7 Structural Estimation

The reduced-form analysis documents how willingness to pay to prevent the dissemination of opposing views varies across treatments and individuals. We now provide a structural decomposition of willingness to pay by estimating the structural parameters of intolerance at the population and the individual levels.

7.1 Estimation and Identification

Estimation. To quantify the structural parameters governing intolerance, we estimate the model introduced in Section 4 using non-linear least squares:

$$\text{WTP}_{jk} = \alpha \cdot d_{jk}^{\gamma} \cdot [1 + \beta(n_j - 1)] \tag{13}$$

where d_{jk} denotes the ideological distance between participant j and the view expressed in the vignette k, and $n \in \{1, 100\}$ is the number of people the view will reach. The parameters (α, γ, β) correspond to the *intolerance parameter*, the *curvature* of disutility with ideological distance, and

the *spread sensitivity*, respectively. Using a structural approach allows us to recover (α, γ, β) directly from the data, which was not possible with the reduced-form regressions.

Ideological distance is constructed from the personal norm rating, pv_{jk}:

$$d_{jk} = \frac{2 - \text{pv}_{jk}}{4} \in [0, 1],$$

where $pv_{jk} \in \{-2, -1, 0, +1, +2\}$. Full approval corresponds to $d = 0$ and full disapproval to $d = 1$.

Equation (13) is non-linear in parameters because of the term d^{γ}. Therefore, we use non-linear least squares (NLS) to estimate the (α, γ, β) that minimize:

$$\sum_{j,k} \left(\text{WTP}_{jk} - \alpha \cdot d_{jk}^{\gamma} \cdot [1 + \beta(n_j - 1)] \right)^2$$

Parameters are estimated on the data from Experiment 2, which yield $N = 4{,}028$ person–vignette observations (1,007 participants × 4 vignettes), with participant-level clustered standard errors.

Identification. Identification follows directly from the experimental design.

α (level of intolerance) is identified by the overall scale of WTP. At $d = 1$ and $n = 1$, the model predicts WTP = α. α is the dollar amount a person would pay to block a maximally distant view that reaches only one person.

γ (curvature) is identified by how WTP varies across different levels of d within the estimation sample. Since d takes five values (0, 0.25, 0.50, 0.75, 1.00), the shape of the WTP–distance relationship pins down γ. If WTP increases linearly in d, then $\gamma \approx 1$; if it increases at an increasing (decreasing) rate, then $\gamma > 1$ ($\gamma < 1$), implying a convex (concave) relationship between WTP and ideological distance.

β (spread sensitivity) is identified by the randomized between-subject Spread-1 ($n = 1$) vs. Spread-100 treatment ($n = 100$), holding d constant. This is the additional amount people pay because the view will reach more people. Specifically:[19]

$$\text{WTP}(n{=}100) - \text{WTP}(n{=}1) = 99\beta \cdot \alpha \cdot d^{\gamma}$$

7.2 Population-Level Estimates

Model (1) in Table 6 reports the estimates from the pooled model. Model (2) allows parameters to differ between liberals and conservatives; we comment on its results in the next Section.

[19] The Spread treatments are essential because they separately identify β from α; without variation in reach, only the product $(\alpha \cdot [1 + \beta \cdot 99])$ would be identified. This is why we use Experiment 2 data for the structural estimation.

The estimated intolerance parameter, $\hat{\alpha} = 3.24$ ($p < 0.01$), implies that an individual who maximally disapproves of a view ($d = 1$) is willing to pay \$3.24–about one-third of a \$10 endowment–to prevent its dissemination. This indicates that intolerance is economically meaningful.

The estimated curvature parameter, $\hat{\gamma} = 0.97$ ($p < 0.01$), is statistically indistinguishable from one, implying an approximately linear relationship between willingness to pay and ideological distance. There is no evidence that intolerance accelerates or decelerates as distance increases.

The estimated spread parameter, $\hat{\beta} = 0.0009$, is small and insignificant ($p = 0.33$). The implied spread multiplier at $n = 100$ is 1.092, meaning that moving from the Spread-1 to the Spread-100 treatment raises WTP by just 9% (from 3.240 to 3.537). Approximately 92% of WTP is attributable to the mere presence of an opposing view and only 8.4% to concern about that view reaching more others. Thus, intolerance is primarily about the *presence* of disagreement, not its *dissemination*.

Table 6: Structural Parameter Estimates

	(1)	(2)		
Parameters	Pooled	Liberals	Conservatives	p-value (L=C)
α (Intolerance intensity)	3.240***	3.219***	3.326***	0.749
	(0.235)	(0.289)	(0.288)	
γ (Distance curvature)	0.968***	1.263***	0.839***	0.029
	(0.088)	(0.165)	(0.103)	
β (Spread sensitivity)	0.0009	0.001		–
	(0.0009)	(0.0009)		
Observations	4,028	4,028		
R^2	0.343	0.345		

Notes: Model (1) reports nonlinear least-squares estimates of $\text{WTP}_{jk} = \alpha d_{jk}^{\gamma}[1 + \beta(n_j - 1)]$. Model (2) reports nonlinear least-squares estimates of $\text{WTP}_{jk} = [\alpha_L(1-C_j) + \alpha_C C_j] \cdot d_{jk}^{[\gamma_L(1-C_j)+\gamma_C C_j]} \cdot [1 + \beta(n_j - 1)]$, where $C_j = 1$ if participant j is conservative. Standard errors clustered at the participant level are in parentheses. The last column reports a Wald test of equality between liberal and conservative estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.3 Individual-Level Estimates

To examine individual-level heterogeneity in intolerance intensity (α_j) and distance curvature (γ_j), we estimate these parameters separately for each participant using a grid-search procedure (see Appendix D). Appendix Tables D.1 and D.2 report the distribution of individual estimates of α_j and γ_j, while Figures D.1 and D.2 display the corresponding Kernel densities by ideological orientation. The mean $\hat{\alpha}_j$ is 8.53 overall (SD = 18.21), with a median of 4.58. The mean $\hat{\gamma}_j$ is 1.93 overall, with a median of 1.41, indicating that most individuals exhibit convex or near-linear disutility in distance.

7.4 Predictive Validity

To assess the external validity of the structural model, we examine whether the recovered individual parameters estimated from vignette-based WTP predict willingness to pay to block real political fundaising posters elicited three days before the 2024 U.S. presidential election.

Appendix Table D.3 reports the results. The estimated intolerance intensity, $\hat{\alpha}_j$, is not significantly associated with poster-blocking WTP in any specification ($p > 0.30$). By contrast, the curvature parameter, $\hat{\gamma}_j$, is a robust predictor: each one-unit increase in $\hat{\gamma}_j$ is associated with a \$0.21 lower willingness to pay to block the posters ($p = 0.002$), and the estimate is virtually unchanged after controlling for ideological orientation and demographics characteristics. One possible interpretation is that individuals with more convex disutility functions concentrate their intolerance on views they perceive as extremely distant from their own, whereas the posters used in the experiment represent less extreme stimuli. By contrast, individuals with flatter curvature respond more uniformly to ideological disagreement and are thus more willing to block even moderate posters.

Overall, these results offer out-of-sample validation of the model: the curvature of the disutility function, rather than its overall level, predicts consequential behavior beyond the vignette task.

8 Comparing Liberals and Conservatives

We finally turn to a politically consequential question: do liberals and conservatives differ in intolerance, and if so, why? By using three complementary approaches (reduced-form treatment effects, structural parameter estimates, and survey-based measures), we find a consistent pattern: liberals exhibit greater intolerance than conservatives, but the nature of this asymmetry is nuanced.

8.1 Reduced-Form Evidence

To examine whether the effect of treatments' ideological incongruence differs between liberals and conservatives, we estimate the following specification:

$$\text{WTP}_{ij} = \beta_0 + \beta_1 \text{Incongruent}_i + \beta_2 \text{Liberal}_i + \beta_3 (\text{Incongruent}_i \times \text{Liberal}_i) + X_i' \delta + \varepsilon_{ij}$$

where Incongruent_i indicates that participant i was assigned a treatment inconsistent with their ideology, Liberal_i identifies liberals, and X_i is a vector of demographic controls and state fixed effects. The coefficient of interest is β_3, which measures whether liberals respond more strongly than conservatives to ideological incongruence.

Experiment 1. Appendix Table B.17 shows that the interaction term is positive and statistically significant at the 1% level for immigration, social security, and LGBTQ, as well as for the pooled ideologically salient vignettes. The liberal incongruence premium on the pooled vignettes is roughly 29% of the $10 endowment, compared to a near-zero premium for conservatives. The extensive margin results (Appendix Table B.18) lead to the same conclusion: liberals are more likely than conservatives to pay a positive amount to block ideologically incongruent views. By contrast, the coefficient on the Liberal indicator is generally negative, implying that under a congruent treatment liberals are actually less willing to block than conservatives. The liberal–conservative gap in intolerance therefore emerges primarily in response to incongruent views.

Experiment 2. Appendix Table C.14 replicates this pattern. The interaction between incongruence and liberal ideology remains positive and significant for immigration ($p < 0.05$), social security ($p < 0.01$), and for the pooled ideologically salient vignettes ($p < 0.01$), with similar results on the extensive margin (Appendix Table C.15). Liberals pay more than conservatives to block incongruent views, and they are substantially more likely to pay any positive amount.

The liberal–conservative asymmetry extends to real political fundraising posters. As shown by the significant ($p < 0.01$) and positive interaction term in model (6) in Appendix Table C.14, liberals exhibit a larger increase in willingness to pay to block incongruent posters than conservatives. Table C.12, where regressions are run separately by ideological orientation, indicates that the treatment effect for liberals is 54% larger than for conservatives. Thus, the ideological asymmetry documented for hypothetical vignettes also generalizes to consequential political content.

8.2 Structural Evidence

The reduced-form results establish that liberals display a larger incongruence premium than conservatives, but they do not reveal *why*. Are liberals more intensely intolerant of any disagreement, or do they react more sharply as ideological distance increases? The structural estimation presented in Section 7 addresses this question by decomposing willingness to pay into its underlying components. In particular, model (2) in Table 6 allows both the intolerance intensity (α) and the curvature of disutility with ideological distance (γ) to differ between liberals and conservatives.

The estimated intolerance intensities are remarkably similar: $\hat{\alpha}_L = 3.22$ for liberals and $\hat{\alpha}_C = 3.33$ for conservatives, and the equality of α_L and α_C cannot be rejected ($\chi^2 = 0.10$, $p = 0.75$). Thus, both groups are willing to pay comparable amounts to block a maximally distant view.

The key difference lies in the curvature parameter. Liberals exhibit a significantly greater curvature than conservatives ($\hat{\gamma}_L = 1.26$ vs. $\hat{\gamma}_C = 0.84$; $\chi^2 = 4.74$, $p = 0.03$). For liberals, intolerance is *convex* in ideological distance: willingness to pay increases disproportionately as

views become more distant from their own. In contrast, conservatives display a *concave* relationship, with willingness to pay increasing more gradually with ideological distance. Moving from moderate disagreement ($d = 0.5$) to strong disagreement ($d = 1$) more than doubles liberals' WTP (by a factor of $2^{1.26} = 2.39$) compared to only $2^{0.84} = 1.79$ for conservatives. Liberals therefore react especially strongly to extreme disagreement, whereas conservatives are more uniformly averse to opposing views.

The individual-level estimates tell the same story. As also visible in Appendix Figure D.1, which displays Kernel densities of $\hat{\alpha}$ by ideological orientation, mean individual intolerance does not differ significantly between liberals ($\hat{\alpha}_L = 7.48$) and conservatives ($\hat{\alpha}_C = 9.34$; $t = -1.08$, $p = 0.28$).

Although the difference in mean curvature is not statistically significant ($t = 0.95$, $p = 0.34$), reflecting the noise inherent in individual-level estimates, liberals exhibit a higher average curvature than conservatives ($\hat{\gamma}_L = 2.05$ vs. $\hat{\gamma}_C = 1.83$). Moreover, among the 455 participants from Experiment 2 for whom individual $\hat{\gamma}_j$ can be recovered, 60.1% of liberals, compared with 52.1% of conservatives–an 8 percentage point gap–, have an estimated curvature greater than one, indicating a higher prevalence of convex intolerance among liberals. This difference is also apparent in Appendix Figure D.2, where the kernel density of $\hat{\gamma}$ for liberals is shifted to the right relative to that for conservatives.

The structural decomposition thus resolves the apparent puzzle raised by the reduced-form results. Liberals are not more intolerant in *level* than conservatives—their baseline aversion to opposing views is the same; instead, their intolerance *escalates more sharply* with ideological distance. Consequently, incongruent treatments–which expose individuals to more ideologically distant views–generate larger treatment effects for liberals, even though the two groups exhibit similar baseline intolerance.

8.3 Survey-Based Measures

Finally, Table 7 compares liberals and conservatives on three unincentivized, context-free measures of intolerance that do not reference any specific political issue, pooling both experiments ($N = 1{,}781$).[20] Liberals score significantly higher on the Staircase measure of intolerance ($p < 0.01$), which elicits hypothetical willingness to pay to prevent the dissemination of opposing views. This pattern is also evident in Appendix Figure C.10, where the kernel density of Staircase scores is shifted rightward, indicating higher intolerance, for liberals relative to conservatives. They are also significantly less likely to agree with the statement that it is important to *let people spread* views opposite to their own ($p < 0.01$). By contrast, liberals and conservatives do not differ in their

[20] We only retain the two survey measures that are significantly associated with the willingness to pay to prevent the dissemination of views in the vignette task in Table 3.

agreement with the more abstract statement that it is important to *remain open to opposing views* ($p > 0.10$).

Table 7: Liberal-Conservative Differences in Hypothetical Measures of Intolerance–Experiments 1 and 2

Dependent variable:	(1) Staircase	(2) Important to be open to opposite views	(3) Important to let people spread opposite views
Liberal Participant	0.827***	0.183	-0.411***
	(0.274)	(0.122)	(0.123)
Experiment 2	1.536***	-0.181	-0.073
	(0.248)	(0.111)	(0.110)
Observations	1,781	1,781	1,781
Demographic Controls	Yes	Yes	Yes

Notes: This table reports linear regressions of participants' Staircase scores (1–16, with higher values indicating greater intolerance) in column (1) and their responses to two survey questions (1–5 Likert scale) in columns (2)–(3). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

These survey measures reinforce the behavioral evidence. Although liberals express a similar abstract commitment to openness, they are less supportive of the dissemination of opposing views and report higher hypothetical intolerance. Together with the structural estimates, these findings suggest that the larger reduced-form treatment effects observed among liberals reflect a steeper response to ideological disagreement rather than greater baseline intolerance.

9 Discussion and Conclusion

This paper develops and validates an incentivized measure of intolerance–the willingness to pay to prevent the dissemination of views one disagrees with. Combining two controlled experiments with a structural model, we quantify how intolerance depends on ideological distance, audience reach, and personal norms.

Three main findings emerge. First, individuals are willing to incur real monetary costs to prevent the dissemination of opposing views, and this willingness increases in the ideological distance between the view and one's personal norm. Second, personal norms, rather than social norms, are the primary determinant of blocking behavior, echoing and extending the recent finding by Bašić and Verrina (2024) that personal norms shape economic behavior independently of social influence. Third, although liberals exhibit larger reduced-form treatment effects than conservatives when exposed to incongruent views, the structural analysis shows that this reflects differences in the curvature of the disutility function rather than in its baseline level: liberals' willingness to block rises more sharply with ideological distance.

The recovered structural parameters predict consequential political behavior: willingness to block the views expressed in vignettes correlates with willingness to block consequential political

fundraising posters. They are also consistent with independent survey measures of intolerance, supporting the external validity of the proposed measure. From a methodological point of view, because they correlate with our incentivized measures of willingness to pay to block views, we validate our unincentivized staircase measure of intolerance and two survey questions about intolerance.

These findings have implications for the economics of information and political economy. Intolerance is not merely expressive: individuals are willing to incur monetary costs to suppress opposing views, potentially reducing informational diversity and the quality of democratic deliberation. The structural framework developed here provides a tractable foundation for studying the welfare consequences of intolerance and evaluating policies aiming at reducing ideological polarization and preserving open information environments.

Finally, revealed intolerance also generates economic costs beyond the information environment. In Experiment 2, participants played a Prisoner's Dilemma and learned whether their partner was tolerant or intolerant toward views they support. As shown in Appendix Figure C.11 and Table C.16, cooperation is lower against intolerant partners (71% versus 75%), a gap driven almost entirely by conservatives whose cooperation drops from 75% to 67%. Revealed intolerance thus tends to erode cooperation and the gains from trade it would otherwise sustain.

Future research should examine the external validity of these findings in other institutional settings, investigate how intolerance operates in group decision-making, and evaluate interventions capable of reducing costly ideological suppression.

References

Akerlof, George A and Rachel E Kranton, "Economics and identity," *The Quarterly Journal of Economics*, 2000, *115* (3), 715–753.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow, "The Welfare Effects of Social Media," *American Economic Review*, 2020, *110* (3), 629–676.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder, "The strength of issues: Using multiple measures to gauge preference stability, ideological constraint, and issue voting," *American Political Science Review*, 2008, *102* (2), 215–232.

Banerjee, Ritwik, Lata Gangadharan, and Marie Claire Villeval, "Towards understanding Intolerance: An Experimental Approach.," *AEA Social Science Registry*, 2026, *AEARCTR-0013848*.

Bašić, Zvonimir and Eugenio Verrina, "Personal norms — and not only social norms — shape economic behavior," *Journal of Public Economics*, 2024, *239*, 105255.

Bauer, Kevin, Yan Chen, Florian Hett, and Michael Kosfeld, "Group Identity and Belief Formation: A Decomposition of Political Polarization," SAFE Working Paper 409, Leibniz Institute for Financial Research SAFE December 2023.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9* (3), 226–232.

Belot, Michèle and Guglielmo Briscese, "Breaking Echo Chambers: An Experimental Study on the Willingness to Listen to Opposite Views," *Economic Journal*, 2026. Forthcoming. Advance access: https://doi.org/10.1093/ej/ueag072.

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini, "Identity, beliefs, and political conflict," *The Quarterly Journal of Economics*, 2021, *136* (4), 2371–2411.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro, "Cross-country trends in affective polarization," *Review of Economics and Statistics*, 2024, *106* (2), 557–565.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin, "From Extreme to Mainstream: The Erosion of Social Norms," *American Economic Review*, 2020, *110* (11), 3522–3548.

Clinton, Joshua, Simon Jackman, and Douglas Rivers, "The Statistical Analysis of Roll Call Data," *American Political Science Review*, 2004, *98* (2), 355–370.

Crick, Bernard, *Political Theory and Practice*, reprint ed., London: Allen Lane, 1971. Originally published 1963.

Duclos, Jean-Yves, Joan Esteban, and Debraj Ray, "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 2004, *72* (6), 1737–1772.

Enke, Benjamin, "Moral values and voting," *Journal of Political Economy*, 2020, *128* (10), 3679–3729.

_ , Ricardo Rodríguez-Padilla, and Florian Zimmermann, "Moral Universalism and the Structure of Ideology," *The Review of Economic Studies*, 2023, *90* (4), 1934–1962.

_ , _ , and _ , "Moral Universalism: Global Evidence," *American Economic Review*, 2025, *115* (1), 43–76.

Esteban, Joan and Debraj Ray, "Linking Conflict to Inequality and Polarization," *American Economic Review*, 2011, *101* (4), 1345–1374.

Esteban, Joan-María and Debraj Ray, "On the Measurement of Polarization," *Econometrica*, 1994, *62* (4), 819–851.

Genicot, Garance, "Tolerance and compromise in social networks," *Journal of Political Economy*, 2022, *130* (1), 94–120.

Gentzkow, Matthew, Jesse M. Shapiro, and Matt Taddy, "Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech," *Econometrica*, 2019, *87* (4), 1307–1340.

Glaeser, Edward L, "The political economy of hatred," *The Quarterly Journal of Economics*, 2005, *120* (1), 45–86.

Hill, Seth J and Chris Tausanovitch, "A disconnect in representation? Comparison of trends in congressional and public polarization," *The Journal of Politics*, 2015, *77* (4), 1058–1075.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J Westwood, "The origins and consequences of affective polarization in the United States," *Annual Review of Political Science*, 2019, *22*, 129–146.

Jackson, J. C. and D. Medvedev, "Worldwide divergence of values," *Nature Communications*, 2024, *15*, 2650.

Krupka, Erin L and Roberto A Weber, "Identifying social norms using coordination games: Why does dictator game sharing vary?," *Journal of the European Economic Association*, 2013, *11* (3), 495–524.

Levy, Ro'ee, "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment," *American Economic Review*, 2021, *111* (3), 831–870.

Martin, Gregory J. and Ali Yurukoglu, "Bias in Cable News: Persuasion and Polarization," *American Economic Review*, 2017, *107* (9), 2565–2599.

Nie, Norman H., Jane Junn, and Kenneth Stehlik-Barry, *Education and Democratic Citizenship in America*, Chicago: University of Chicago Press, 1996.

Poole, Keith T and Howard Rosenthal, "A spatial model for legislative roll call analysis," *American journal of political science*, 1985, pp. 357–384.

Sullivan, John L., James Piereson, and George E. Marcus, "An Alternative Conceptualization of Political Tolerance: Illusory Increases, 1950s–1970s," *American Political Science Review*, 1979, *73* (3), 781–794.

Sullivan, John L, James Piereson, and George E Marcus, *Political Tolerance and American Democracy*, Chicago: University of Chicago Press, 1982.

Tabellini, Guido, "The Scope of Cooperation: Values and Incentives," *The Quarterly Journal of Economics*, 2008, *123* (3), 905–950.

A Appendix A: Proofs

A.1 Proof of Proposition 1 (Existence)

Proof. By Assumption 1(i), $h(|v_i - v_j|) > 0$ when $v_i \neq v_j$. By Assumption 1(iii), $\psi(n_i) \geq 0$, so $[1 + \psi(n_i)] \geq 1 > 0$. Since $\alpha_j > 0$, equation (5) implies $\text{WTP}_j > 0$. ■

A.2 Proof of Proposition 2 (Ideological congruence)

Proof. For a liberal j^L with $v_j^L < \bar{v}$, we have $|v^C - v_j^L| > |v^L - v_j^L|$ since conservative views are further from liberal personal norms. By Assumption 1(ii), $h(|v^C - v_j^L|) > h(|v^L - v_j^L|)$. Since $[1 + \psi(n)]$ is positive and identical for both comparisons, $\text{WTP}_{j^L}(v^C, n) > \text{WTP}_{j^L}(v^L, n)$. The argument for conservatives is symmetric. ■

A.3 Proof of Proposition 3 (Ideological distance)

Proof. Immediate from Assumption 1(ii) and Equation (5). ■

A.4 Proof of Proposition 4 (Spread)

Proof. By Assumption 1(iii), $\psi(n)$ is increasing in n. Thus $[1 + \psi(n')] > [1 + \psi(n)]$ when $n' > n$. Since $\alpha_j \cdot h(|v_i - v_j|) > 0$, we have $\text{WTP}_j(v_i, n') > \text{WTP}_j(v_i, n)$. ■

A.5 Proof of Proposition 5 (Salience)

Proof. By Equation (6), $\alpha_j(\theta^N) = 0$, so $\text{WTP}_j(v_i, n; \theta^N) = 0$ regardless of ideological distance. For salient issues, $\alpha_j(\theta^S) > 0$, and by earlier results, $\text{WTP}_j(v_i, n; \theta^S) > 0$ when $v_i \neq v_j$. ■

B Appendix B - Figures and Tables from Experiment 1

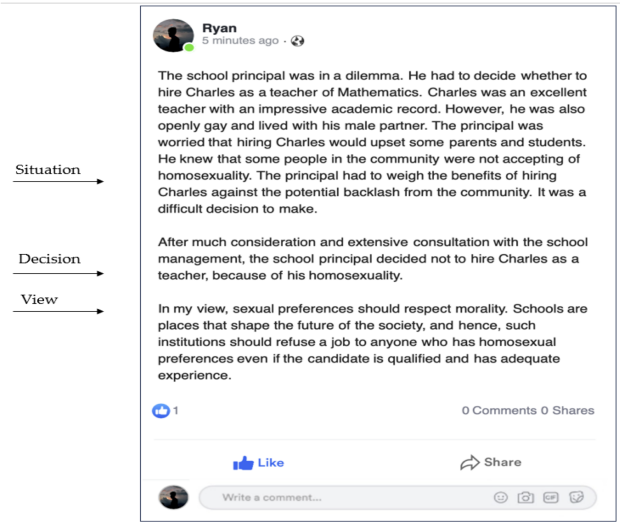


Figure B.1: Vignettes presented in the form of a facebook post

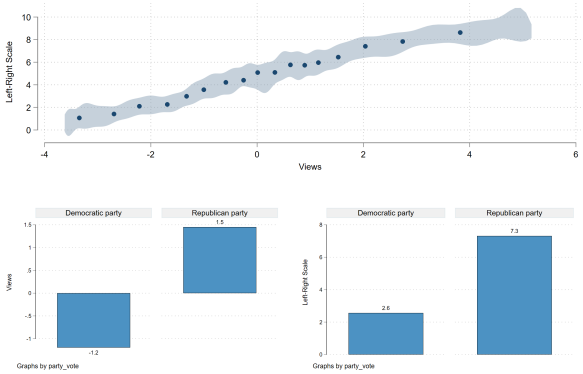


Figure B.2: Relationship between the three measures of ideal points: Left Right, Views, Vote

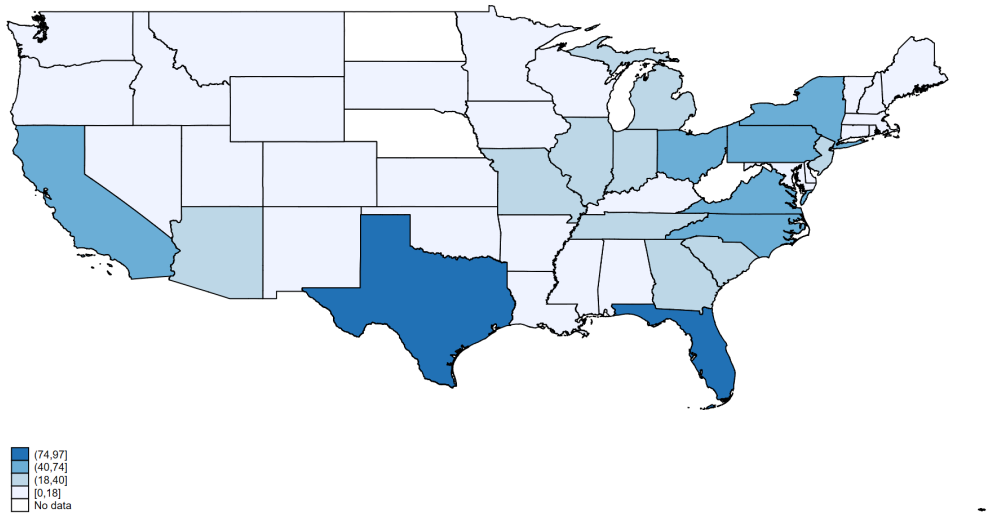


Figure B.3: US Map Weight–Experiment 1



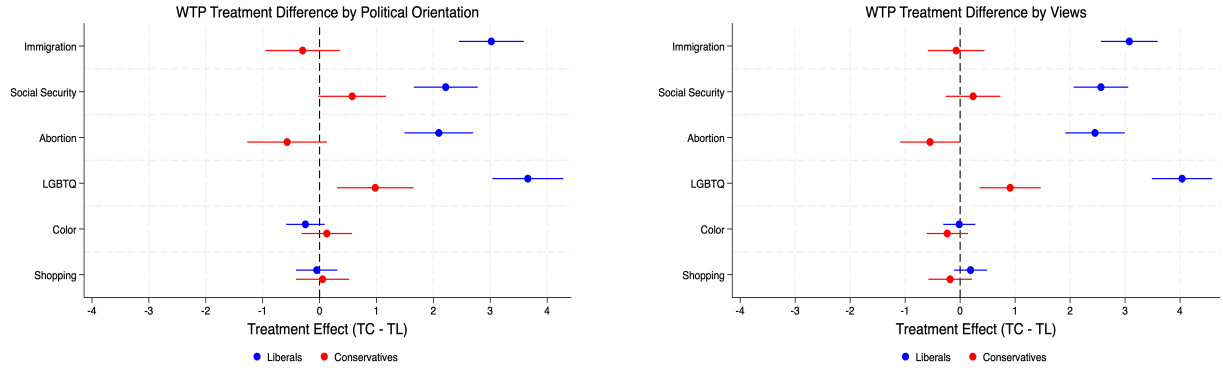
Figure B.4: Proportion with Positive Willingness to Pay to Prevent the Dissemination of Views (Binary) in Congruent and Incongruent Treatments–Experiment 1

Notes: The bars show the proportion of participants with WTP > 0 for congruent and incongruent treatments, separately for liberals and conservatives. The red dots indicate the mean difference (congruent-incongruent), with 95% confidence intervals.



Figure B.5: Proportion of Participants with Positive Willingness to Pay to Prevent the Dissemination of Views (Binary) among Liberals and Conservatives, by Treatment–Experiment 1

Notes: The bars show the proportion of participants with WTP > 0 for Treatment Liberal (TL) and Treatment Conservative (TC) vignettes, separately for liberals and conservatives. The red dots indicate the mean difference (TC-TL), with 95% confidence intervals.



(a) By party affiliation

(b) By expressed views

Figure B.6: Treatment Difference (TC–TL) in the Willingness to Pay under Alternative Definitions of Ideology

Notes: The figure plots the treatment difference in WTP to block (TC–TL) for liberals (blue) and conservatives (red). Panel (a) defines ideology by party affiliation (Democrat vs. Republican); panel (b) classifies ideology based on participants' expressed agreement with views on the vignette issues. Positive values indicate higher willingness to pay to prevent the dissemination of conservative views; negative values indicate higher willingness to pay to prevent the dissemination of liberal views.

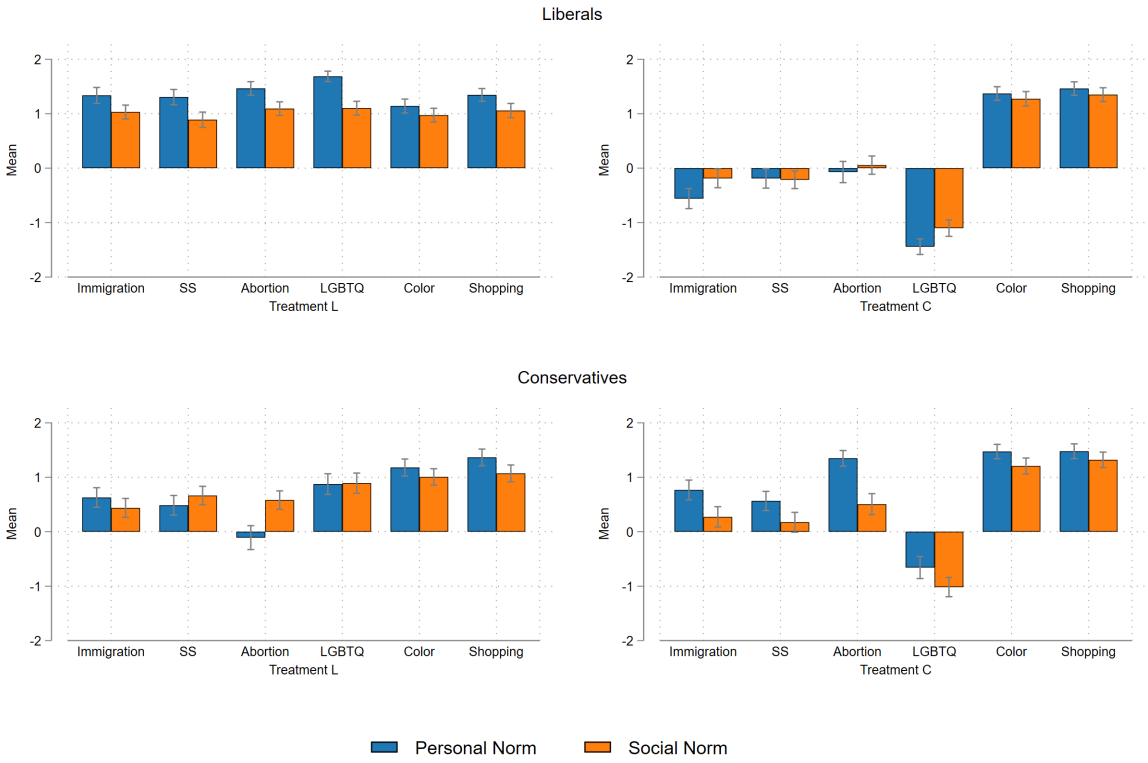


Figure B.7: Personal and Social Norms by Treatments for Liberals and Conservatives

Notes: The figure represents personal and social norms, separately for each vignette and for Treatment C and Treatment L, with 95% confidence intervals. The top panel is for liberal participants, and the bottom panel is for conservative participants.

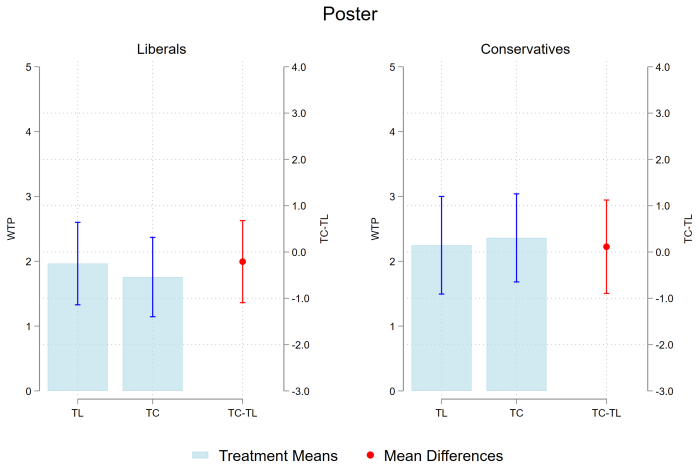


Figure B.8: Willingness to Pay to Prevent the Dissemination of Fundraising Political Posters, by Political Orientation– Experiment 1

Table B.1: Summary Statistics–Validation Experiment for Experiment 1

Variable	Definition	(1) Overall	(2) Treatment C	(3) Treatment L	(4) TC-TL
Age	Age in years	44.63	43.95	45.37	-1.42
		(14.93)	(14.72)	(15.14)	[0.32]
Female	=1 if female, 0 otherwise	0.47	0.49	0.44	0.05
		(0.50)	(0.50)	(0.50)	[0.31]
Religion = Protestant	=1 if religion is Protestant, 0 otherwise	0.17	0.17	0.16	0.01
		(0.37)	(0.38)	(0.37)	[0.81]
Religion = Catholic	=1 if religion is Catholic, 0 otherwise	0.14	0.15	0.14	0.003
		(0.35)	(0.35)	(0.35)	[1.00]
Religion = Other Christians	=1 if religion is Other Christians, 0 otherwise	0.11	0.12	0.10	0.02
		(0.32)	(0.33)	(0.30)	[0.47]
Religion = Atheist	=1 if religion is Atheist, 0 otherwise	0.11	0.13	0.09	0.03
		(0.31)	(0.34)	(0.29)	[0.24]
Religion = Agnostic	=1 if religion is Agnostic, 0 otherwise	0.13	0.14	0.13	0.01
		(0.34)	(0.35)	(0.33)	[0.69]
Religion = Other	=1 if religion is Other, 0 otherwise	0.33	0.29	0.37	-0.08*
		(0.47)	(0.45)	(0.48)	[0.06]
Ethnicity = White	=1 if ethnicity is White, 0 otherwise	0.57	0.65	0.50	0.14***
		(0.49)	(0.48)	(0.50)	[0.002]
Ethnicity = Black	=1 if ethnicity is Black, 0 otherwise	0.13	0.14	0.13	0.01
		(0.34)	(0.34)	(0.33)	[0.79]
Ethnicity = Hispanic	=1 if ethnicity is Hispanic, 0 otherwise	0.07	0.06	0.08	-0.02
		(0.26)	(0.24)	(0.27)	[0.48]
Ethnicity = Other	=1 if ethnicity is Other, 0 otherwise	0.22	0.15	0.29	-0.13***
		(0.42)	(0.36)	(0.45)	[0.001]
Education	Takes value 1 -12, where 1 = Nursery school, and 12 = Doctorate Degree. For full list of categories see the instructions in Appendix	7.73	7.57	7.90	-0.33
		(2.23)	(2.23)	(2.22)	[0.12]
Pre-tax monthly income	Takes value 1-11, where 0-$1,000 takes value 1, and More than $10,000 takes value 11. For the full list of categories see the instruction in Appendix	5.86	5.79	5.93	-0.14
		(3.30)	(3.29)	(3.31)	[0.67]
Ideal Point	Index of ideology calculated from 7-point responses to 10 questions on abortion, transgender rights, healthcare, immigration, weldare, military spending, affirmative action, guns, taxes, equal pay. For full list of categories see the instructions in Appendix	-0.00	0.06	-0.07	0.13
		(1.00)	(1.04)	(0.95)	[0.17]
Left Right	Take value 0-10, where Left i.e. more liberal is 0, and Right i.e. more conservative is 10	4.26	4.49	4.02	0.46
		(3.15)	(3.23)	(3.06)	[0.12]
Vote = Democratic Party	=1 if voting preference is Democratic Party, 0 otherwise	0.53	0.50	0.57	-0.07
		(0.50)	(0.50)	(0.50)	[0.25]
Vote = Republican Party	=1 if voting preference is Republican Party, 0 otherwise	0.29	0.32	0.27	0.05
		(0.46)	(0.47)	(0.44)	[0.37]
Observations		470	226	244	470

Notes: The table reports the variable definitions, the means, and the treatment differences of the control variables for the validation experiment data (for Experiment 1). A t-test is used to test the statistical significance of the difference in the variables between Treatment C and Treatment L. *** p<0.01, ** p<0.05, * p<0.1

Table B.2: Treatment Effect–Validation Experiment for Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Own View	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping
Treatment C	3.969***	4.114***	2.090***	5.472***	0.466	0.093
	(0.311)	(0.341)	(0.262)	(0.307)	(0.506)	(0.224)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.738]]	[[0.738]]
Constant	3.908	-2.045	0.445	3.232	4.297*	3.644*
	(2.765)	(3.032)	(2.331)	(2.731)	(2.273)	(1.991)
Observations	294	294	294	294	172	294
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Others' View	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping
Treatment C	3.654***	3.865***	2.101***	5.260***	0.070	-0.037
	(0.327)	(0.336)	(0.264)	(0.305)	(0.517)	(0.214)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.983]]	[[0.983]]
Constant	2.656	-1.153	2.627	0.155	4.962**	6.287***
	(2.907)	(2.991)	(2.347)	(2.714)	(2.320)	(1.905)
Observations	294	294	294	294	172	294
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Panel A regresses *OwnView*, i.e. one's own rating of the decision and the view expressed in a post, on a scale of 0 (left) to 10 (right), on the conservative treatment dummy and the control variables. Panel B regresses *Others'View*, i.e. one's guess about others' ratings of the decision and the view expressed in a post, on a scale of 0 (left) to 10 (right) and the control variables. The control variables included in the regression are listed in Table B.1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Summary Statistics of Treatment C and Treatment L over Liberals and Conservatives–Experiment 1

	Overall	Liberals/Conservatives	Liberals			Conservatives		
Variable			Treatment C	Treatment L	TC-TL	Treatment C	Treatment L	TC-TL
Age	45.57	46.70	45.40	45.11	0.29	49.06	47.79	1.27
	(15.63)	(15.81)	(16.08)	(15.90)	[0.85]	(15.32)	(15.61)	[0.44]
Female	0.50	0.50	0.46	0.51	-0.05	0.55	0.49	0.06
	(0.50)	(0.50)	(0.50)	(0.50)	[0.33]	(0.50)	(0.50)	[0.29]
Religion = Protestant	0.22	0.21	0.14	0.15	-0.004	0.30	0.29	0.01
	(0.41)	(0.41)	(0.35)	(0.35)	[1.00]	(0.46)	(0.46)	[0.91]
Religion = Catholic	0.19	0.21	0.14	0.19	-0.05	0.26	0.25	0.01
	(0.40)	(0.41)	(0.35)	(0.40)	[0.16]	(0.44)	(0.43)	[0.90]
Religion = Other Christians	0.14	0.14	0.08	0.08	-0.01	0.23	0.18	0.05
	(0.35)	(0.34)	(0.27)	(0.28)	[0.86]	(0.42)	(0.38)	[0.24]
Religion = Atheist	0.09	0.10	0.15	0.17	-0.02	0.01	0.05	-0.04**
	(0.29)	(0.30)	(0.35)	(0.37)	[0.59]	(0.10)	(0.21)	[0.05]
Religion = Agnostic	0.11	0.12	0.18	0.16	0.02	0.05	0.06	-0.01
	(0.32)	(0.32)	(0.39)	(0.37)	[0.61]	(0.22)	(0.24)	[0.81]
Religion = Other	0.23	0.22	0.31	0.24	0.07	0.14	0.17	-0.03
	(0.42)	(0.41)	(0.46)	(0.43)	[0.15]	(0.35)	(0.37)	[0.55]
Ethnicity = White	0.72	0.72	0.68	0.66	0.03	0.79	0.78	0.01
	(0.45)	(0.45)	(0.47)	(0.47)	[0.60]	(0.41)	(0.41)	[0.90]
Ethnicity = Black	0.17	0.18	0.22	0.23	-0.02	0.14	0.13	0.004
	(0.38)	(0.39)	(0.41)	(0.42)	[0.73]	(0.34)	(0.34)	[1.00]
Ethnicity = Hispanic	0.04	0.04	0.04	0.04	-0.004	0.03	0.03	0.003
	(0.20)	(0.19)	(0.20)	(0.21)	[1.00]	(0.18)	(0.17)	[1.00]
Ethnicity = Other	0.06	0.05	0.06	0.06	-0.01	0.04	0.05	-0.02
	(0.24)	(0.22)	(0.23)	(0.24)	[0.84]	(0.19)	(0.23)	[0.61]
Education	7.73	7.87	8.06	7.93	0.13	7.74	7.69	0.04
	(2.12)	(2.07)	(1.90)	(1.97)	[0.50]	(2.23)	(2.18)	[0.85]
Pre-tax monthly income	3.58	3.60	3.49	3.62	-0.14	3.64	3.68	-0.03
	(1.59)	(1.59)	(1.56)	(1.60)	[0.38]	(1.53)	(1.66)	[0.84]
Left Right	4.71	4.62	1.74	1.62	0.12	8.20	8.13	0.07
	(3.08)	(3.51)	(1.37)	(1.33)	[0.38]	(1.41)	(1.41)	[0.62]
Ideal Point	-0.004	-0.01	-0.72	-0.72	-0.004	0.87	0.81	0.06
	(0.99)	(1.07)	(0.73)	(0.73)	[0.95]	(0.76)	(0.70)	[0.48]
Vote = Democratic Party	0.49	0.53	0.91	0.90	0.01	0.07	0.07	0.005
	(0.50)	(0.50)	(0.28)	(0.30)	[0.74]	(0.26)	(0.25)	[1.00]
Vote = Republican Party	0.40	0.43	0.05	0.04	0.001	0.87	0.92	-0.04
	(0.49)	(0.49)	(0.21)	(0.21)	[1.00]	(0.33)	(0.28)	[0.23]
Observations	1,009	774	197	226	1,009	184	167	1,009

Notes: This table reports the mean values of the control variables in the main experiment 1. The first column reports the overall mean and the second column reports the mean for liberals and conservatives (i.e., excluding those observations with *LeftRight*=5). The subsequent columns report the treatment-wise means and the mean differences, separately for liberals and conservatives. A t-test is used to test the statistical significance of the treatment difference. *** p<0.01, ** p<0.05, * p<0.1

Table B.4: Willingness to Pay in Congruent and Incongruent Treatments (Linear specifications)–Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WTP in $	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Incongruent Treatment	1.768***	0.976***	1.277***	1.458***	-0.254*	-0.080	1.458***	-0.152
	(0.233)	(0.217)	(0.242)	(0.240)	(0.149)	(0.151)	(0.193)	(0.139)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.110]]	[[0.575]]	[[0.003]]	[[0.289]]
Constant	4.133**	4.648***	3.814**	4.460	1.426	0.847	4.672***	1.338*
	(1.883)	(1.062)	(1.549)	(3.058)	(0.867)	(0.745)	(1.273)	(0.776)
Observations	774	774	774	774	774	774	774	774
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares estimates of the WTP in $ to block views in the congruent and incongruent treatments, separately for each vignette, with standard errors in parentheses. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.5: Willingness to Pay in Congruent and Incongruent Treatments (Tobit specifications)–Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WTP in $	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Incongruent Treatment	4.409***	2.601***	3.854***	3.083***	-0.440	0.513	2.518***	-0.123
	(0.485)	(0.451)	(0.530)	(0.480)	(0.753)	(0.764)	(0.286)	(0.610)
	[[0.047]]	[[0.126]]	[[0.093]]	[[0.093]]	[[0.847]]	[[0.847]]	[[0.847]]	[[0.897]]
Constant	1.358	3.311	2.563	3.419	2.673	-1.190	4.343**	0.865
	(3.151)	(2.752)	(3.254)	(3.774)	(4.730)	(4.837)	(1.987)	(3.977)
	774	774	774	774	774	774	774	774
Observations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Tobit estimates of the WTP in $ to block views in the congruent and incongruent treatments, separately for each vignette, with standard errors in parentheses. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Positive Willingness to Pay (Binary) in Congruent and Incongruent Treatments–Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$WTP > 0$	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Incongruent Treatment	0.280***	0.181***	0.247***	0.175***	-0.010	-0.008	0.263***	-0.007
	(0.034)	(0.035)	(0.035)	(0.035)	(0.027)	(0.026)	(0.033)	(0.028)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.927]]	[[0.927]]	[[0.003]]	[[0.927]]
Constant	0.724***	0.841***	0.939***	0.881***	0.656*	0.547	0.899***	0.594*
	(0.183)	(0.187)	(0.204)	(0.177)	(0.340)	(0.335)	(0.200)	(0.330)
Observations	774	774	774	774	774	774	774	774
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Probit estimates of the decision to pay a positive amount to block views in the congruent and incongruent treatments, separately for each vignette, with standard errors in parentheses. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Willingness to Pay for Liberals and Conservatives (with Lasso Controls)–Experiment 1

Dependent variable: WTP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Liberals	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	3.047***	2.231***	2.026***	3.588***	-0.265	-0.047	2.723***	-0.156
	(0.286)	(0.282)	(0.300)	(0.317)	(0.162)	(0.176)	(0.240)	(0.158)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.196]]	[[0.777]]	[[0.003]]	[[0.432]]
Constant	2.640	0.504	4.265*	5.542**	0.064	1.180	3.238*	0.622
	(1.945)	(2.382)	(2.557)	(2.256)	(1.790)	(1.794)	(1.776)	(1.750)
Observations	423	423	423	423	423	423	423	423
Lasso Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Conservatives	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	0.341	-0.614**	0.613*	-0.983***	-0.210	-0.163	-0.161	-0.187
	(0.339)	(0.307)	(0.364)	(0.347)	(0.230)	(0.234)	(0.265)	(0.221)
	[[0.618]]	[[0.159]]	[[0.037]]	[[0.037]]	[[0.618]]	[[0.678]]	[[0.678]]	[[0.625]]
Constant	1.003	2.880*	-0.948	3.081*	0.124	0.269	1.504	0.197
	(1.644)	(1.507)	(1.955)	(1.737)	(1.243)	(1.266)	(1.410)	(1.204)
Observations	351	351	351	351	351	351	351	351
Lasso Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear estimates of the WTP in $ to block views, separately for each vignette, with standard errors in parentheses and inclusion of Lasso control variables. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: Willingness to Pay among Liberals and Conservatives (Tobit specifications)–Experiment 1

Dependent variable: WTP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Liberals	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	7.435***	5.743***	6.310***	7.867***	-0.972	0.839	4.757***	0.221
	(0.695)	(0.663)	(0.795)	(0.689)	(1.311)	(1.228)	(0.399)	(0.997)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	-3.382	-0.243	3.335	3.400	-35.117	-32.404	2.050	-28.117
	(5.385)	(5.476)	(6.220)	(5.514)	(13,736.012)	(4,783.198)	(3.616)	(2,664.168)
Observations	423	423	423	423	423	423	423	423
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Conservatives	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	0.755	-1.427**	1.623**	-2.529***	-0.781	-1.129	-0.165	-0.971
	(0.719)	(0.671)	(0.767)	(0.743)	(0.984)	(1.096)	(0.378)	(0.822)
	[[0.601]]	[[0.130]]	[[0.017]]	[[0.017]]	[[0.601]]	[[0.601]]	[[0.694]]	[[0.532]]
Constant	4.307	4.888	1.177	-4.809	4.900	1.930	2.534	3.765
	(6.660)	(6.071)	(7.021)	(6.671)	(7.637)	(8.453)	(3.699)	(6.594)
Observations	351	351	351	351	351	351	351	351
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Tobit estimates of the WTP in $ to block views, separately for each vignette, with standard errors in parentheses. Panel A is for liberal participants, and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Positive Willingness to Pay (Binary) among Liberals and Conservatives–Experiment 1

Dependent variable: WTP> 0	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Liberals	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	0.514***	0.447***	0.401***	0.542***	-0.006	0.017	0.502***	0.016
	(0.046)	(0.046)	(0.046)	(0.045)	(0.034)	(0.034)	(0.042)	(0.036)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.864]]	[[0.807]]	[[0.003]]	[[0.807]]
Constant	0.595***	0.663***	0.624***	0.583***	0.102	0.102	0.635***	0.098
	(0.158)	(0.182)	(0.207)	(0.188)	(0.125)	(0.112)	(0.162)	(0.120)
Observations	423	423	423	423	423	423	423	423
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Conservatives	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	-0.060	0.072	-0.120**	0.178***	-0.007	0.029	-0.004	0.026
	(0.057)	(0.057)	(0.056)	(0.057)	(0.046)	(0.043)	(0.055)	(0.047)
	[[0.658]]	[[0.561]]	[[0.173]]	[[0.013]]	[[0.993]]	[[0.844]]	[[0.993]]	[[0.854]]
Constant	0.829**	0.970***	1.263***	0.751**	1.254***	1.031***	1.280***	1.122***
	(0.327)	(0.341)	(0.347)	(0.349)	(0.278)	(0.287)	(0.328)	(0.277)
Observations	351	351	351	351	351	351	351	351
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Probit estimates of the decision to pay a positive amount to block views, separately for each vignette, with standard errors in parentheses. Panel A is for liberal participants and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: Personal and Social Norms across Congruent and Incongruent Treatments–Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Personal Norm	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Incongruent Treatment	-1.006***	-0.804***	-1.487***	-0.797***	0.009	0.006	-1.111***	0.028
	(0.094)	(0.092)	(0.094)	(0.094)	(0.071)	(0.067)	(0.067)	(0.062)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.993]]	[[0.993]]	[[0.003]]	[[0.870]]
Constant	-1.051	-0.669	1.124**	-0.867	1.438***	1.945***	-0.658	1.563***
	(0.743)	(0.541)	(0.471)	(0.687)	(0.463)	(0.416)	(0.464)	(0.373)
Observations	774	774	774	774	774	774	774	774
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(B) Social Norm	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Incongruent Treatment	-0.514***	-0.358***	-0.472***	-0.189**	0.072	0.056	-0.452***	0.083
	(0.090)	(0.091)	(0.088)	(0.091)	(0.072)	(0.073)	(0.068)	(0.066)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.163]]	[[0.611]]	[[0.668]]	[[0.003]]	[[0.688]]
Constant	-0.781	0.180	0.368	-1.318**	1.233***	1.143**	-0.641	1.216***
	(0.631)	(1.215)	(0.861)	(0.656)	(0.417)	(0.458)	(0.566)	(0.417)
Observations	774	774	774	774	774	774	774	774
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of the effects of incongruent treatment on judgments about the appropriateness of views, separately for each vignette, with standard errors in parentheses. Panel A is for personal norms, and panel B is for social norms. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: Personal Norms across Treatments among Liberals and Conservatives–Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Liberals	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	-1.835***	-1.513***	-1.509***	-3.128***	0.252***	0.113	-1.993***	0.188**
	(0.131)	(0.124)	(0.127)	(0.094)	(0.094)	(0.090)	(0.075)	(0.084)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.013]]	[[0.233]]	[[0.003]]	[[0.030]]
Constant	-1.168	-0.745	0.895	0.700	-0.369	1.202*	-0.065	0.239
	(0.718)	(0.640)	(0.641)	(0.547)	(0.396)	(0.648)	(0.509)	(0.432)
Observations	423	423	423	423	423	423	423	423
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Conservatives	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	0.198	0.041	1.521***	-1.486***	0.254**	0.108	0.079	0.167*
	(0.146)	(0.148)	(0.154)	(0.154)	(0.109)	(0.110)	(0.095)	(0.095)
	[[0.482]]	[[0.767]]	[[0.003]]	[[0.003]]	[[0.086]]	[[0.658]]	[[0.658]]	[[0.299]]
Constant	-1.655**	-0.308	-0.090	1.203	2.216***	1.956***	-0.315	1.695***
	(0.734)	(0.931)	(0.821)	(0.988)	(0.679)	(0.623)	(0.550)	(0.511)
Observations	351	351	351	351	351	351	351	351
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of the effects of Treatment C (conservative) on personal judgments about the appropriateness of views, separately for each vignette, with standard errors in parentheses. "Loaded" aggregates the four ideologically salient vignettes, and "neutral" aggregates the two ideologically neutral vignettes. Panel A is for liberal participants, and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.12: Social Norms across Treatments among Liberals and Conservatives-Experiment 1

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Liberals	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	-1.207***	-1.138***	-0.992***	-2.155***	0.336***	0.280***	-1.376***	0.301***
	(0.119)	(0.119)	(0.112)	(0.109)	(0.095)	(0.099)	(0.076)	(0.090)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.007]]	[[0.003]]	[[0.003]]
Constant	0.145	1.194*	0.003	1.332	0.675*	0.372	0.667	0.389
	(0.775)	(0.692)	(0.606)	(0.888)	(0.388)	(0.488)	(0.636)	(0.387)
Observations	423	423	423	423	423	423	423	423
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Conservatives	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Treatment C	-0.085	-0.413***	-0.035	-1.895***	0.162	0.198*	-0.608***	0.181*
	(0.144)	(0.144)	(0.145)	(0.139)	(0.116)	(0.117)	(0.095)	(0.105)
	[[0.781]]	[[0.030]]	[[0.797]]	[[0.003]]	[[0.449]]	[[0.316]]	[[0.003]]	[[0.312]]
Constant	-1.726**	-0.544	1.338	0.001	1.661**	1.322**	-0.247	1.527**
	(0.870)	(0.879)	(0.921)	(0.966)	(0.725)	(0.652)	(0.664)	(0.629)
Observations	351	351	351	351	351	351	351	351
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of the effects of Treatment C (conservative) on second-order beliefs about majority judgments on the appropriateness of views, separately for each vignette, with standard errors in parentheses. "Loaded" aggregates the four ideologically salient vignettes, and "neutral" aggregates the two ideologically neutral vignettes. Panel A is for liberal participants, and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$.

Table B.13: Effects of Personal Norms on Willingness to Pay: Liberals and Conservatives–Experiment 1

Dependent variable: WTP (A) Liberals — Treatment L	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
Personal Norm	-0.076	-0.447***	-0.975***	-0.545*	-0.766***
	(0.142)	(0.169)	(0.235)	(0.315)	(0.262)
Constant	0.862	-0.166	2.724	1.385	1.257
	(0.940)	(1.269)	(1.961)	(1.650)	(1.313)
Observations	226	226	226	226	226
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.259	0.298	0.425	0.324	0.318

Dependent variable: WTP (B) Liberals — Treatment C	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
Personal Norm	-0.650**	-0.571**	-0.617***	-1.244***	-1.092***
	(0.250)	(0.231)	(0.234)	(0.415)	(0.351)
Constant	8.154***	6.464***	1.448	12.598***	7.026***
	(2.983)	(2.332)	(3.310)	(3.376)	(2.501)
Observations	197	197	197	197	197
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.323	0.353	0.353	0.354	0.319

Dependent variable: WTP (C) Conservatives — Treatment L	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
Personal Norm	-0.057	-0.382*	-0.453*	-0.571**	-0.390
	(0.265)	(0.208)	(0.236)	(0.243)	(0.253)
Constant	-5.733**	-0.632	-1.280	0.185	-2.246
	(2.524)	(1.638)	(2.207)	(1.846)	(1.501)
Observations	167	167	167	167	167
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.417	0.449	0.464	0.460	0.439

Dependent variable: WTP (D) Conservatives — Treatment C	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
Personal Norm	-0.580***	-0.389*	-0.697**	-0.552**	-1.084***
	(0.212)	(0.224)	(0.351)	(0.227)	(0.326)
Constant	9.528***	10.569***	9.473**	4.681	7.734**
	(3.350)	(3.672)	(4.255)	(3.423)	(3.382)
Observations	184	184	184	184	184
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.405	0.451	0.379	0.465	0.448

Notes: This table compares linear regressions of WTP on personal norms, estimated separately by vignette, ideology, and treatment, with standard errors in parentheses. "Loaded" pools the four ideologically salient vignettes. Panel A reports results for liberals in Treatment Liberal (L), panel B for liberals in Treatment Conservative (C), panel C for conservatives in Treatment Liberal, and panel D for conservatives in Treatment Conservative. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.14: Effects of Social Norms on Willingness to Pay: Liberals and Conservatives–Experiment 1

Dependent variable: WTP	(1)	(2)	(3)	(4)	(5)
(A) Liberals — Treatment L	Immigration	Social Security	Abortion	LGBTQ	Loaded
Social Norm	0.117	-0.354*	-0.125	-0.079	0.035
	(0.178)	(0.191)	(0.191)	(0.232)	(0.234)
Constant	0.754	-0.333	2.671	1.021	0.789
	(0.933)	(1.243)	(1.971)	(1.542)	(1.111)
Observations	226	226	226	226	226
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.260	0.285	0.335	0.307	0.279

Dependent variable: WTP	(6)	(7)	(8)	(9)	(10)
(B) Liberals — Treatment C	Immigration	Social Security	Abortion	LGBTQ	Loaded
Social Norm	-0.388	0.143	-0.395	-0.067	-0.269
	(0.306)	(0.270)	(0.278)	(0.366)	(0.424)
Constant	10.842***	7.340***	1.314	15.503***	9.350***
	(2.892)	(2.354)	(3.385)	(3.126)	(2.343)
Observations	197	197	197	197	197
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.300	0.326	0.329	0.294	0.276

Dependent variable: WTP	(11)	(12)	(13)	(14)	(15)
(C) Conservatives — Treatment L	Immigration	Social Security	Abortion	LGBTQ	Loaded
Social Norm	-0.274	0.254	-0.063	-0.582**	-0.231
	(0.302)	(0.216)	(0.282)	(0.253)	(0.262)
Constant	-5.357**	-2.371	-0.889	-0.294	-2.509
	(2.521)	(1.676)	(2.197)	(1.952)	(1.563)
Observations	167	167	167	167	167
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.422	0.437	0.442	0.463	0.431

Dependent variable: WTP	(16)	(17)	(18)	(19)	(20)
(D) Conservatives — Treatment C	Immigration	Social Security	Abortion	LGBTQ	Loaded
Social Norm	-0.478**	-0.374	0.001	-0.096	-0.448*
	(0.203)	(0.232)	(0.215)	(0.258)	(0.264)
Constant	9.547***	11.060***	7.726*	4.811	7.816***
	(3.159)	(3.347)	(4.421)	(3.231)	(2.972)
Observations	184	184	184	184	184
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.399	0.452	0.353	0.437	0.414

Notes: This table compares linear regressions of WTP on social norms, estimated separately by vignette, ideology, and treatment, with standard errors in parentheses. "Loaded" pools the four ideologically salient vignettes. Panel A reports results for liberals in Treatment Liberal (L), panel B for liberals in Treatment Conservative (C), panel C for conservatives in Treatment Liberal, and panel D for conservatives in Treatment Conservative. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.15: Mediation Analysis of the Effects of Personal and Social Norms on Willingness to Pay: Liberals–Experiment 1

Dependent variable: WTP in $	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
(A) Liberals—Personal Norms					
Natural Direct Effect	1.657	1.356	1.039	-0.276	0.685
	(0.00)	(0.00)	(0.00)	(0.70)	(0.18)
Natural Indirect Effect	1.361	0.861	1.057	3.936	2.063
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total Effect	3.019	2.217	2.096	3.660	2.748
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Proportion Mediated	0.451	0.388	0.504	1.075	0.751
Observations	423	423	423	423	423

Dependent variable: WTP in $	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
(B) Liberals—Social Norms					
Natural Direct Effect	2.396	2.134	1.533	2.783	2.030
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Natural Indirect Effect	0.623	0.084	0.563	0.877	0.718
	(0.02)	(0.72)	(0.01)	(0.15)	(0.11)
Total Effect	3.019	2.217	2.096	3.660	2.748
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Proportion Mediated	0.206	0.038	0.269	0.240	0.261
Observations	423	423	423	423	423

Notes: This table reports a formal mediation analysis estimating the effects of personal norms (panel A) and social norms (panel B) on liberal participants' WTP, with standard errors in parentheses. "Loaded" pools the four ideologically salient vignettes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.16: Mediation Analysis of the Effects of Personal and Social Norms on Willingness to Pay: Conservatives–Experiment 1

Dependent variable: WTP in $	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
(A) Conservatives—Personal Norms					
Natural Direct Effect	-0.205	0.625	0.309	0.147	0.219
	(0.54)	(0.04)	(0.54)	(0.70)	(0.39)
Natural Indirect Effect	-0.093	-0.051	-0.880	0.832	-0.048
	(0.30)	(0.53)	(0.02)	(0.00)	(0.66)
Total Effect	-0.298	0.573	-0.570	0.978	0.171
	(0.37)	(0.06)	(0.11)	(0.00)	(0.51)
Proportion Mediated	0.312	-0.089	1.542	0.850	-0.279
Observations	351	351	351	351	351

Dependent variable: WTP in $	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
(B) Conservatives—Social Norms					
Natural Direct Effect	-0.371	0.423	-0.566	0.997	-0.059
	(0.26)	(0.17)	(0.11)	(0.05)	(0.84)
Natural Indirect Effect	0.074	0.150	-0.005	-0.019	0.230
	(0.25)	(0.13)	(0.76)	(0.96)	(0.12)
Total Effect	-0.298	0.573	-0.570	0.978	0.171
	(0.37)	(0.06)	(0.11)	(0.00)	(0.52)
Proportion Mediated	-0.248	0.262	0.008	-0.019	1.345
Observations	351	351	351	351	351

Notes: This table reports a formal mediation analysis estimating the effects of personal norms (panel A) and social norms (panel B) on conservative participants' WTP, with standard errors in parentheses. "Loaded" pools the four ideologically salient vignettes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.17: Regression Analysis of WTP on Incongruent Treatment and Liberals

WTP	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
Incongruent Treatment	0.203	-1.188**	0.758	-0.395	-0.122
	(0.583)	(0.561)	(0.612)	(0.603)	(0.266)
	[[0.894]]	[[0.213]]	[[0.651]]	[[0.894]]	[[0.894]]
Liberals	-1.247**	-1.913***	-0.436	-1.192**	-1.158***
	(0.577)	(0.583)	(0.610)	(0.585)	(0.263)
	[[0.213]]	[[0.010]]	[[0.894]]	[[0.213]]	[[0.003]]
Incongruent Treatment x Liberals	3.033***	4.158***	1.040	3.670***	2.913***
	(0.999)	(1.014)	(1.087)	(1.094)	(0.367)
	[[0.020]]	[[0.003]]	[[0.817]]	[[0.010]]	[[0.003]]
Constant	4.392**	5.177***	4.144**	5.117*	4.687***
	(1.999)	(1.069)	(1.632)	(3.062)	(1.141)
Observations	774	774	774	774	774
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.18: Regression Analysis of Binary WTP on Incongruent Treatment and Liberals

Binary WTP	(1) Immigration	(2) Social Security	(3) Abortion	(4) LGBTQ	(5) Loaded
Incongruent Treatment	-0.066	-0.254***	0.107	-0.223***	-0.032
	(0.085)	(0.088)	(0.089)	(0.086)	(0.051)
	[[0.698]]	[[0.020]]	[[0.538]]	[[0.043]]	[[0.698]]
Liberals	-0.287***	-0.382***	-0.111	-0.306***	-0.241***
	(0.084)	(0.086)	(0.085)	(0.082)	(0.051)
	[[0.003]]	[[0.003]]	[[0.538]]	[[0.003]]	[[0.003]]
Incongruent Treatment x Liberals	0.671***	0.834***	0.283*	0.786***	0.542***
	(0.145)	(0.153)	(0.159)	(0.153)	(0.066)
	[[0.003]]	[[0.003]]	[[0.269]]	[[0.003]]	[[0.003]]
Constant	0.786***	0.946***	1.025***	1.041***	0.911***
	(0.190)	(0.187)	(0.212)	(0.179)	(0.178)
Observations	774	774	774	774	774
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.19: Regression Analysis of WTP in Congruent and Incongruent Treatments with Lasso Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WTP	Immigration	Social Security	Abortion	LGBTQ	Color	Shopping	Loaded	Neutral
Incongruent Treatment	1.795***	0.959***	1.396***	1.543***	-0.211	-0.081	1.423***	-0.146
	(0.221)	(0.212)	(0.233)	(0.247)	(0.136)	(0.142)	(0.184)	(0.131)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.206]]	[[0.575]]	[[0.003]]	[[0.326]]
Constant	0.685	1.139	-0.294	2.420	0.901	1.151	0.987	1.026
	(1.393)	(1.409)	(1.722)	(1.671)	(0.906)	(0.927)	(1.217)	(0.881)
Observations	774	774	774	774	774	774	774	774
Lasso Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

C Appendix C - Figures and Tables from Experiment 2

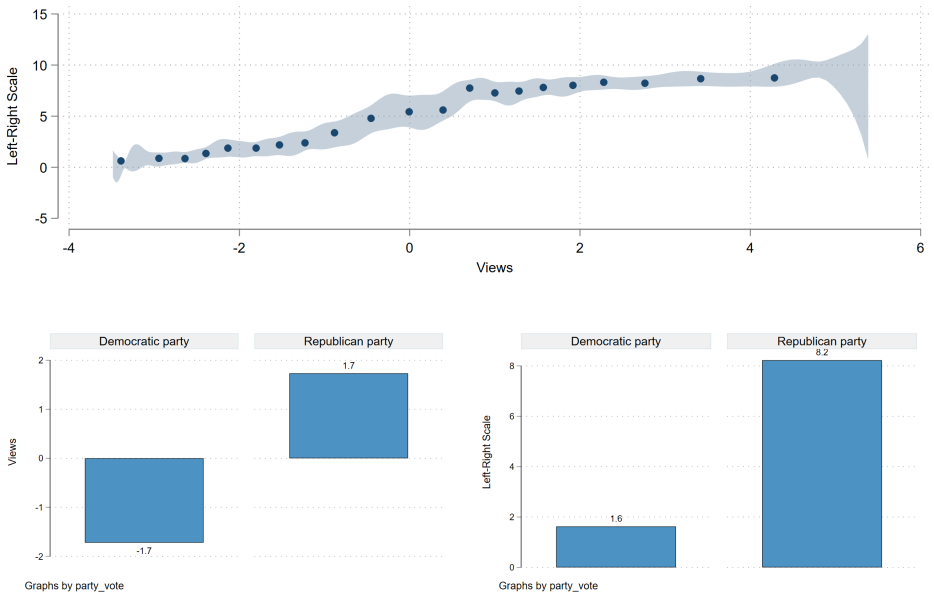


Figure C.1: Relationship between the Three Measures of Ideal Points: Left Right, Views, Vote–Experiment 2

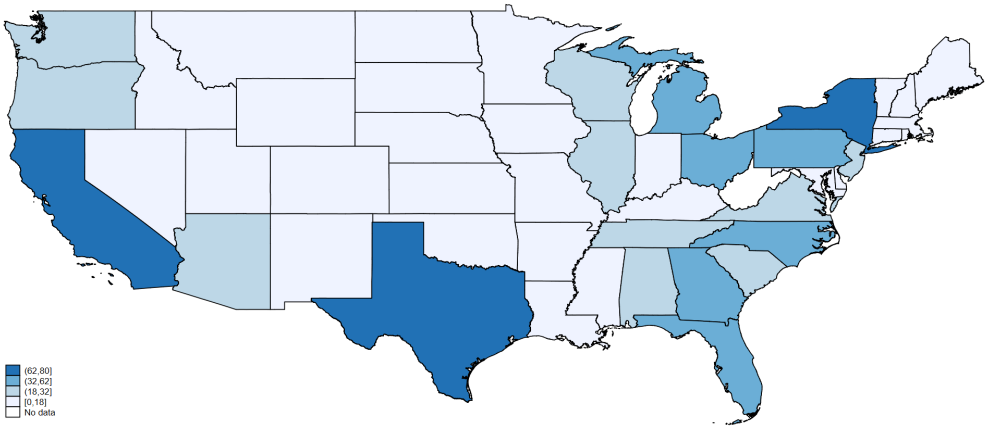


Figure C.2: US Map Weight–Experiment 2

How to Contribute to the Democratic Party



Make a difference to American democracy by supporting the Democratic Party! Here's how you can get involved:

Volunteer for Campaigns

Participate in canvassing, phone or text banking, or use social media to engage voters.

Attend Democratic Party Events

Show your support at town halls, rallies, and fundraisers.

Get Educated & Spread the Word

Stay informed and discuss key issues with your community and spread this poster widely among friends and family.

Support Voting Rights

Volunteer with voter protection programs and help people get to the polls.

Run for Office or Join Local Party Groups

Strengthen local democracy by running for local office or joining committees.

Donate to the Democratic Party Campaign Fund

Support grassroots campaigns with small contributions to DNC, DSCC, DCCC, or candidates.

Donation Links

- Democratic National Committee (DNC): www.democrats.org
- Democratic Congressional Campaign Committee (DCCC): secure.actblue.com/donate/dccc
- Democratic Senatorial Campaign Committee (DSCC): secure.actblue.com/donate/dscc

(a) Democratic Party Poster

How to Contribute to the Republican Party



Make a difference to American democracy by supporting the Republican Party (GOP)! Here's how you can get involved:

Volunteer for Campaigns

Participate in canvassing, phone or text banking, or use social media to engage voters.

Attend Republican Party Events

Show your support at town halls, rallies, and fundraisers.

Get Educated & Spread the Word

Stay informed and discuss key issues with your community and spread this poster widely among friends and family

Support Voting Rights

Volunteer with voter protection programs and help people get to the polls.

Run for Office or Join Local Party Groups

Strengthen local democracy by running for local office or joining committees.

Donate to the Republican Party Campaign Fund

Support grassroots campaigns with small contributions to RNC, NRCC, NRSC, or candidates.

Donation Links

- Republican National Committee (RNC): https://www.gop.gov/
- National Republican Congressional Committee (NRCC): https://www.nrcc.org/
- National Republican Senatorial Committee (NRSC): https://www.nrsc.org/

(b) Republican Party Poster

Figure C.3: Political Fundraising Posters–Experiment 2



Figure C.4: Mean Willingness to Pay to Prevent the Dissemination of Views in Congruent and Incongruent Treatments–Experiment 2

Notes: The bars show the mean WTP in $ for congruent and incongruent treatments. The two upper figures compare the mean WTP across vignettes in Experiment 1 and Experiment 2. The middle and lower figures present the mean WTP per vignette in Experiment 2. The red dots indicate the mean difference (congruent-incongruent), with 95% confidence intervals.

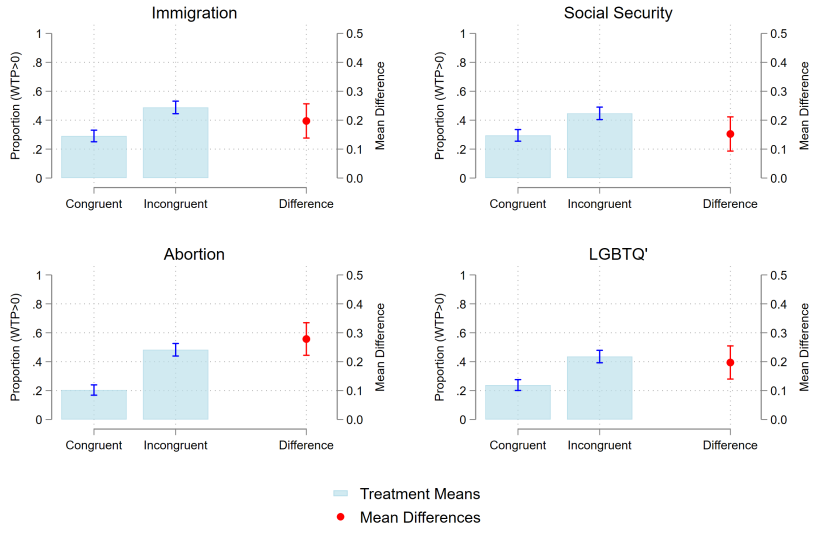


Figure C.5: Proportion of Participants with Positive Willingness to Pay to Prevent the Dissemination of Views (Binary) in Congruent and Incongruent Treatments–Experiment 2

Notes: The bars show the proportion of participants with WTP > 0 for congruent and incongruent treatments. The red dots indicate the mean difference (congruent-incongruent), with 95% confidence intervals.

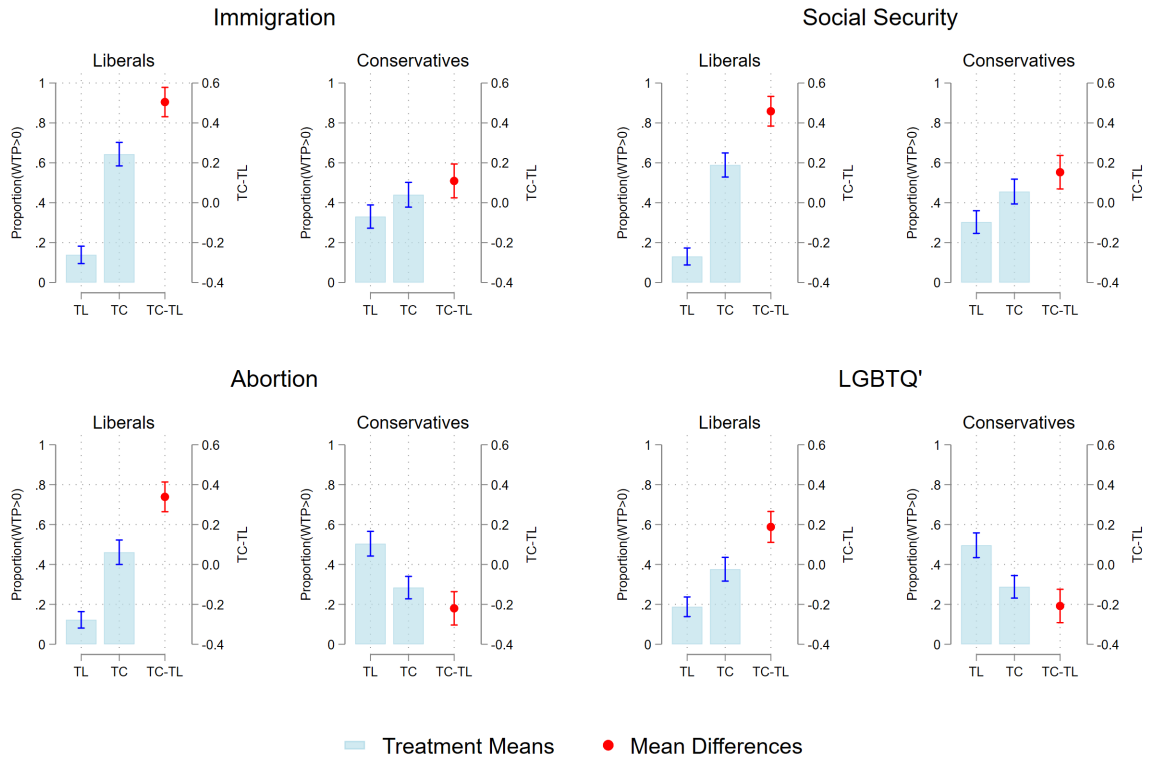
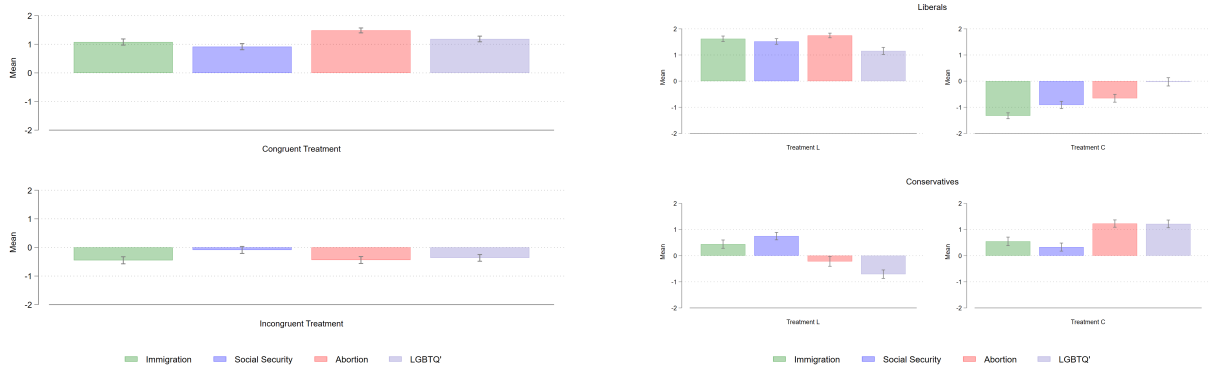


Figure C.6: Proportion of Participants with Positive Willingness to Pay to Prevent the Dissemination of Views (Binary) among Liberals and Conservatives–Experiment 2

Notes: The bars show the proportion of participants with WTP > 0 for Treatment Liberal (TL) and Treatment Conservative (TC) vignettes, separately for liberals and conservatives. The red dots indicate the mean difference (TC–TL), with 95% confidence intervals.



(a) Congruent vs. Incongruent Treatments

(b) By Political Orientation

Figure C.7: Personal norms–Experiment 2

Notes: Panel (a) shows the mean personal norm rating for congruent and incongruent treatments across the four vignettes. Panel (b) shows the mean personal norm rating for Treatment Liberal (TL) and Treatment Conservative (TC) vignettes, separately for liberals and conservatives. Positive values indicate approval of the decision, while negative values indicate disapproval.

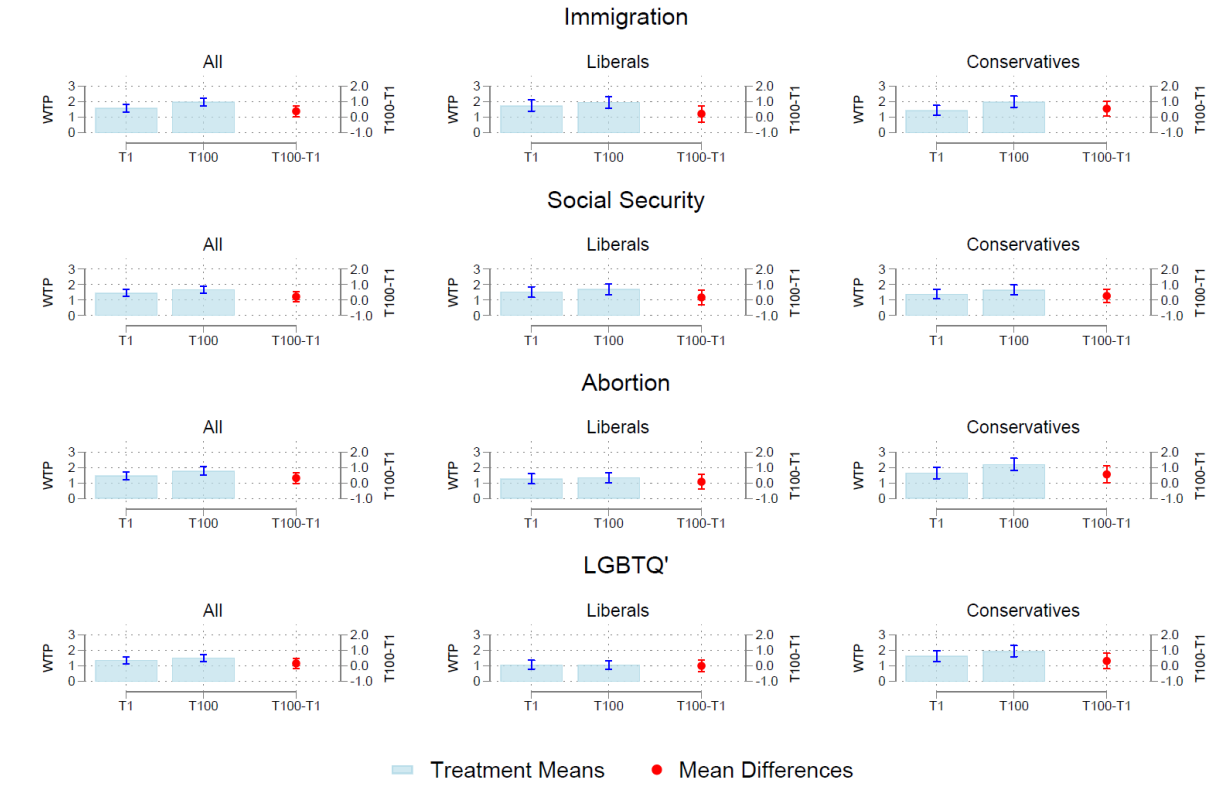


Figure C.8: Willingness to Pay to Prevent the Dissemination of Views, by Spread Treatment–Experiment 2

Notes: The bars show the mean WTP in $ for the Spread-1 (T1) and Spread-100 (T100) treatments, separately for each vignette, for all participants and for liberals and conservatives. The red dots indicate the mean treatment difference (T100–T1), with 95% confidence intervals.

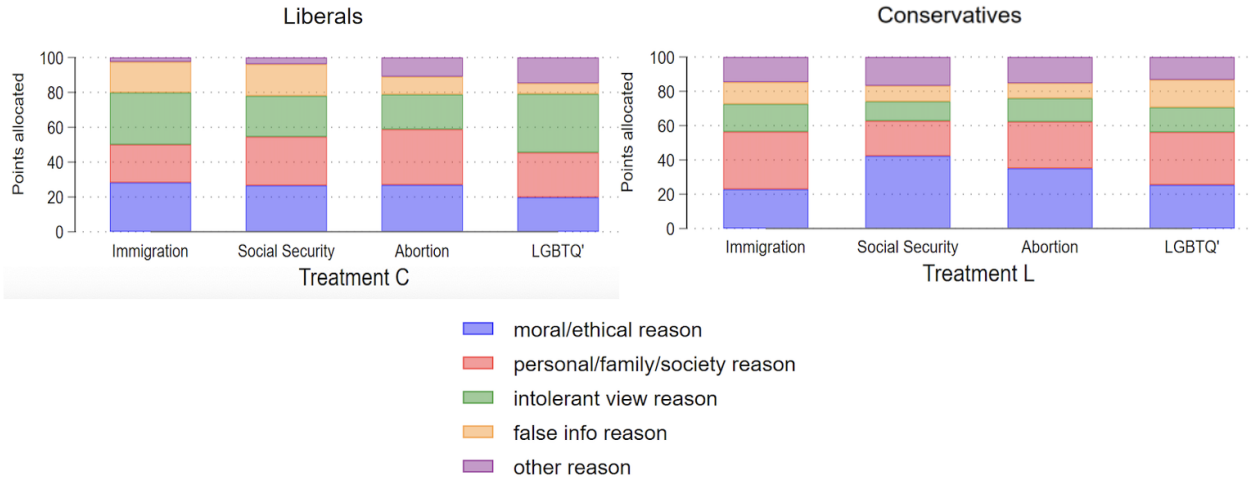


Figure C.9: Reasons for Blocking Views Provided by Participants with a Positive Willingness to Pay–Experiment 2

Notes: The stacked bars show the mean points (out of 100) allocated to each reason for blocking a view, among participants with WTP > 0 in incongruent treatments. Results are shown separately for liberals in Treatment Conservative (TC) and for conservatives in Treatment Liberal (TL) vignettes.

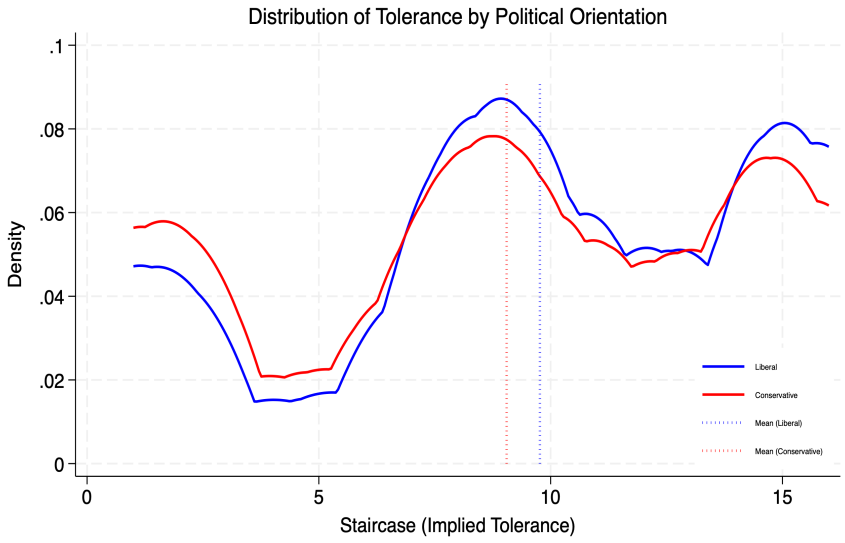


Figure C.10: Kernel Density of Intolerance as Measured by the Staircase Scores, by Ideology–Experiments 1 and 2.

Notes: This figure displays Epanechnikov kernel density of Staircase scores with optimal bandwidth, pooling data from Experiments 1 and 2. The Staircase measure elicits hypothetical willingness to pay to block the spread of views one disagrees with, without reference to any specific political issue.

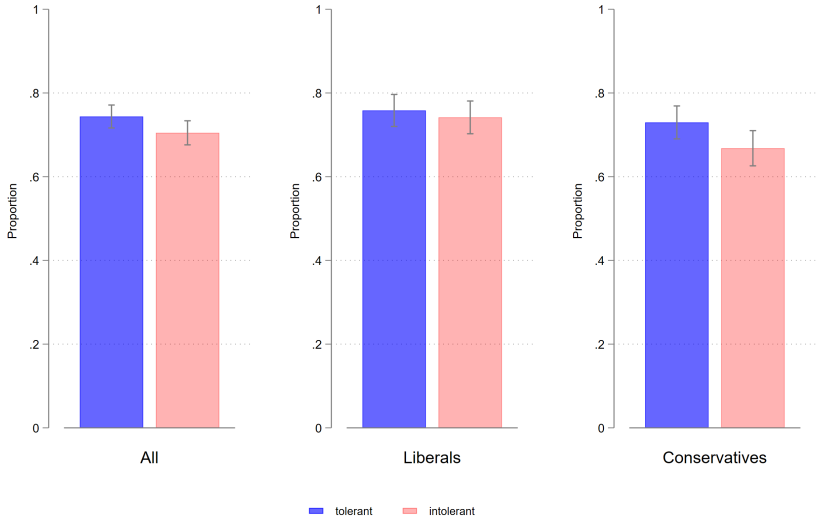


Figure C.11: Cooperation Rates in the Prisoner's Dilemma, by Partner's Tolerance and Ideological Orientation–Experiment 2

Notes: The bars show the proportion of participants who chose to cooperate in the Prisoner's Dilemma, separately by whether their partner was described as tolerant (blue) or intolerant (red).

Table C.1: Summary Statistics–Validation Experiment for Experiment 2

Variable	Definition	(1) Overall	(2) Treatment C	(3) Treatment L	(4) TC-TL
Age	Age in years	46.38	46.21	46.56	-0.35
		(14.78)	(14.08)	(15.49)	[0.84]
Female	=1 if female, 0 otherwise	0.52	0.50	0.54	-0.04
		(0.50)	(0.50)	(0.50)	[0.55]
Religion = Protestant	=1 if religion is Protestant, 0 otherwise	0.23	0.24	0.23	0.005
		(0.42)	(0.43)	(0.42)	[1.00]
Religion = Catholic	=1 if religion is Catholic, 0 otherwise	0.19	0.17	0.21	-0.04
		(0.39)	(0.38)	(0.41)	[0.44]
Religion = Other Christians	=1 if religion is Other Christians, 0 otherwise	0.11	0.12	0.10	0.02
		(0.31)	(0.32)	(0.30)	[0.70]
Religion = Atheist	=1 if religion is Atheist, 0 otherwise	0.10	0.13	0.07	0.06*
		(0.30)	(0.34)	(0.26)	[0.11]
Religion = Agnostic	=1 if religion is Agnostic, 0 otherwise	0.13	0.16	0.10	0.05
		(0.33)	(0.36)	(0.30)	[0.21]
Religion = Other	=1 if religion is Other, 0 otherwise	0.23	0.18	0.28	-0.10*
		(0.42)	(0.39)	(0.45)	[0.06]
Ethnicity = White	=1 if ethnicity is White, 0 otherwise	0.71	0.71	0.70	0.01
		(0.46)	(0.45)	(0.46)	[0.89]
Ethnicity = Black	=1 if ethnicity is Black, 0 otherwise	0.11	0.13	0.09	0.04
		(0.32)	(0.34)	(0.29)	[0.34]
Ethnicity = Hispanic	=1 if ethnicity is Hispanic, 0 otherwise	0.07	0.06	0.07	-0.01
		(0.25)	(0.24)	(0.26)	[0.81]
Ethnicity = Other	=1 if ethnicity is Other, 0 otherwise	0.11	0.10	0.13	-0.03
		(0.32)	(0.30)	(0.34)	[0.45]
Education	Takes value 1 -12, where 1 = Nursery school, and 12 = Doctorate Degree. For full list of categories see the instructions in Appendix	8.00	8.16	7.83	0.33
		(2.11)	(2.05)	(2.17)	[0.20]
Pre-tax monthly income	Takes value 1-11, where 0-$1,000 takes value 1, and More than $10,000 takes value 11. For the full list of categories see the instruction in Appendix	3.68	3.67	3.70	-0.02
		(1.52)	(1.51)	(1.54)	[0.91]
Ideal Point	Index of ideology calculated from 7-point responses to 10 questions on abortion, transgender rights, healthcare, immigration, weldare, military spending, affirmative action, guns, taxes, equal pay. For full list of categories see the instructions in Appendix	-0.00	-0.09	0.09	-0.19
		(1.00)	(0.93)	(1.06)	[0.12]
Left Right	Take value 0-10, where Left i.e. more liberal is 0, and Right i.e. more conservative is 10	4.73	4.50	4.96	-0.46
		(3.16)	(3.14)	(3.17)	[0.23]
Vote = Democratic Party	=1 if voting preference is Democratic Party, =0 otherwise	0.49	0.51	0.48	0.03
		(0.50)	(0.50)	(0.50)	[0.63]
Vote = Republican Party	=1 if voting preference is Republican Party, =0 otherwise	0.39	0.35	0.43	-0.08
		(0.49)	(0.48)	(0.50)	[0.21]
Observations		273	135	138	273

Notes: This table reports the variable definitions, the means, and the treatment differences of the control variables for the validation experiment data used for preparing Experiment 2. A t-test is used to test the statistical significance of the difference in the variables between Treatment C and Treatment L. *** p<0.01, ** p<0.05, * p<0.1.

Table C.2: Treatment Effect–Validation Experiment for Experiment 2

LGBTQ'	(1) Own View	(2) Others' View
Treatment C	4.804***	4.775***
	(0.282)	(0.284)
	[[0.003]]	[[0.003]]
Constant	5.445**	5.407**
	(2.369)	(2.381)
Observations	273	273
Control Variables	Yes	Yes
State FE	Yes	Yes

Notes: Column (1) reports a linear regression of *Own View*, i.e., one's own rating of the decision and the view expressed in the LGBTQ vignette in the validation experiment used for Experiment 2, on a scale of 0 (left) to 10 (right), on the conservative treatment dummy and the control variables. Column (2) reports a linear regression of *Others' View*, i.e., one's guess about others' ratings of the decision and the view expressed in the LGBTQ' vignette, on a scale of 0 (left) to 10 (right), and the control variables. The control variables included in the regressions are listed in Table C.1. Standard errors are reported in parentheses and numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Summary Statistics of Treatment C and Treatment L over Liberals and Conservatives–Experiment 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Overall		Liberals			Conservatives	
Variable		Treatment C	Treatment L	TC-TL	Treatment C	Treatment L	TC-TL
Age	45.20	44.12	45.25	-1.14	47.05	44.41	2.63**
	(14.62)	(14.85)	(15.93)	[0.41]	(13.42)	(14.11)	[0.03]
Female	0.52	0.55	0.54	0.01	0.46	0.52	-0.06
	(0.50)	(0.50)	(0.50)	[0.93]	(0.50)	(0.50)	[0.21]
Religion = Protestant	0.24	0.14	0.11	0.03	0.35	0.35	-0.002
	(0.43)	(0.35)	(0.32)	[0.35]	(0.48)	(0.48)	[1.00]
Religion = Catholic	0.20	0.14	0.11	0.02	0.30	0.27	0.03
	(0.40)	(0.35)	(0.32)	[0.42]	(0.46)	(0.44)	[0.49]
Religion = Other Christians	0.13	0.10	0.08	0.01	0.16	0.17	-0.01
	(0.33)	(0.30)	(0.27)	[0.64]	(0.37)	(0.38)	[0.81]
Religion = Atheist	0.11	0.19	0.20	-0.01	0.02	0.03	-0.01
	(0.31)	(0.39)	(0.40)	[0.82]	(0.15)	(0.18)	[0.60]
Religion = Agnostic	0.13	0.18	0.22	-0.04	0.04	0.06	-0.01
	(0.33)	(0.39)	(0.42)	[0.27]	(0.21)	(0.24)	[0.55]
Religion = Other	0.19	0.25	0.27	-0.02	0.12	0.11	0.01
	(0.39)	(0.43)	(0.44)	[0.68]	(0.33)	(0.31)	[0.78]
Ethnicity = White	0.75	0.67	0.72	-0.04	0.80	0.80	-0.003
	(0.43)	(0.47)	(0.45)	[0.29]	(0.40)	(0.40)	[1.00]
Ethnicity = Black	0.12	0.14	0.14	0.005	0.10	0.10	-0.002
	(0.32)	(0.35)	(0.35)	[0.90]	(0.29)	(0.30)	[1.00]
Ethnicity = Hispanic	0.05	0.07	0.04	0.03	0.05	0.05	-0.003
	(0.22)	(0.25)	(0.19)	[0.11]	(0.21)	(0.22)	[1.00]
Ethnicity = Other	0.08	0.11	0.11	0.01	0.06	0.05	0.01
	(0.27)	(0.32)	(0.31)	[0.89]	(0.23)	(0.21)	[0.69]
Education	8.03	8.17	8.08	0.08	8.004	7.86	0.14
	(2.08)	(2.04)	(1.94)	[0.63]	(2.02)	(2.30)	[0.46]
Pre-tax monthly income	3.66	3.51	3.55	-0.04	3.87	3.70	0.17
	(1.52)	(1.59)	(1.46)	[0.75]	(1.48)	(1.53)	[0.21]
Left Right	4.93	1.70	1.54	0.16	8.22	8.23	-0.02
	(3.57)	(1.44)	(1.34)	[0.19]	(1.34)	(1.34)	[0.89]
Ideal Point	0.004	-0.76	-0.84	0.08	0.85	0.76	0.10*
	(1.00)	(0.59)	(0.52)	[0.12]	(0.65)	(0.65)	[0.09]
Observations	1,007	258	245	1,007	250	254	1,007

Notes: This table reports the mean values of the control variables in the main Experiment 2. The first column reports the overall means. Columns (2)-(4) report the means for liberals and columns (5)-(7) the means for conservatives. Columns (4) and (7) report the mean differences between treatments. A t-test is used to test the statistical significance of the treatment differences. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.4: Summary Statistics of Treatment C, Treatment L, Treatment 100 and Treatment 1–Experiment 2

Variable	(1) Overall	(2) Treatment C	(3) Treatment L	(4) TC-TL	(5) Treatment 100	(6) Treatment 1	(7) T100-T1
Age	45.20	45.56	44.83	0.73	45.85	44.52	1.33
	(14.62)	(14.23)	(15.02)	[0.43]	(14.45)	(14.78)	[0.15]
Female	0.52	0.51	0.53	-0.02	0.52	0.51	0.01
	(0.50)	(0.50)	(0.50)	[0.49]	(0.50)	(0.50)	[0.75]
Religion = Protestant	0.24	0.25	0.24	0.01	0.25	0.23	0.03
	(0.43)	(0.43)	(0.42)	[0.77]	(0.44)	(0.42)	[0.30]
Religion = Catholic	0.20	0.22	0.19	0.02	0.19	0.22	-0.04
	(0.40)	(0.41)	(0.39)	[0.35]	(0.39)	(0.42)	[0.18]
Religion = Other Christians	0.13	0.13	0.13	0.002	0.14	0.12	0.02
	(0.33)	(0.34)	(0.33)	[1.00]	(0.35)	(0.32)	[0.30]
Religion = Atheist	0.11	0.11	0.11	-0.01	0.12	0.09	0.03
	(0.31)	(0.31)	(0.32)	[0.76]	(0.33)	(0.29)	[0.16]
Religion = Agnostic	0.13	0.11	0.14	-0.03	0.11	0.14	-0.03
	(0.33)	(0.32)	(0.35)	[0.22]	(0.31)	(0.35)	[0.13]
Religion = Other	0.19	0.19	0.19	-0.001	0.18	0.19	-0.01
	(0.39)	(0.39)	(0.39)	[1.00]	(0.39)	(0.40)	[0.63]
Ethnicity = White	0.75	0.74	0.76	-0.02	0.78	0.71	0.07***
	(0.43)	(0.44)	(0.43)	[0.38]	(0.41)	(0.45)	[0.01]
Ethnicity = Black	0.12	0.12	0.12	0.002	0.12	0.11	0.01
	(0.32)	(0.32)	(0.32)	[1.00]	(0.33)	(0.32)	[0.77]
Ethnicity = Hispanic	0.05	0.06	0.04	0.01	0.04	0.07	-0.03**
	(0.22)	(0.24)	(0.20)	[0.32]	(0.19)	(0.25)	[0.04]
Ethnicity = Other	0.08	0.08	0.08	0.01	0.06	0.10	-0.05***
	(0.27)	(0.28)	(0.27)	[0.64]	(0.23)	(0.31)	[0.005]
Education	8.03	8.09	7.97	0.12	8.04	8.02	0.02
	(2.08)	(2.03)	(2.13)	[0.37]	(2.07)	(2.10)	[0.90]
Pre-tax monthly income	3.66	3.68	3.63	0.06	3.66	3.65	0.005
	(1.52)	(1.54)	(1.50)	[0.55]	(1.50)	(1.54)	[0.96]
Left Right	4.93	4.91	4.95	-0.04	4.84	5.01	-0.17
	(3.57)	(3.54)	(3.61)	[0.86]	(3.63)	(3.52)	[0.46]
Ideal Point	0.004	0.03	-0.03	0.06	0.002	0.01	-0.003
	(1.01)	(1.02)	(0.99)	[0.34]	(0.98)	(1.03)	[0.96]
Vote = Democratic Party	0.50	0.51	0.49	0.02	0.50	0.50	-0.001
	(0.50)	(0.50)	(0.50)	[0.61]	(0.50)	(0.50)	[1.00]
Vote = Republican Party	0.50	0.49	0.51	-0.02	0.50	0.50	0.001
	(0.50)	(0.50)	(0.50)	[0.61]	(0.50)	(0.50)	[1.00]
Observations	1,007	508	499	1,007	513	494	1,007

Notes: This table reports the mean values of the control variables in the main Experiment 2. The first column reports overall means. Columns (4) and (7) report the mean differences between treatments. A t-test is used to test the statistical significance of the treatment differences. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.5: Willingness to Pay in Congruent and Incongruent Treatments (Linear specifications)–Experiment 2

Dependent variable:	(1)	(2)	(3)	(4)	(5)
$WTPin\$$	Immigration	Social Security	Abortion	LGBTQ'	All
Incongruent Treatment	1.077***	0.631***	1.463***	0.874***	1.015***
	(0.178)	(0.169)	(0.180)	(0.161)	(0.140)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	-0.499	-0.227	0.967	1.696	0.738
	(1.003)	(0.855)	(1.388)	(1.188)	(0.795)
Observations	1,007	1,007	1,007	1,007	1,007
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares estimates of the WTP in $ to block views in the congruent and incongruent treatments, separately for each vignette, with standard errors in parentheses. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. This table is equivalent to Table B.4 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.6: Positive Willingness to Pay (Binary) in Congruent and Incongruent Treatments–Experiment 2

Dependent variable	(1)	(2)	(3)	(4)	(5)
$I(WTP > 0)$	Immigration	Social Security	Abortion	LGBTQ'	All
Incongruent Treatment	0.188***	0.148***	0.265***	0.177***	0.196***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.024)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	0.386**	0.403*	0.285	0.574***	0.446**
	(0.186)	(0.232)	(0.250)	(0.219)	(0.174)
Observations	1,007	1,007	1,007	1,007	1,007
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Probit estimates of the decision to pay a positive amount to block views in the congruent and incongruent treatments, separately for each vignette. Standard errors are reported in parentheses and numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. This table is equivalent to Table B.6 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.7: Willingness to Pay among Liberals and Conservatives–Experiment 2

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)
(A) Liberals	Immigration	Social Security	Abortion	LGBTQ'	All
Treatment C	2.763***	2.249***	1.728***	0.743***	1.864***
	(0.267)	(0.248)	(0.247)	(0.207)	(0.196)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	-0.857	-1.243	-0.056	0.996	-0.088
	(1.284)	(1.199)	(1.355)	(1.088)	(0.762)
Observations	503	503	503	503	503
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)
(B) Conservatives	Immigration	Social Security	Abortion	LGBTQ'	All
Treatment C	0.713***	0.840***	-1.188***	-0.990***	-0.161
	(0.254)	(0.232)	(0.284)	(0.260)	(0.205)
	[[0.010]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.462]]
Constant	-1.259	-1.385	2.428	2.646	0.840
	(1.287)	(1.292)	(2.199)	(1.896)	(1.133)
Observations	504	504	504	504	504
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear estimates of the WTP in $ to block views, separately for each vignette, with standard errors in parentheses. Standard errors are reported in parentheses, and numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. Panel A is for liberal participants, and panel B is for conservative participants. This table is equivalent to Table 1 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: Positive Willingness to Pay (Binary) among Liberals and Conservatives–Experiment 2

Dependent variable: $I(WTP > 0)$	(1)	(2)	(3)	(4)	(5)
(A) Liberals	Immigration	Social Security	Abortion	LGBTQ'	All
Treatment C	0.492***	0.474***	0.340***	0.146***	0.362***
	(0.041)	(0.040)	(0.042)	(0.043)	(0.032)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	0.567***	0.485**	0.256	0.788***	0.557***
	(0.198)	(0.242)	(0.387)	(0.292)	(0.191)
Observations	503	503	503	503	503
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Dependent variable: $I(WTP > 0)$	(1)	(2)	(3)	(4)	(5)
(B) Conservatives	Immigration	Social Security	Abortion	LGBTQ'	All
Treatment C	0.136***	0.151***	-0.198***	-0.199***	-0.028
	(0.046)	(0.047)	(0.046)	(0.046)	(0.035)
	[[0.007]]	[[0.007]]	[[0.003]]	[[0.003]]	[[0.415]]
Constant	0.077	0.016	0.459	0.511*	0.294
	(0.314)	(0.366)	(0.391)	(0.271)	(0.253)
Observations	504	504	504	504	504
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Probit estimates of the decision to pay a positive amount to block views, separately for each vignette. Standard errors are reported in parentheses and numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. Panel A is for liberal participants and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. This table is equivalent to Table B.9 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.9: Personal Norms across Congruent and Incongruent Treatments–Experiment 2

Dependent variable:	(1)	(2)	(3)	(4)	(5)
Personal Norm	Immigration	Social Security	Abortion	LGBTQ	All
Incongruent Treatment	-1.506***	-0.989***	-1.918***	-1.513***	-1.486***
	(0.074)	(0.076)	(0.076)	(0.081)	(0.059)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	1.709***	1.343***	1.653***	1.720***	1.435***
	(0.558)	(0.519)	(0.477)	(0.660)	(0.523)
Observations	1,007	1,007	1,007	1,007	1,007
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of the effects of incongruent treatment on personal norms about the appropriateness of views, separately for each vignette, with standard errors in parentheses. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. This table is equivalent to Table B.10 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.10: Personal Norms across Treatments among Liberals and Conservatives–Experiment 2

Dependent variable: Norm	(1)	(2)	(3)	(4)	(5)
(A) Liberals	Immigration	Social Security	Abortion	LGBTQ'	All
Treatment C	-2.913***	-2.472***	-2.479***	-1.133***	-2.254***
	(0.086)	(0.095)	(0.090)	(0.111)	(0.069)
	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	1.849***	1.328*	1.730**	0.582	1.308**
	(0.462)	(0.684)	(0.749)	(0.961)	(0.525)
Observations	503	503	503	503	503
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Dependent variable: Norm	(1)	(2)	(3)	(4)	(5)
(B) Conservatives	Immigration	Social Security	Abortion	LGBTQ'	All
Treatment C	-0.003	-0.459***	1.366***	1.849***	0.689***
	(0.122)	(0.116)	(0.134)	(0.124)	(0.084)
	[[0.970]]	[[0.003]]	[[0.003]]	[[0.003]]	[[0.003]]
Constant	2.460***	1.530***	0.549	1.528**	1.244**
	(0.829)	(0.573)	(0.615)	(0.776)	(0.503)
Observations	504	504	504	504	504
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of the effects of Treatment C (conservative) on personal judgments about the appropriateness of views, separately for each vignette, with standard errors in parentheses. Panel A is for liberal participants, and panel B is for conservative participants. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. This table is equivalent to Table B.11 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.11: Effects of Personal Norms on Willingness to Pay: Liberals and Conservatives–Experiment 2

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)
(A) Liberals — Treatment L	Immigration	Social Security	Abortion	LGBTQ	All
Personal Norm	-0.076	-0.213	-0.627***	-0.574***	-0.688***
	(0.134)	(0.152)	(0.166)	(0.113)	(0.161)
Constant	1.129	-0.658	0.977	1.314	1.346
	(2.122)	(2.486)	(2.030)	(2.232)	(1.723)
Observations	245	245	245	245	245
R-squared	0.341	0.296	0.381	0.423	0.414
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)
(B) Liberals — Treatment C	Immigration	Social Security	Abortion	LGBTQ	All
Personal Norm	-1.023***	-0.082	-0.480**	-0.830***	-0.708***
	(0.250)	(0.203)	(0.192)	(0.144)	(0.217)
Constant	-2.222	-0.630	1.157	0.612	0.217
	(3.586)	(3.395)	(3.347)	(2.672)	(2.630)
Observations	258	258	258	258	258
R-squared	0.339	0.271	0.277	0.365	0.299
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)
(C) Conservatives — Treatment L	Immigration	Social Security	Abortion	LGBTQ	All
Personal Norm	-0.402***	-0.482***	-0.825***	-0.689***	-0.720***
	(0.152)	(0.130)	(0.162)	(0.169)	(0.191)
Constant	2.388	0.750	3.526	6.333**	3.718*
	(2.599)	(1.941)	(3.302)	(3.006)	(2.227)
Observations	254	254	254	254	254
R-squared	0.325	0.425	0.339	0.335	0.322
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Dependent variable: *WTP*	(1)	(2)	(3)	(4)	(5)
(D) Conservatives — Treatment C	Immigration	Social Security	Abortion	LGBTQ	All
Personal Norm	-0.900***	-0.454***	-1.192***	-0.824***	-1.166***
	(0.156)	(0.170)	(0.141)	(0.139)	(0.167)
Constant	1.817	-2.563	5.666**	1.765	1.958
	(3.536)	(3.690)	(2.725)	(2.893)	(2.424)
Observations	250	250	250	250	250
R-squared	0.351	0.233	0.488	0.411	0.395
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of WTP on personal norms, estimated separately by vignette, ideology, and treatment, with standard errors in parentheses. Panel A reports results for liberals in Treatment Liberal (L), panel B for liberals in Treatment Conservative (C), panel C for conservatives in Treatment Liberal, and panel D for conservatives in Treatment Conservative. This table is equivalent to Table B.13 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.12: Willingness to Pay to Prevent the Dissemination of Political Posters among Liberals and Conservatives–Experiment 2

Dependent variable:	(1)	(2)
WTP in$	Liberals	Conservatives
Incongruent Treatment	2.514***	1.630***
	(0.255)	(0.273)
	[[0.003]]	[[0.003]]
Constant	3.035	-1.077
	(2.349)	(1.229)
Observations	503	504
Control Variables	Yes	Yes
State FE	Yes	Yes

Notes: This table reports linear regressions of WTP in $ to prevent the dissemination of real political fundraising posters, with standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.13: Correlation between Willingness to Pay to Block Political Posters and Experimental Measures of Intolerance–Experiment 2

(A)	(1)	(2)	(3)	(4)
WTP to block posters	WTP to block a vignette	Staircase measure	Important to be open to opposite views	Important to let people spread opposite views
Predictor	0.72***	0.15***	-0.13***	-0.26***
	(0.04)	(0.02)	(0.04)	(0.04)
Constant	1.22	0.78	3.46**	3.96***
	(1.43)	(1.34)	(1.49)	(1.35)
Observations	1,007	1,007	1,007	1,007
R-squared	0.35	0.15	0.10	0.13

(B)	(5)	(6)	(7)	(8)
I(WTP to block posters > 0)	WTP to block a vignette	Staircase measure	Important to be open to opposite views	Important to let people spread opposite views
Predictor	0.10***	0.02***	-0.01*	-0.04***
	(0.01)	(0.00)	(0.01)	(0.01)
Constant	0.36*	0.26	0.62***	0.75***
	(0.20)	(0.21)	(0.23)	(0.20)
Observations	1,007	1,007	1,007	1,007
R-squared	0.28	0.15	0.09	0.12

Notes: This table reports linear regressions of WTP in $ in panel (A) and Probit regressions in panel (B), with standard errors in parentheses. In panel A, the dependent variable is the willingness to pay to prevent the dissemination of political posters. In panel B, the dependent variable is whether the participant had a positive willingness to pay or not. Predictors are indicated in the headings of columns. All regressions include demographic controls. This table is equivalent to Table 4 in Experiment 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.14: Willingness to Pay: Effects of Incongruent Treatment by Ideological Orientation–Experiment 2

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
WTP in $	Immigration	Social Security	Abortion	LGBTQ'	All	Poster
Incongruent Treatment	0.201	-1.116***	1.079**	0.978**	0.191	1.567***
	(0.461)	(0.414)	(0.489)	(0.443)	(0.200)	(0.258)
	[[0.894]]	[[0.063]]	[[0.146]]	[[0.146]]	[[0.771]]	[[0.003]]
Liberal Participant	-0.765	-1.469***	-1.007**	-0.529	-1.045***	-0.675***
	(0.491)	(0.448)	(0.456)	(0.433)	(0.195)	(0.228)
	[[0.442]]	[[0.010]]	[[0.146]]	[[0.625]]	[[0.003]]	[[0.027]]
Incongruent Treatment x Liberal	1.753**	3.523***	0.773	-0.193	1.648***	0.933***
	(0.850)	(0.807)	(0.854)	(0.797)	(0.278)	(0.361)
	[[0.176]]	[[0.003]]	[[0.771]]	[[0.894]]	[[0.003]]	[[0.076]]
Constant	-0.175	0.541	1.377	1.881	1.137	1.062
	(1.012)	(0.928)	(1.337)	(1.167)	(0.811)	(1.349)
Observations	1,007	1,007	1,007	1,007	1,007	1,007
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table compares linear regressions of WTP in the vignette task (columns (1)-(4)) and for real political fundraising posters (column (6)) on Incongruent Treatment and liberal participants. All regressions include demographic controls. Standard errors are reported in parentheses and numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.15: Positive Willingness to Pay (Binary): Effects of Incongruent Treatment by Ideological Orientation–Experiment 2

Dependent variable:	(1)	(2)	(3)	(4)	(5)
$I(WTP > 0)$	Immigration	Social Security	Abortion	LGBTQ'	All
Incongruent Treatment	-0.033	-0.252***	0.149*	0.183**	0.031
	(0.079)	(0.077)	(0.078)	(0.081)	(0.034)
	[[0.880]]	[[0.010]]	[[0.216]]	[[0.096]]	[[0.671]]
Liberal Participant	-0.191**	-0.396***	-0.228***	-0.094	-0.210***
	(0.082)	(0.078)	(0.077)	(0.078)	(0.033)
	[[0.096]]	[[0.003]]	[[0.020]]	[[0.535]]	[[0.003]]
Incongruent Treatment x Liberal	0.442***	0.807***	0.232	-0.011	0.329***
	(0.145)	(0.142)	(0.141)	(0.145)	(0.045)
	[[0.020]]	[[0.003]]	[[0.359]]	[[0.914]]	[[0.003]]
Constant	0.467**	0.603***	0.378	0.609***	0.526***
	(0.199)	(0.227)	(0.242)	(0.221)	(0.165)
Observations	1,007	1,007	1,007	1,007	1,007
Control Variables	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes

Notes: This table compares Probit regressions of having a positive WTP on Incongruent Treatment and liberal participants. All regressions include demographic controls. Standard errors are reported in parentheses and numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.16: Cooperation and Intolerance–Experiment 2

Dependent variable:	(1)	(2)
Decision to cooperate	Binary Person 1 tolerance	Person 1 tolerance
Tolerant Partner	0.962**	
	(0.382)	
	[[0.860]]	
Staircase		-0.085**
		(0.034)
		[[0.860]]
Constant	0.726	5.222***
	(2.155)	(1.481)
Control Variables	Yes	Yes
Observations	1,945	1,945
State FE	Yes	Yes
Individual FE	Yes	Yes

Notes: This table reports Probit regressions of the decision to cooperate in the Prisoner's Dilemma played in Experiment 2, with standard errors in parentheses. Numbers in double square brackets represent Romano-Wolf p-values corrected for multiple hypothesis testing. In model (1) Tolerant Partner is a dummy variable equal to 1 if the partner's Staircase score is below median, and 0 otherwise. In model (2), this dummy is replaced with the partner's Staircase score itself. Each participant yields two observations, one for each partner type. All regressions include demographic controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D Appendix D - Individual-Level Structural Estimation

Procedure. To understand individual-level heterogeneity in intolerance intensity (α_j) and distance curvature (γ_j), we recover these parameters for each participant separately using a grid-search approach.

We first divide out the known spread component by defining:

$$\text{WTP}_{jk}^{\text{adj}} \equiv \frac{\text{WTP}_{jk}}{1 + \hat{\beta}(n_j - 1)}$$

with $\hat{\beta}$ fixed at the model (1) estimate, reducing the within-person equation to $\text{WTP}_{jk}^{\text{adj}} = \alpha_j \cdot d_{jk}^{\gamma_j}$.

For each person j, we search over a fine grid of $\gamma \in [-2, 5]$ in steps of 0.01. For each candidate γ, the optimal α conditional on γ is available in closed form:

$$\hat{\alpha}(\gamma) = \frac{\sum_k \text{WTP}_{jk}^{\text{adj}} \cdot d_{jk}^{\gamma}}{\sum_k d_{jk}^{2\gamma}}$$

and we select the γ minimizing the residual sum of squares $\sum_k (\text{WTP}_{jk}^{\text{adj}} - \hat{\alpha}(\gamma) \cdot d_{jk}^{\gamma})^2$. This jointly recovers both $\hat{\gamma}_j$ and $\hat{\alpha}_j$.

The grid search operates on raw WTP and retains WTP = 0 observations, which are informative: a person who pays nothing at moderate distance but positive amounts at high distance reveals steep curvature. The eligible sample consists of 455 participants from Experiment XXX—those with $d > 0$ for at least two vignettes at distinct distances, and at least one positive WTP. Because $\hat{\alpha}_j$ and $\hat{\gamma}_j$ are estimated jointly, neither parameter is confounded by the other.

Estimating Individual-Level $\hat{\alpha}_j$. For each person's optimal $\hat{\gamma}_j$, the corresponding intolerance intensity is:

$$\hat{\alpha}_j(\hat{\gamma}_j) = \frac{\sum_k \text{WTP}_{jk}^{\text{adj}} \cdot d_{jk}^{\hat{\gamma}_j}}{\sum_k d_{jk}^{2\hat{\gamma}_j}}$$

Table D.1 reports the distribution, and Figure D.1 displays the Kernel density of individual-level intolerance $\hat{\alpha}_j$, by ideological orientation. The mean $\hat{\alpha}_j$ is 8.53 overall (SD = 18.21); the median is 4.58.

Estimating Individual-Level $\hat{\gamma}_j$. Table D.2 reports the distribution, and Figure D.2 displays the Kernel density of individual-level curvature $\hat{\gamma}_j$, by ideological orientation. The mean $\hat{\gamma}_j$ is 1.93 overall, with a median of 1.41, indicating that most individuals exhibit convex or near-linear disutility in distance.

Predictive Validity The recovered $\hat{\alpha}_j$, and $\hat{\gamma}_j$, estimated from vignette-based WTP, can be used to predict poster-blocking WTP, a separate incentivized behavior elicited three days before the 2024 U.S. presidential election. If the structural model captures meaningful latent traits, these person-level parameters should predict out-of-sample behavior:

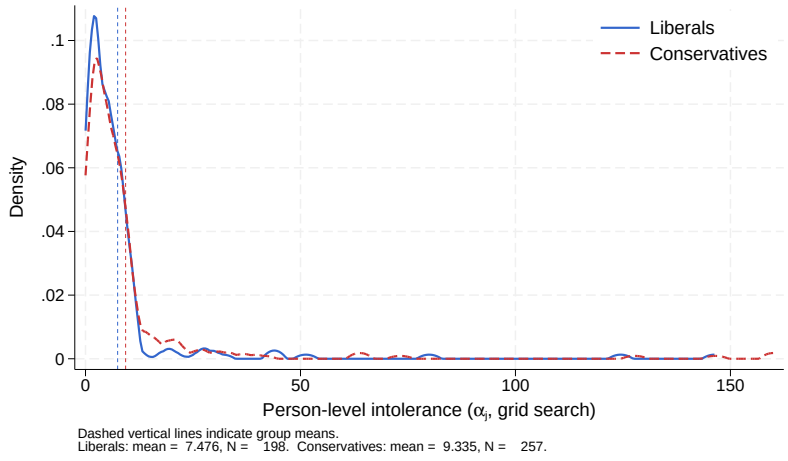
$$\text{PosterWTP}_j = \delta_0 + \delta_1 \hat{\alpha}_j + \delta_2 \hat{\gamma}_j + X_j'\delta + u_j \tag{14}$$

Table D.3 shows that the estimated intolerance intensity, $\hat{\alpha}_j$, is not significantly associated with poster-blocking WTP in any specification ($p > 0.30$). By contrast, the curvature parameter, $\hat{\gamma}_j$, is a robust predictor: each one-unit increase in $\hat{\gamma}_j$ is associated with a \$0.21 lower willingness to pay to block the posters ($p = 0.002$), and the estimate is virtually unchanged after controlling for ideological orientation and demographic characteristics.

Table D.1: Distribution of Individual-Level Intolerance Parameter $\hat{\alpha}_j$ (Grid Search)

	All	Liberals	Conservatives
Mean	8.526	7.476	9.335
Std. Dev.	18.211	15.712	19.917
Median	4.580	4.461	4.649
25th percentile	1.912	1.696	2.005
75th percentile	8.238	7.329	8.563
% with $\hat{\alpha}_j > 0$	100.0%	100.0%	100.0%
Observations	455	198	257
Test: H_0: mean(Liberals) = mean(Conservatives) t = -1.08, p = 0.281			

Notes: $\hat{\alpha}_j$ is estimated jointly with $\hat{\gamma}_j$ via grid search. For each individual's optimal $\hat{\gamma}_j$, the corresponding $\hat{\alpha}_j(\hat{\gamma}_j) = \sum_k \text{WTP}_{jk}^{\text{adj}} \cdot d_{jk}^{\hat{\gamma}_j} / \sum_k d_{jk}^{2\hat{\gamma}_j}$. Same sample as Table D.2 (N = 455).



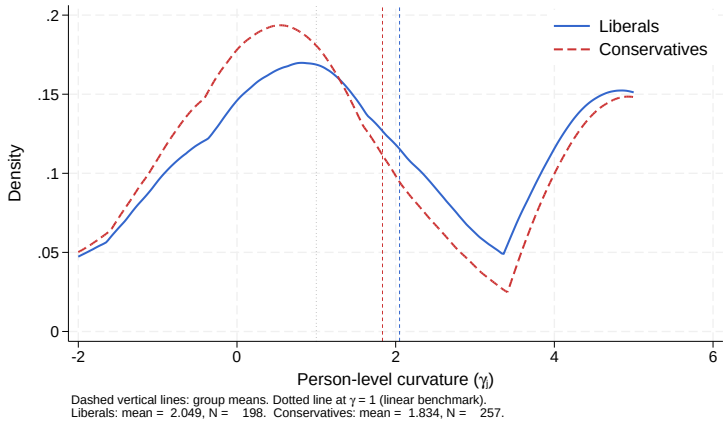
Notes: This figure displays Epanechnikov kernel density with optimal bandwidth. Dashed vertical lines indicate group means. $\hat{\alpha}_j$ is estimated jointly with $\hat{\gamma}_j$ via grid search on raw WTP. Same sample as Table D.2 (N = 455).

Figure D.1: Kernel Density of individual-Level Intolerance $\hat{\alpha}_j$, by Ideological Orientation

Table D.2: Distribution of Individual-Level Curvature Parameter $\hat{\gamma}_j$

	All	Liberals	Conservatives
Mean	1.928	2.049	1.834
Std. Dev.	2.376	2.351	2.396
Median	1.410	1.570	1.160
25th percentile	0.000	0.170	0.000
75th percentile	5.000	5.000	5.000
% with $\hat{\gamma}_j > 1$ (convex)	55.6%	60.1%	52.1%
Observations	455	198	257
Test: H_0: mean(Liberals) = mean(Conservatives)			
$t = 0.95$, $p = 0.340$			

Notes: $\hat{\gamma}_j$ is estimated by grid search over $\gamma \in [-2, 5]$, minimizing $\sum_k (\text{WTP}_{jk}^{\text{adj}} - \alpha_j \cdot d_{jk}^{\gamma_j})^2$ where $\text{WTP}^{\text{adj}} = \text{WTP}/[1 + \hat{\beta}(n_j - 1)]$. For each candidate γ, the optimal $\alpha_j(\gamma)$ is computed in closed form. This retains observations with WTP = 0 (unlike the log-linearized alternative). Restricted to participants with $d > 0$ for ≥ 2 vignettes at ≥ 2 distinct distance values, and ≥ 1 positive WTP. $\hat{\gamma}_j > 1$: convex (accelerating) disutility; $\hat{\gamma}_j < 1$: concave (decelerating).



Notes: This figure displays Epanechnikov kernel density with optimal bandwidth. Dashed vertical lines indicate group means; dotted line at $\gamma = 1$ marks the linear benchmark. $\hat{\gamma}_j$ is estimated via grid search on raw WTP (including WTP = 0 observations). Sample: participants with $d > 0$ for ≥ 2 vignettes at distinct distances and ≥ 1 positive WTP ($N = 455$).

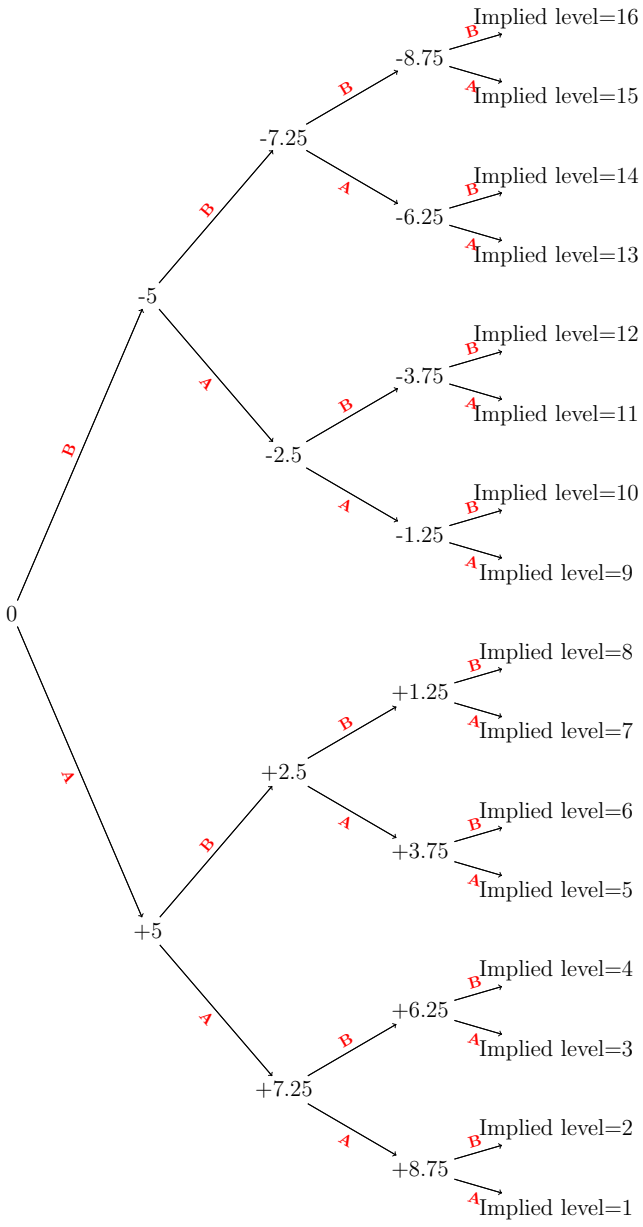
Figure D.2: Kernel Density of Individual-Level Curvature $\hat{\gamma}_j$, by Ideological Orientation

Table D.3: Predictive Validity: Structural Parameters and Poster-Blocking WTP

Dependent variable: WTP to block a real political fundraising poster	(1)	(2)	(3)	(4)
$\hat{\alpha}_j$ (Intolerance intensity)	-0.007	0.002	0.003	0.003
	(0.006)	(0.007)	(0.007)	(0.007)
$\hat{\gamma}_j$ (Distance curvature)		-0.210***	-0.219***	-0.231***
		(0.069)	(0.068)	(0.069)
Conservative participant			-0.624*	-0.453
			(0.322)	(0.323)
Demographics	No	No	No	Yes
Observations	455	455	455	455
R^2	0.001	0.021	0.029	0.101

Notes: This table reports OLS estimates with robust standard errors in parentheses. $\hat{\alpha}_j$ and $\hat{\gamma}_j$ are individual-level structural parameters recovered from WTP to prevent the dissemination of vignettes. Demographic controls include age, gender, education, and income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Appendix E



1

Notes: This figure illustrates the unincentivized Staircase measure of intolerance. Participants completed five consecutive binary choices, with each subsequent question depending on the previous response. The first choice was between (A) receiving a sure payment of $10 without the option to block the dissemination of a view and (B) receiving a sure payment of $10 with the option to block the view from reaching 1 or 100 people, depending on the treatment. The Staircase score ranges from 1 to 16: higher values indicate a greater willingness to forgo earnings to obtain the blocking option (higher intolerance), whereas lower values indicate that participants require increasingly larger compensation to accept the blocking option (higher tolerance).

Figure E.1: Staircase Method of Eliciting Intolerance

F Appendix F - Instructions

F.1 Vignettes (Common for both Experiment 1 and Experiment 2)

The vignettes used in both experiments are given below.

–Immigration

Situation: *Emily stared at the anti-immigration petition that would be submitted to the Congress, trying to decide whether to sign it or not. She had been out of work for months and had no prospects or hope. She blamed the immigrants. They were taking her jobs and lowering her wages. She wanted to sign the petition and make a difference. But she hesitated, as she knew some immigrants. They were not all bad. They were not all illegal. They were hardworking, well-qualified people, they were Americans. She was torn and conflicted.*

Treatment Conservative
Decision: *Finally, she decided to sign the petition.*
View: *In my view, immigration in the US threatens national security, cultural identity, and social cohesion. It undermines the rule of law and hurts the natives in their pursuit of the American dream. Immigration should be discouraged.*

Treatment Liberal
Decision: *Finally, she decided not to sign the petition.*
View: *In my view, immigration in the US adds value, talent, and opportunity. It boosts the labor force, the consumer market, and the tax base. It reflects the spirit and vision of America. Immigration should be encouraged.*

–Social Security

Situation: *Claudia read the petition that mentioned that the people who were out of jobs should be provided social security benefits such as meals at subsidised rate, free housing and access to free health care. Claudia knew that there were many such people who were old or ill. These people needed care and had no support for sustenance. But she also knew there were many people who were voluntarily out of jobs, because they were lazy or didn't have the required qualifications. She didn't want the taxes she was paying from her hard-earned money to benefit such people who were lazy. She was in a dilemma about whether to sign the petition or not.*

Treatment Conservative
Decision: *After much thought, Claudia decided not to sign the petition.*
View: *In my view, the state should keep itself away from providing social security and let the market function, because it is a matter of liberty, efficiency, and personal responsibility. Social security creates a dependency*

culture and discourages people from working hard and saving for their future. Social security also wastes public resources and increases the burden on taxpayers and businesses.

Treatment Liberal
Decision: *After much thought, Claudia decided to sign the petition.*
View: *In my view, the state should support social security because it is a matter of justice, solidarity, and human dignity. Social security ensures that everyone has access to basic needs and opportunities, regardless of their income, health, or social status. Social security also strengthens the social fabric and promotes the common good of society.*

–Abortion

Situation: *Maria sat in her room, the leaflet that had been distributed glaring back at her. The upcoming referendum on abortion was fast approaching, and she had to choose a side: pro-life or pro-choice. Her mind raced – what did she really believe in? Pro-life echoed her concern that human life begins at conception, but pro-choice resonated with her unflinching belief in personal freedom and autonomy. She knew she would have to take a stand and vote, but she was conflicted.*

Treatment Conservative
Decision: *After much deliberation, she decided to vote in favor of Life.*
View: *In my view, abortion or taking away life should not be allowed because it is a moral crime and a health risk. Everyone should protect and respect life from conception.*

Treatment Liberal
Decision: *After much deliberation, she decided to vote in favor of Choice.*
View: *In my view, abortions should be allowed because it is a human right, a personal choice, and a health issue. No one should force or judge women's decisions.*

F.2 Vignettes (Only Experiment 1)

The following vignettes were used only in Experiment 1.

–LGBTQ

Situation: *The school principal was in a dilemma. He had to decide whether to hire Charles as a teacher of Mathematics. Charles was an excellent teacher with an impressive academic record. However, he was also openly gay and lived with his male partner. The principal was worried that hiring Charles would upset some parents and students. He knew that some people in the community were not accepting of homosexuality. The principal had to weigh the benefits of hiring Charles against the potential backlash from the community. It was a difficult decision to make.*

Treatment Conservative

Decision: *After much consideration and extensive consultation with the school management, the school principal decided not to hire Charles as a teacher, because of his homosexuality.*

View: *In my view, sexual preferences should respect morality. Schools are places that shape the future of the society, and hence, such institutions should refuse a job to anyone who has homosexual preferences even if the candidate is qualified and has adequate experience.*

Treatment Liberal

Decision: *After much consideration and extensive consultation with the school management, the school principal decided to hire Charles as a teacher, despite his homosexuality.*

View: *In my view, sexual preferences are individual choices. Schools are places that shape the future of the society, and hence, such institutions should not refuse a job to someone on the basis of his/her homosexual preferences if the candidate is qualified and has adequate experience.*

–Shopping

Situation: *Albert was strolling down the shopping centre and he saw a stylish lifestyle store. He stepped in, rows of neatly folded shirts and pants lined the walls. A dotted, cotton shirt and a pair of wide cotton trousers immediately caught his attention. However, Albert knew he couldn't indulge in both. He carefully weighed his options and eventually made a decision.*

Treatment Conservative

Decision: *Albert decided to buy the shirt as he thought it would complement his existing wardrobe.*

View: *In my view, undoubtedly, a dotted shirt is a stylish addition that any wardrobe should have. It makes the wearer look elegant and chic.*

Treatment Liberal

Decision: *Albert decided to buy the cotton pants as he thought it would complement his existing wardrobe.*

View: *In my view, undoubtedly, a pair of wide trousers is a stylish addition that any wardrobe should have. It makes the wearer look elegant and chic.*

–Color

Situation: *Jacob stands at the crossroads of fashion, pondering between the allure of a crisp blue shirt and the whimsical charm of a soft beige one. Each hue whispers promises of different personas—blue, the epitome of calm sophistication; beige, a daring expression of elegance. Contemplating his choice, Jacob weighs not only color but also the statement each shirt makes about his identity.*

Treatment Conservative

Decision: *Ultimately, he decided to go with the blue shirt, feeling confident in his choice to embrace a*

sophisticated look.
View: *In my view, blue is timeless. No matter what the occasion is, you can never go wrong when you are wearing blue.*

Treatment Liberal
Decision: *Ultimately, he decided to go with the beige shirt, feeling confident in his choice to embrace a bold and novel look.*
View: *In my view, beige is a bold color. On any occasion, a beige shirt asserts itself in its vibrant and unapologetic energy.*

F.3 Vignette (Only Experiment 2)

The following vignette was used only in Experiment 2.

LGBTQ'

Situation: *The human resource manager in a well-known tech company is considering distributing free coffee mugs to employees during Pride Month. The mugs feature a rainbow design and the phrase "Support LGBTQ Rights" printed on them. Some employees may praise the initiative, saying it promotes inclusion and shows the company's commitment to diversity. Others may criticize it, arguing that a company should remain neutral on social matters and avoid endorsing political or ideological messages. The HR manager knows the decision could boost morale among some employees but risk backlash from others. It is a tough call.*

Treatment Conservative
Decision: *After much discussion, the manager decided not to distribute the rainbow-themed mugs, because they believed staying neutral on social matters was important for maintaining harmony among employees.*
View: *Businesses should not actively support the LGBTQ+ movement as this will help maintain harmony and avoid creating clashes among employees or customers. Focusing on core operations rather than ideological messaging ensures inclusivity without endorsing any particular stance.*

Treatment Liberal
Decision: *After much discussion, the manager decided to go ahead and distribute the rainbow-themed mugs, because they believed supporting LGBTQ rights was part of the company's core values.*
View: *Businesses should actively support the LGBTQ+ movement because it promotes equality, inclusion, and respect for human rights. Visible actions, like distributing Pride-themed items, demonstrate a company's commitment to diversity and create a workplace where everyone feels valued.*