Targeted Advertising in Elections[*]

Maria Titova[†]

December 2024

Abstract

How does targeted advertising influence electoral outcomes? This paper presents a one-dimensional spatial model of voting in which a privately informed challenger persuades voters to support him over the status quo. I show that targeted advertising enables the challenger to persuade voters with opposing preferences and swing elections decided by such voters; under simple majority, the challenger can defeat the status quo even when it is located at the median voter's bliss point. Ex-ante commitment power is unnecessary—the challenger succeeds by strategically revealing different pieces of verifiable information to different voters. Publicizing all political ads would mitigate the negative effects of targeted advertising and help voters collectively make the right choice.

KEYWORDS: Persuasion, verifiable information, targeted advertising, elections

JEL CLASSIFICATION: D72, D82, D83

[*]I thank Avidit Acharya, Sandeep Baliga, Arjada Bardhi, Renee Bowen, Steve Callander, Georgy Egorov, Catherine Hafer, Federica Izzo, Aleksandr Levkun, Navin Kartik, Joel Sobel, Francesco Trebbi, Richard Van Weelden and the audience members at Bilkent, CEU, CMU/Pittsburgh, Columbia, Essex, Glasgow, Linz, LSE, NBER Political Economy, NYU, Princeton, SAET, SITE, Stanford GSB, UC Riverside, UC San Diego, UNC Chapel Hill, Warwick. All errors are my own.

[†]Department of Economics, Vanderbilt University. E-mail: motitova@gmail.com.

1. Introduction

Targeted advertising, broadly defined as sending private messages tailored to certain groups of voters, was part of many successful electoral campaigns. In 1960, John F. Kennedy's campaign distributed two million copies of "the blue bomb," a pamphlet advertising his support of civil rights, across African American churches. Decades later, George W. Bush's 2004 reelection campaign used direct mail to highlight his opposition to gay marriage and support for "traditional family values" among evangelical Christians. More recently, the 2016 Brexit referendum and the Trump presidential campaign employed the services of Cambridge Analytica, a data mining firm, to design and distribute thousands of targeted ads to diverse audiences. Although these examples suggest a correlation between targeted advertising and razor-thin electoral success, the precise mechanisms by which tailoring messages to different voters influences election outcomes remain poorly understood.

This paper proposes a simple theoretical model that fills this gap. The model is grounded in three stylized facts about electoral campaigns. First, voters have incomplete information and update their beliefs in response to campaign messages (Kendall, Nannicini, and Trebbi, 2015; Spenkuch and Toniatti, 2018; Le Pennec and Pons, 2023). Second, politicians use the strategy of ambiguity and avoid making precise statements about their positions on issues (Page, 1978; Druckman, Kifer, and Parkin, 2009; Fowler et al., 2021). Third, campaigns tailor messages to specific groups of voters (Hillygus and Shields, 2014). I identify a novel mechanism for exactly how targeted advertising changes electoral outcomes. As it turns out, privately revealing different pieces of information to different voters allows politicians to persuade voters with diametrically opposing preferences and win otherwise unwinnable elections, in which such voters are pivotal.

The model has two components: an advertising campaign followed by an election. In the election, a unit mass of voters chooses between two candidates, the challenger and the status quo. Voters care about the candidates' *policy outcomes*, which represent proposed policies, their implementation or welfare consequences. The challenger's policy outcome $x \in [-1, 1]$ is initially unknown to the voters, while the status quo policy outcome is commonly known and normalized to zero.[1] Voters have quadratic spatial preferences: they are risk averse and prefer to vote for policy outcomes closest to their bliss points. The goal of the office-motivated challenger is to convince a decisive coalition of voters to approve (his metaphorical proposal). I model the challenger's advertising campaign as a game of persuasion with verifiable information (Milgrom and Roberts, 1986). That is, I assume that the challenger privately knows his policy outcome x and can send any subset of $[-1, 1]$ that contains x. Conceptually, this communication protocol allows the challenger to lie by omission but not commission: a message $[-0.5, 0]$ informs a voter that the challenger's policy outcome is moderately left, but is only partially informative because x could be anywhere between -0.5 and 0. Communication with verifiable information is a reasonable middle ground between the possibilities identified by Persson and Tabellini (2002), who famously wrote (p. 483), "It is thus somewhat schizophrenic to study either extreme: where promises have no meaning or where they are all that matter." I consider two versions of the game: public disclosure (PD) and targeted advertising (TA). The former models a public advertising campaign in which the challenger sends the same message to all voters. In the targeted advertising game, the challenger knows the voters' bliss points and can send private messages to different groups of voters.

The main point of the paper is that targeted advertising allows the challenger to win elections that are unwinnable with public disclosure.[2] My first two results (Theorems 1 and 2) characterize elections (described by a set of voters' bliss points and a set of decisive coalitions) that are unwinnable for the challenger with PD and TA, respectively. Theorem 1 states that an election is unwinnable for the challenger with public disclosure if and only if there is no decisive coalition of left or right voters. The intuition is simple: if there is a decisive coalition of left (right) voters, then the challenger can use a fully revealing strategy and win when his policy outcome is left (right) of the status quo. Otherwise, all decisive coalitions include status quo voters (whose bliss point is zero so they always vote for the status quo) or left and right voters (who have diametrically opposing preferences, so the status quo is already the best compromise). Theorem 2 states

[1]Note that the candidates do not choose their policy outcomes (like they do in the Downsian model of electoral competition); they are endowed with them. The assumptions that the status quo is known and not a strategic player vastly simplify exposition and can be loosened in a number of ways (as long as the challenger is the only one advertising *privately*, see Section 5).

[2]I say that an election is unwinnable with PD or TA if the challenger's ex-ante odds of winning are zero in every (perfect Bayesian) equilibrium of the corresponding game. An election is winnable otherwise.

that an election is unwinnable for the challenger with targeted advertising if and only if every decisive coalition includes a status quo voter. In particular, targeted advertising makes it possible for the challenger to convince voters on the opposite sides of the status quo with a positive probability by telling them different things. Under simple majority, Theorems 1 and 2 classify elections as follows:

I. If the median voter's bliss point is left or right or the status quo, then the election is winnable with public disclosure.

II. If the median voter's bliss point is at the status quo but status quo voters do not form a majority, then the election is unwinnable with public disclosure but winnable with targeted advertising.

III. If the median voter's bliss point is at the status quo and status quo voters form a majority, then the election is unwinnable with public disclosure and targeted advertising.

The second part of the paper focuses on the optimal targeted advertising strategy that maximizes the challenger's (ex-ante) odds of winning elections that are unwinnable with public disclosure. I make a further simplification that all left voters have the same bliss point $L < 0$ and all right voters have the same bliss point $R > 0$.[3] Below I use a motivating example in which $L = -0.2$, $R = 0.4$, the minimal decisive coalition includes left and right voters, and $x \sim U[-1, 1]$ to illustrate the remaining two results: Proposition 2 identifies the optimal targeted advertising strategy, while Proposition 3 describes the comparative statics as the right voters become more extreme/the electorate becomes more polarized.

Consider the following strategy of the challenger: to the left voters, he sends the message $[-0.4, 0.2]$ whenever his policy outcome is in $[-0.4, 0.2]$, and the message $[-1, 1]$ otherwise. To the right voters, the challenger sends the message $[-0.4, 0.8]$ whenever his policy outcome is in $[-0.4, 0.8]$ and $[-1, 1]$ otherwise. When a left voter receives message $[-0.4, 0.2]$, she learns that the challenger's policy outcome could be anywhere

[3] As I discuss in Section 4, we can approximate any election by letting L and R be the bliss points of the least extreme left and right voters, respectively. Messages that persuade voter L (R) would also persuade all left (right) voters because more extreme voters are easier to persuade (see Lemma 2).

in $[-0.4, 0.2]$, which is just enough information to convince her to approve.[4] By similar reasoning, a right voter is convinced after message $[-0.4, 0.8]$. These strategies lead to the following electoral outcome. The left voters approve if and only if $x \in [-0.4, 0.2]$ and the right voters approve if and only if $x \in [-0.4, 0.8]$. Given that left and right voters form a decisive coalition, the challenger wins the election if and only if his policy outcome is between -0.4 and 0.2. His ex-ante odds of winning are 30% – a massive improvement over his odds of winning without targeted advertising, which are 0%. Proposition 2 confirms that the described electoral outcome is an equilibrium outcome with the highest odds of the challenger winning across all equilibria of the TA game.

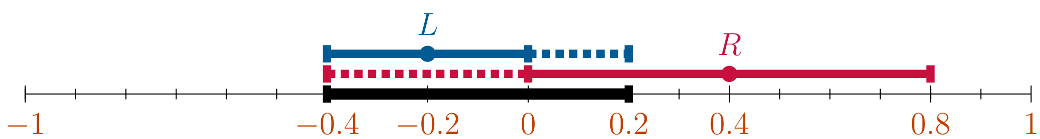


Figure 1. *Targeted messages that convince left voters (in blue) and right voters (in red). The challenger wins the election whenever his policy outcome lies in the intersection of the convincing messages (in black).*

To see how this challenger-preferred equilibrium outcome changes as right voters become more extreme, suppose the right voters' bliss point increases from $R = 0.4$ to $R' = 0.5$. Following the same logic as above, we find that the convincing messages are $[-0.4, 0.2]$ for left voters and $[-0.5, 1]$ for right voters. The challenger wins when his policy outcome is between -0.4 and 0.2, exactly as before. However, the challenger's equilibrium odds of winning can be even higher. Specifically, notice that when his policy outcome is between -0.5 and -0.4, the strategy described above convinces right but not left voters. However, left voters actually prefer policy outcomes in $[-0.5, 0.4]$ to those in $[0.1, 0.2]$ as they are closer to their bliss point. Hence, we can recalculate the message that convinces left voters (making them indifferent between approval and rejection), forcing it to start at -0.5. That message is $[-0.5, 0.179]$. Figure 2 illustrates the electoral outcome after right voters become more extreme.

In the new equilibrium, the set of winning policy outcomes is $[-0.5, 0.179]$ and the

[4] If the prior is $x \sim U[-1, 1]$ and the challenger sends message $[-0.4, 0.2]$ with probability one whenever $x \in [-0.4, 0.2]$, then a left voter's posterior belief (calculated via Bayes rule) after message $[-0.4, 0.2]$ is uniform on $[-0.4, 0.2]$. Her expected utility is $\int_{-0.4}^{0.2} -\frac{(x+0.2)^2}{0.6} dx = -0.04$ if she approves and $-0.2^2 = -0.04$ if she rejects.

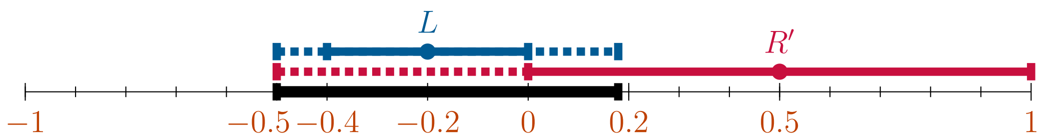


Figure 2. *More extreme right voters are persuadable by policy outcomes further to the left. As a result, the set of challenger's winning policy outcomes (in black) is larger and shifts to the left.*

challenger's odds of winning are 33.96%. Proposition 3 confirms that when right voters become more extreme, the set of winning policy outcomes shift in the opposite direction, to the left; also, the challenger's odds of winning increase. Intuitively, when right voters become more extreme, their dissatisfaction with the status quo grows, which makes them persuadable by wider ranges of policy outcomes.

My findings suggest a novel explanation for why politicians use the strategy of ambiguity — because advertising different ranges of policy outcomes to different voters allows them to persuade voters with diametrically opposing preferences without lying (by commission) to any of them. Previous explanations for why politicians use the strategy of ambiguity include voters' risk-seeking behavior (Shepsle, 1972); candidates' preference for ambiguity (Aragonès and Neeman, 2000); subsequent elections (Meirowitz, 2005; Alesina and Holden, 2008); resolution of uncertainty after the election (Kartik, Van Weelden, and Wolton, 2017). Notably, two previous papers find that ambiguity enables politicians to persuade voters with opposing preferences: in Callander and Wilson (2008), voters have context-dependent preferences, and in Tolvanen (2021), the voters' preferences are correlated with the state of the world. I reach a similar conclusion in a setting where voters have standard quadratic spatial preferences.

This paper builds on the literature comparing public and private communication. When information is verifiable (like in this paper), information unravels whether advertising is public or private if the candidates are symmetric (Janssen and Teteryatnikova, 2017; Schipper and Woo, 2019). However, I show that when candidates are asymmetric – specifically, when only the challenger advertises privately – unraveling does not occur in every equilibrium, and the sender generally prefers private to public communication. In contrast, in cheap talk models, senders often favor public communication because it limits the number of possible deviations in each state of the world (Farrell and Gibbons, 1989, Koessler, 2008, Goltsman and Pavlov, 2011, Bar-Isaac and Deb, 2014). Consequently, targeted advertising cannot swing unwinnable elections if ads consist only of cheap talk.[5] My analysis highlights that persuading voters with opposing preferences requires providing selective evidence or easily verifiable facts. In information design, senders have ex-ante commitment power and generally prefer private communication (Arieli and Babichenko, 2019, Bardhi and Guo, 2018, Chan et al., 2019, Heese and Lauermann, 2021). Since the challenger-preferred equilibrium described in Proposition 2 happens to be a commitment outcome, I also contribute to this strand of literature by identifying the class of elections in which the challenger's odds of winning not only increase, but become positive.

My analysis suggests that targeted advertising is bad for democracy because it elects politicians who are guaranteed to lose when voters possess the same information. For example, under simple majority, targeted advertising allows the challenger to beat the status quo located at the median voter's bliss point, which, according to various versions of the median voter theorem, is unbeatable. The most effective policy to make targeted advertising obsolete is to publicize all the ads transmitted during electoral campaigns. While voters may still make mistakes due to incomplete (but public) information, having a common belief would be sufficient for them to collectively make the right choice.

2. Model

There is a challenger (he/him) and a unit mass of voters (she/her). The space of policy outcomes is $X := [-1, 1]$. Each voter is characterized by her bliss point $v \in X$; I refer to a voter with bliss point $v \in X$ as "voter v" when there is no possibility of confusion. The *election* is a pair $(\mathcal{V}, \mathcal{D})$, where $\mathcal{V} \subseteq X$ is the set of voters' bliss points (the electorate) and $\mathcal{D} \subseteq 2^{\mathcal{V}} \setminus \varnothing$ is the set of decisive coalitions (associated with the social choice rule which I do not model explicitly). I assume that $\mathcal{D}$ is monotonic ($D \in \mathcal{D}$ and $D \subset D' \subseteq \mathcal{V}$ imply $D' \in \mathcal{D}$) and proper ($D \in \mathcal{D}$ implies $\mathcal{V} \setminus D \notin \mathcal{D}$). These assumptions are satisfied for any preference aggregation rule (Austen-Smith and Banks, 2000). I further assume that $\mathcal{V} \in \mathcal{D}$. The game proceeds as follows.

1. The challenger learns his policy outcome $x \in X$ drawn from a common prior distri-

[5] An exception is Schnakenberg (2015) where a cheap-talk sender prefers private communication when the policy space is multi-dimensional, because then he can make statements about different dimensions of the policy to different voters. In contrast, my sender prefers private communication even when the policy space is one-dimensional.

bution $\mu_0 \in \Delta X$ that has a full support and no atoms.[6]

2. The challenger sends messages to voters. Each message is a Borel subset of X (a statement about his policy outcome) that contains a grain of truth, $x \in m$. This communication protocol (introduced by Milgrom and Roberts, 1986) allows the challenger to *lie by omission* and send messages that contain policy outcomes other than x. However, it does not allow the challenger to *lie by commission* and send messages that do not include x. I consider two versions of the game:
   - **public disclosure (PD)**: the challenger chooses a public message m that is the same for all $v \in \mathcal{V}$;
   - **targeted advertising (TA)**: the challenger chooses a collection of private messages $\{m_v\}_{v \in \mathcal{V}}$; voters with bliss point $v \in \mathcal{V}$ observe message m_v only.
3. Each voter decides whether to *approve* the challenger's policy outcome or *reject* it in favor of the status quo. I normalize the status quo (policy outcome) to 0.
4. Payoffs are realized. The challenger is office-motivated: his payoff is 1 if a decisive coalition of voters approves and 0 otherwise. Voters are expressive and have quadratic spatial preferences over policy outcomes.[7] Specifically, when the challenger's policy outcome is $x \in X$, voter v's payoff is $u_v(\text{approve}, x) = -(v-x)^2$ and $u_v(\text{reject}, x) = -(v-0)^2$. I define voter v's *net payoff from approval* as $\alpha_v(x) := u_v(\text{approve}, x) - u_v(\text{reject}, x) = -x^2 + 2vx$ so that v's best response is to approve $x \in X$ whenever $\alpha_v(x) \geq 0$. I let voter v's *approval set* be the set $A_v := \{x \in X \mid \alpha_v(x) \geq 0\}$ of policy outcomes that she prefers to approve under complete information.[8] Figure 3 illustrates the preferences of voter $v < 0$.

I refer to voters with bliss point $v < 0$ as *left voters*; voters $v = 0$ as *status quo voters*; voters $v > 0$ as *right voters*. For any two voters on the same side of the status quo, I say

[6] For a compact metrizable space Y, I let ΔY denote the set of all Borel probability measures over Y, endowed with the weak* topology. I say that $\gamma \in \Delta Y$ is degenerate if $\gamma(y) = 1$ for some $y \in Y$, denoted by $\gamma = \delta_{\{y\}}$, and non-degenerate otherwise. For $W \subseteq X$ such that $\mu_0(W) > 0$, I let $\mu_0(\cdot \mid W) \in \Delta X$ be the prior distribution conditional on W, $\mu_0(x \mid W) := \frac{\mu_0(x) \cdot \mathbb{1}(x \in W)}{\mu_0(W)}$.

[7] Expressive voters derive utility from expressing support (based on ethics, identity, or ideology) for one of the candidates, independent of any effect of the voting act on the electoral outcome. See Brennan and Lomasky (1993) and Hamlin and Jennings (2011) for theory, and Felsenthal and Brichta (1985), Kan and Yang (2001), Artabe and Gardeazabal (2014) for empirical evidence of expressive voting behavior.

[8] Given a subset $W \subseteq X$ of the policy outcome space, I let $W^c := X \setminus W$ be its complement and $\lfloor W \rfloor := \min W$ and $\lceil W \rceil := \max W$ its smallest and largest elements, respectively.

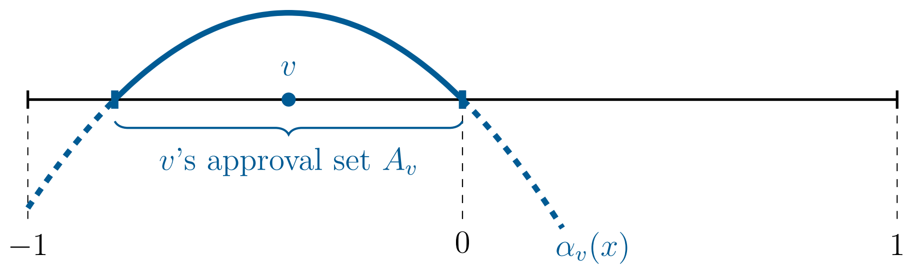


Figure 3. *The policy outcome space $X = [-1, 1]$, the status quo policy outcome 0, a voter's bliss point $v < 0$, her net payoff from approval $\alpha_v(x)$, and her approval set A_v. Under complete information, this voter prefers to approve policy outcomes left, but not too far left, of the status quo.*

that the one with a bliss point closer to the status quo is less extreme:

Definition 1. *A left voter w is* more extreme *than a left voter v if $w < v < 0$. A right voter w is* more extreme *than a right voter v if $0 < v < w$.*

I focus on perfect Bayesian equilibria of this game. In equilibrium, (i) the challenger sends messages that maximize his payoff, (ii) each voter approves the challenger's policy whenever her expected net payoff from approval is non-negative under her posterior belief, and (iii) each voter calculates her posterior using Bayes' rule. I restrict attention to equilibria in which all voters with bliss point $v \in X$ act the same. For ease of exposition, I also assume that status quo voters always vote for the status quo.[9] I refer to the challenger's equilibrium ex-ante utility as his *odds of winning*.

3. Analysis

Public Disclosure

In the public disclosure game, the voters' common prior belief is updated to a common posterior. Therefore, the electorate faces a collective choice problem between a safe option (the status quo) and a lottery over the challenger's policy outcomes (represented by

[9]A status quo voter's net payoff from approval $\alpha_0(x) = -x^2$ is strictly negative unless the challenger's policy outcome coincides with the status quo. The only belief under which this voter weakly prefers to approve is $\delta_{\{0\}}$. While there exist equilibria in which status quo voters approve if and only if $x = 0$, the prior measure of that event is zero since μ_0 is atomless.

their common posterior belief $\mu \in \Delta X$). The first result describes which elections are "unwinnable" for the challenger with public disclosure.

THEOREM 1. *The challenger's odds of winning are zero in every equilibrium of the public disclosure game if and only if there is no left or right decisive coalition.*

Proof. We prove necessity by contraposition. Suppose that there exists a decisive coalition $D \in \mathcal{D}$ of voters on the same side of the status quo. Then, there exists a full disclosure equilibrium in which the challenger sends message $\{x\}$ for each $x \in X$ and voters approve whenever x is in their approval set. The challenger's odds of winning are positive because $\mu_0\Big(\bigcap_{v \in D} A_v\Big) > 0$ for the decisive coalition of left or right voters.

We prove sufficiency directly. Two cases are possible: (1) every decisive coalition includes a status quo voter, or (2) there exists a decisive coalition $D \in \mathcal{D}$ that consists of left and right voters. In both cases, the challenger convinces a decisive coalition only if the public belief is $\delta_{\{0\}}$, or whenever $x = 0$, which has zero prior measure. To see this, observe that for any belief other than $\delta_{\{0\}}$, (1) status quo voters strictly prefer to reject, while (2) left and right voters never weakly prefer to approve at the same time. More precisely, for any belief $\mu \in \Delta X \setminus \delta_{\{0\}}$, if a left voter $L < 0$ prefers to approve, we have

$$\mathbb{E}_\mu[\alpha_L(x)] \geq 0 \implies \alpha_L(\mathbb{E}_\mu[x]) > 0 \implies \mathbb{E}_\mu[x] \in (\lfloor A_L \rfloor, \lceil A_L \rceil) \implies$$
$$\mathbb{E}_\mu[x] \notin (\lfloor A_R \rfloor, \lceil A_R \rceil) \implies \alpha_R(\mathbb{E}_\mu[x]) \leq 0 \implies \mathbb{E}_\mu[\alpha_L(x)] < 0 \text{ for all } R > 0,$$

so all right voters strictly prefer to reject, and vice versa. Here, the first and last implications use Jensen's inequality for the strictly concave α_v and non-degenerate μ. We conclude that there is no common belief (with the exception of $\delta_{\{0\}}$) under which a decisive coalition of left and right voters approves. That completes the proof. ∎

Under simple majority, we get a familiar characterization of elections that are unwinnable with public disclosure.

COROLLARY 1. *Under simple majority, the challenger's odds of winning are zero in every equilibrium of the public disclosure game if and only if the status quo is the median voter's bliss point.*

The proof of the corollary is obvious: if there is no left or right decisive coalition, then the status quo is the median voter's bliss point. Note that Corollary 1 is a special case of median voter theorems for collective choice problems under uncertainty. The result holds more generally for a single-peaked and strictly concave net payoff from approval α_v (Shepsle, 1972) and when the voters' utility function u_v satisfies the single-crossing expectational differences property (Kartik, Lee, and Rappoport, 2023). Quadratic spatial preferences assumed in this paper has both of these properties..

Targeted Advertising

The reason why some elections are unwinnable for the challenger with public disclosure is that the status quo beats any lottery over the challenger's policy outcomes. Targeted advertising allows the challenger to induce different beliefs among different voters and win some of these elections. The next result describes which elections are "unwinnable" for the challenger with targeted advertising.

Theorem 2. *The challenger's odds of winning are zero in every equilibrium of the targeted advertising game if and only if every decisive coalition includes a status quo voter.*

For sufficiency, recall that status quo voters always reject. I prove necessity by contraposition. Two cases are possible: (1) there is a left or right decisive coalition and (2) there are no left or right decisive coalitions but there is a decisive coalition $D_{LR} \in \mathcal{D}$ of left and right voters. In case (1), the full disclosure equilibrium described in the proof of Theorem 1 is an equilibrium of the targeted advertising game. For case (2), note that left (right) voters are willing to approve some right (left) policy outcomes as long as their *expected* net payoff from approval is non-negative. Hence, we can construct an equilibrium in which the challenger gets left and right voters to approve intervals of policy outcomes sufficiently close to the status quo. I formalize this argument in the appendix.

Theorems (1) and (2) describe how the challenger advertises his policy outcome depending on the composition of the electorate. If every decisive coalition includes a status quo voter, he loses with public and private advertising. If there is a decisive coalition of left (or right) voters, he wins by advertising publicly and tailoring his messages to the decisive group. If no decisive coalition includes only left (or only right) voters but some decisive coalition does not include status quo voters, then the challenger can win with targeted but not public advertising. In particular, targeted advertising allows the challenger to beat the status quo policy outcome that much of the political economy literature deems unbeatable.

4. Baseline Election

While Theorems (1) and (2) characterize which elections are winnable with public disclosure and targeted advertising, they do not make a unique prediction of how the challenger wins these elections. Specifically, the proofs of both theorems involve providing an example of an equilibrium in which the challenger's odds of winning are positive. The reason is that the model admits multiple equilibria and there are two sources of multiplicity. Firstly, there may be multiple decisive coalitions. Secondly, the verifiable disclosure game has a range of equilibrium outcomes even if there is only one receiver (Titova, 2023). To move forward in the analysis, I consider a class of baseline elections in which the minimal decisive coalition is unique. Furthermore, I focus on the challenger-preferred equilibrium in order to provide the upper bound on his odds of winning across all equilibria.

Definition 2. *A baseline election has electorate* $\{L, 0, R\}$*, where* $-1 \leq L < 0 < R \leq 1$*.*

In the baseline election, all left voters have the same bliss point $L < 0$ and all right voters have the same bliss point $R > 0$. This assumption limits the number of possible decisive coalitions and allows us to focus on the messages to be sent to left (right) voters, all of whom have the same bliss point. A baseline electorate $\{L, 0, R\}$ can be viewed as an approximation of a general electorate $\mathcal{V}$ by letting $L = \max\limits_{v \in \mathcal{V},\, v<0} v$ and $R = \min\limits_{v \in \mathcal{V},\, v>0} v$ be the least extreme left and right voters in $\mathcal{V}$, respectively. Whenever the challenger convinces L and R, he also convinces their more extreme counterparts, because more extreme voters are easier to persuade.[10]

Given our focus on the challenger-preferred equilibrium, it is useful to first find the highest probability of convincing a voter. Consider the following auxiliary problem with parameters l and r such that $-1 \leq l \leq \lfloor A_v \rfloor < \lceil A_v \rceil \leq r \leq 1$:

$$I_v(l, r) := \arg\max_{I \subseteq [l,r]} \int_I d\mu_0(x) \quad \text{subject to} \quad \int_I \alpha_v(x) d\mu_0(x) \geq 0. \tag{AUX}$$

Roughly speaking, Problem (AUX) identifies the largest (in terms of prior measure) set of policy outcomes I such that, if voter v learns that $x \in I$ and nothing else, she prefers to approve. The objective function is the probability of convincing the voter, while the

[10] I formally show that in Lemma 2 in the comparative statics section.

obedience constraint $\int_I \alpha_v(x)d\mu_0(x) \geq 0$ ensures that the voter's expected net payoff from approval is non-negative given her information. Parametrizing the problem with l and r allows us to only look within certain subsets of X. For example, if we are interested in the largest set of right policy outcomes that a left voter $v < 0$ prefers to approve, then we let $l = \lfloor A_v \rfloor$ and $r = 1$.

Problem (AUX) comes from the information design literature (see, e.g., Alonso and Câmara, 2016) and provides a theoretical upper bound on the probability of convincing a Bayesian voter. The solution is an interval characterized by a cutoff value for the voter's net payoff from approval: voter v approves every policy outcome with a not too negative net payoff from approval (i.e., every $x \in X$ for which $\alpha_v(x) \geq -c_v^*$). The cutoff value $c_v^* > 0$ is obtained from the binding obedience constraint. The set $\{x \in [l, r] \mid \alpha_v(x) \geq -c_v^*\}$ is the upper contour set of the strictly concave function $\alpha_v(x)$ and is therefore an interval. Lemma 1 characterizes the solution of the auxiliary problem; the formal proof can be found in the appendix.

Lemma 1. *The solution to Problem* (AUX) *for* $v \neq 0$ *with* $-1 \leq l \leq \lfloor A_v \rfloor < \lceil A_v \rceil \leq r \leq 1$ *is almost surely an interval.*[11] *Furthermore,*

- *if* $\int_l^r \alpha_v(x)d\mu_0(x) \geq 0$, *then* $I_v(l, r) = [l, r]$;
- *otherwise,* $I_v(l, r) = \{x \in [l, r] \mid \alpha_v(x) \geq -c_v^*(l, r)\}$, *where* $c_v^*(l, r) > 0$ *is obtained from* $\int_{I_v(l,r)} \alpha_v(x)d\mu_0(x) = 0$.

Wining Elections with Public Disclosure

Here we find the challenger-preferred equilibrium for a baseline election that is winnable with public disclosure. From Theorem 1, such an election must have a left or right decisive coalition (that is, $\{L\} \in \mathcal{D}$ or $\{R\} \in \mathcal{D}$), and from the properties of $\mathcal{D}$, the minimal decisive coalition is unique.[12] Without loss of generality assume that the unique minimal decisive coalition is $\{L\}$. In this election, targeted advertising is as good as public disclosure. To maximize his odds of winning, the challenger finds the largest subset of

[11] Almost surely with respect to the prior measure μ_0.

[12] If, for example, $\{L\} \in \mathcal{D}$, then $\{L, 0\} \in \mathcal{D}$ (by monotonicity) and $\{R\} = \mathcal{V} \setminus \{L, 0\} \notin \mathcal{D}$ (by properness). Similarly, $\{0\} \notin \mathcal{D}$.

$[-1, 1]$ that L is willing to approve (that is, he solves Problem (AUX) for voter L with parameters $l = -1$ and $r = 1$) and then publicly reveals whether his policy outcome is in that interval or not.

Proposition 1. *Consider a baseline election $(\{L, 0, R\}, \mathcal{D})$ such that $\{L\} \in \mathcal{D}$. Then, the challenger's highest odds of winning across all equilibria of the public disclosure and targeted advertising games are $\mu_0(I_L(-1, 1))$. He achieves these odds by publicly revealing to all voters whether his policy outcome is in $I_L(-1, 1)$ or not.*

I illustrate the equilibrium outcome in Figure 4.[13] Suppose that the challenger sends message $M = I_L(-1, 1)$ when $x \in I_L(-1, 1)$ and message M^c otherwise, thus revealing whether his policy outcome is in $I_L(-1, 1)$ or not. When the public message is M, the decisive coalition prefers to approve, so the challenger wins whenever $x \in I_L(-1, 1)$ and his odds of winning are $\mu_0(I_L(-1, 1))$. Since $I_L(-1, 1)$ by definition maximizes the odds of convincing a Bayesian voter L, the challenger's odds of winning cannot be higher than $\mu_0(I_L(-1, 1))$ is any other equilibrium.

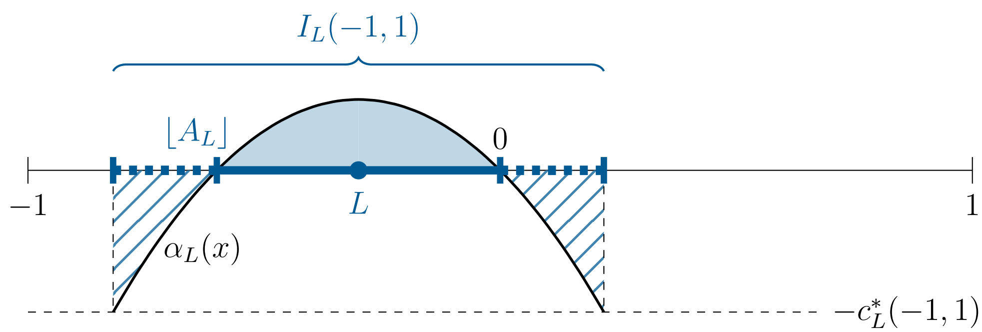


Figure 4. *To maximize his odds of convincing the decisive coalition $\{L\}$, the challenger reveals whether his policy outcome is in $I_L(-1, 1)$. Under uniform prior, c_L^* is obtained from equating the solid area to the dashed area so that voter L is indifferent between approval and rejection when she learns that $x \in I_L(-1, 1)$.*

Swinging Elections with Targeted Advertising

For the remainder of this section, let us focus on a baseline election that is unwinnable with public disclosure but winnable with targeted advertising. From Theorems 1 and 2, the unique minimal decisive coalition is $\{L, R\}$.

[13] Figure 4 presents the numerical solution $I_L(-1, 1) = [-0.82, 0.22]$ for $L = -0.3$ and uniform prior.

Similarly to the previous case, we will build towards an equilibrium in which the challenger reveals to voters $v \in \{L, R\}$ whether $x \in M_v$.[14] Given this strategy of the challenger, when voter v receives message M_v, she learns whether $x \in M_v$ and nothing else, so her best response is determined by her obedience constraint. We build towards a direct equilibrium, in which voter v approves after message M_v and rejects after M_v^c. Then, M_v is both the set of policy outcomes approved by v and the message that convinces her to approve. In a direct equilibrium, the challenger wins whenever his policy outcome is in $M_L \cap M_R$; his odds of winning are $\mu_0(M_L \cap M_R)$. In the challenger-preferred direct equilibrium, his odds of winning are maximized:

$$\begin{gathered} \max_{(M_L, M_R) \subseteq X^2} \int_{M_L \cap M_R} d\mu_0(x) \quad \text{subject to} \\ \int_{M_v} \alpha_v(x) d\mu_0(x) \geq 0 \quad \text{and} \quad \int_{M_v^c} \alpha_v(x) d\mu_0(x) < 0, \quad \forall v \in \{L, R\}. \end{gathered} \tag{1}$$

Once again, Problem (1) describes not only the challenger-preferred direct equilibrium but provides an upper bound on the probability of convincing Bayesian voters L and R.[15] That upper bound is established in information design, wherein the challenger commits not to deviate after learning his policy outcome. However, in this setup, the challenger does not need commitment power because his potential deviations are either unprofitable (when he is winning) or unavailable (because his messages must be verifiable). We find the solution $(\overline{M}_L, \overline{M}_R)$ to Problem (1) next. The first step is calculating the most biased message that convinces each voter to approve.

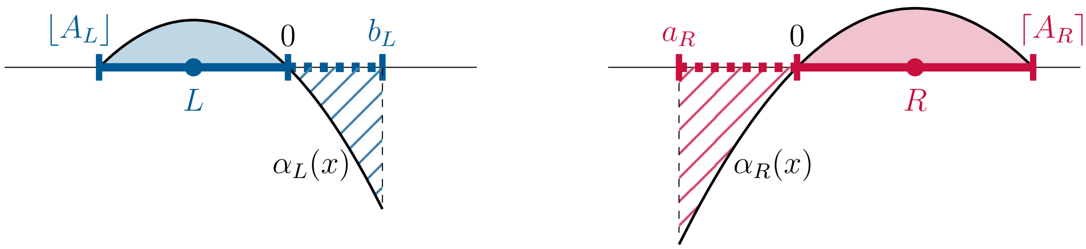
DEFINITION 3. *The largest asymmetric interval of approved policy outcomes I_v of voter v is*

- *if $v = L < 0$, then $I_L = [\lfloor A_L \rfloor, b_L] := I_L(\lfloor A_L \rfloor, 1)$ is the solution of Problem (AUX) for L with $l = \lfloor A_L \rfloor$ and $r = 1$;*
- *if $v = R > 0$, then $I_R = [a_R, \lceil A_R \rceil] := I_R(-1, \lceil A_R \rceil)$ is the solution of Problem (AUX) for R with $l = -1$ and $r = \lceil A_R \rceil$.*

[14] Formally, the challenger's strategy is to send the collection of messages (m_L, m_R) with probability one depending on x, where (m_L, m_R) equals (M_L, M_R) if $x \in M_L \cap M_R$; (M_L, M_R^c) if $x \in M_L \cap M_R^c$; (M_L^c, M_R) if $x \in M_L^c \cap M_R$; (M_L^c, M_R^c) if $x \in M_L^c \cap M_R^c$.

[15] I formalize this argument in the proof of Proposition 2 in the appendix.

Simply put, I_L includes L's approval set $[\lfloor A_L \rfloor, 0]$ plus as many right policy outcomes $(0, b_L]$ as her obedience constraint allows. Similarly, I_R includes R's approval set $[0, \lceil A_R \rceil]$ plus the largest set $[a_R, 0)$ of left policy outcomes satisfying her obedience constraint. Figure 5 illustrates these intervals for $L = -0.2$ and $R = 0.25$ and uniform prior.[16]



(a) $I_L = [\lfloor A_L \rfloor, b_L]$. ***(b)*** $I_R = [a_R, \lceil A_R \rceil]$.

Figure 5. *Largest asymmetric intervals of approved policy outcomes of voters L and R. Under uniform prior, b_L and a_R are obtained from equating solid and dashed areas.*

One might guess that sending each voter her most biased message, that is, letting $\overline{M}_L = I_L$ and $\overline{M}_R = I_R$, maximizes the challenger's odds of winning. It is indeed optimal if $I_L \cap I_R = [a_R, b_L]$ which is when $a_R \geq \lfloor A_L \rfloor$ and $b_L \leq \lceil A_R \rceil$. It is easy to see that the challenger's odds of winning cannot be improved upon $\mu_0([a_R, b_L])$: every policy outcome outside of $[a_R, b_L]$ is further away from at least one voter's bliss point, which makes them more "costly" in terms of that voter's obedience constraint. Figure 6 illustrates this challenger-preferred equilibrium outcome for $L = -0.2$, $R = 0.25$ and uniform prior.

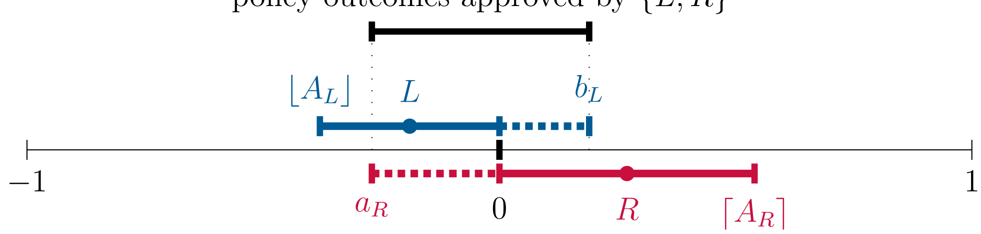


Figure 6. *The electoral outcome when the challenger reveals to left voters whether his policy is in $[\lfloor A_L \rfloor, b_L]$ and to right voters whether his policy outcome is in $[a_R, \lceil A_R \rceil]$. The decisive coalition $\{L, R\}$ approves policy outcomes in $[a_R, b_L]$.*

[16] Here, $I_L = [-0.4, 0.2]$ and $I_R = [-0.25, 0.5]$. Notably, when μ_0 is uniform, then for $L \geq -0.5$ we have $b_L = -L$ and for $R \leq 0.5$ we have $a_R = -R$.

Next, consider the case when $a_R < \lfloor A_L \rfloor$ and $b_L \leq \lceil A_R \rceil$. Now, it is easy to see that sending the most biased message to each voter no longer maximizes the challenger's odds of winning. Indeed, R is now willing to approve L's entire approval set plus policy outcomes left of $\lfloor A_L \rfloor$, which left voters prefer to policy outcomes close to b_L. Hence, $\overline{M}_L$ should start at a_R and span as far right as possible. Formally, in this case, $\overline{M}_L = I_L(a_R, 1)$ so that the challenger wins whenever $x \in I_L(a_R, 1)$. Figure 7 illustrates this outcome for $L = -0.2$, $R = 0.45$ and uniform prior.

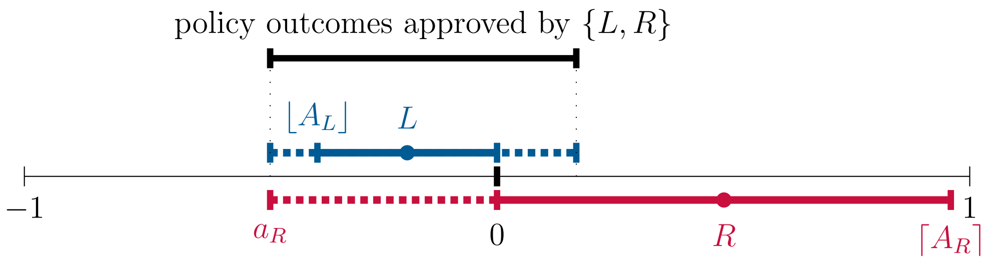


Figure 7. *To maximize the odds of convincing the decisive coalition $\{L, R\}$ when $a_R < \lfloor A_L \rfloor$, the challenger gets left voters to approve the largest subset of $[a_R, 1]$.*

Two cases remain. First, it could be that $a_R \geq \lfloor A_L \rfloor$ and $b_L > \lceil A_R \rceil$. This case is symmetric to the one above. Finally, it is impossible to have $a_R < \lfloor A_L \rfloor$ and $b_L > \lceil A_R \rceil$.[17] The formal result below describes a challenger-preferred equilibrium of a baseline election that is unwinnable with public disclosure but becomes winnable with targeted advertising.

Proposition 2. *Consider a baseline election $(\{L, 0, R\}, \mathcal{D})$ such that $\{L, R\} \in \mathcal{D}$ but $\{L\}, \{R\} \notin \mathcal{D}$. Then, the challenger's highest odds of winning across all equilibria of the targeted advertising game are $\mu_0(\overline{M}_L \cap \overline{M}_R) > 0$, where*

1. $\overline{M}_L = [\lfloor A_L \rfloor, b_L]$, $\overline{M}_R = [a_R, \lceil A_R \rceil]$, $\overline{M}_L \cap \overline{M}_R = [a_R, b_L]$ *if* $a_R \geq \lfloor A_L \rfloor$ *and* $b_L \leq \lceil A_R \rceil$;
2. $\overline{M}_L = I_L(a_R, 1)$, $\overline{M}_R = [a_R, \lceil A_R \rceil]$, $\overline{M}_L \cap \overline{M}_R = \overline{M}_L$ *if* $a_R < \lfloor A_L \rfloor$ *and* $b_L \leq \lceil A_R \rceil$;
3. $\overline{M}_L = [\lfloor A_L \rfloor, b_L]$, $\overline{M}_R = I_R(-1, b_L)$, $\overline{M}_L \cap \overline{M}_R = \overline{M}_R$ *if* $a_R \geq \lfloor A_L \rfloor$ *and* $b_L > \lceil A_R \rceil$.

He achieves these odds of winning by revealing to voter $v \in \{L, R\}$ whether his policy outcome is in $\overline{M}_v$.

[17] If $b_L > \lceil A_R \rceil$ and $a_R < \lfloor A_L \rfloor$, then left and right voters prefer to approve under belief $\mu_0(\cdot \mid [\lfloor A_L \rfloor, \lceil A_R \rceil])$, which contradicts the proof of Theorem 1 (left and right voters never prefer to approve under a common non-degenerate belief).

The formal proof of this result is in the appendix and involves three steps. At Step 1, I confirm that the described $\overline{M}_L$ and $\overline{M}_R$ solve Problem (1). At Step 2, I characterize the direct equilibrium which involves describing voters' skeptical off-path beliefs and showing that no players have profitable deviations. At Step 3, I show that $\mu_0(\overline{M}_L \cap \overline{M}_R)$ is the upper bound on the challenger's odds of winning across all communication protocols.

Comparative Statics

Proposition 2 suggests that the shape of the equilibrium set of winning policy outcomes depends on the voters' bliss points and we explore that relationship here. First, observe that more extreme voters are willing to approve wider ranges of policy outcomes.

Lemma 2. *Suppose that w is a more extreme voter than v. Then, $I_w \supseteq I_v$ and*

- *if $0 < v < w$, then $\lfloor I_w \rfloor = a_w \leq a_v = \lfloor I_v \rfloor$; the inequality is strict unless $a_v = -1$;*
- *if $w < v < 0$, then $\lceil I_w \rceil = b_w \geq b_v = \lceil I_v \rceil$; the inequality is strict unless $b_v = 1$.*

The intuition behind the proof of Lemma 2 is illustrated in Figure 8 for right voters. For the sake of the argument, suppose that we have one right voter who becomes more extreme, meaning that her bliss point increases from $R > 0$ to $E > R$. Then, two effects occur. On the one hand, the voter's approval set expands, and the expected value of her net payoff from approval over her approval set strictly increases. On the other hand, due to the voter's aversion to risk, her net payoff from approval to the left of her approval set strictly decreases. These two effects work in opposite directions, meaning that a_R decreases if the former effect dominates and increases if the latter does. It turns out that the quadratic utility is not concave enough for the latter effect ever to dominate.

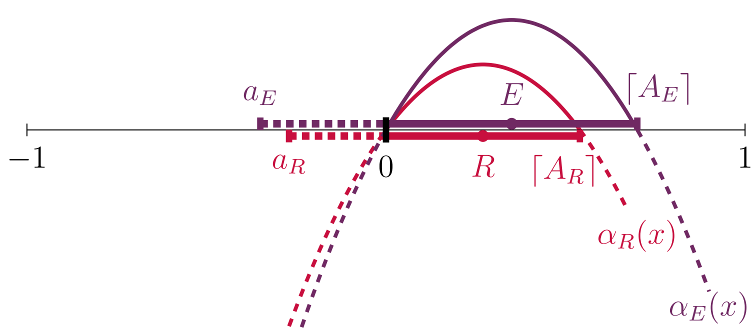


Figure 8. *A right voter becomes more persuadable as her bliss point increases from R to E: her largest asymmetric interval of approved policy outcomes $[a_R, \lceil A_R \rceil]$ expands and its left boundary a_R strictly decreases.*

Next, let us explore how the challenger-preferred equilibrium outcome changes as right voters become more extreme.[18] Note that making right voters more extreme makes the electorate more polarized (Esteban and Ray, 1994).

Proposition 3. *Consider the targeted advertising game with a baseline election* $(\{L, 0, R\}, \mathcal{D})$ *such that* $\{L, R\} \in \mathcal{D}$ *but* $\{L\}, \{R\} \notin \mathcal{D}$. *Let* $(\overline{M}_L, \overline{M}_R)$ *be the challenger-preferred equilibrium intervals of approved policy outcomes described in Proposition 2. Suppose that* $b_L \leq \lceil A_R \rceil$. *Then, as* R *increases,*

- *the challenger's odds of winning* $\mu_0(\overline{M}_L \cap \overline{M}_R)$ *increase;*
- *the set of winning policy outcomes* $\overline{M}_L \cap \overline{M}_R$ *shifts to the left, that is,* $\lfloor \overline{M}_L \cap \overline{M}_R \rfloor$ *and* $\lceil \overline{M}_L \cap \overline{M}_R \rceil$ *decrease.*

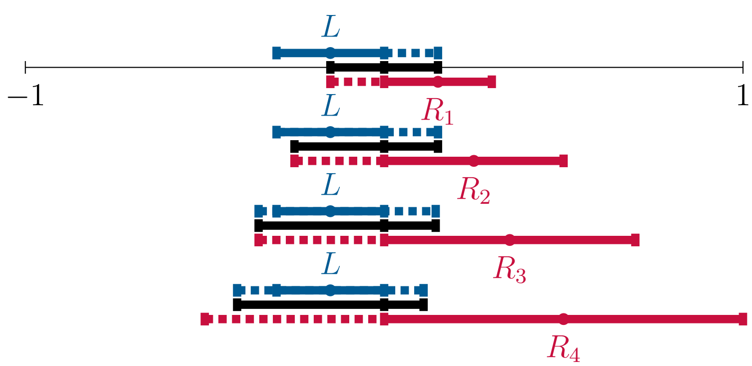


Figure 9. *The challenger-preferred equilibrium outcome as right voters become more extreme (top to bottom). Right voters approve ranges of policy outcomes (in red) that span further left, and the set of winning policy outcomes (in black) shifts left.*

Figure 9 illustrates the equilibrium outcomes of four baseline elections, holding the left voters' bliss point L fixed and increasing the right voters' bliss point from R_1 to R_4 (top to bottom).[19] From Lemma 2, as right voters become more extreme, their largest asymmetric interval of approved policy outcomes expands, and that has two consequences. First, these voters are now more persuadable, which means that the challenger's odds of

[18] Specifically, we compare the challenger-preferred equilibria of baseline elections with electorates $\{L, 0, R\}$ and $\{L, 0, E\}$, both of which satisfy the requirements of Proposition 2, and $R < E$.

[19] Figure 9 presents the numerical solution for the uniform prior, with bliss point $L = -0.15$ for left voters and successive bliss points $R_1 = 0.15$, $R_2 = 0.25$, $R_3 = 0.35$, and $R_4 = 0.50$ (top to bottom) for the right voters. The sets of winning policy outcomes (in black) are $[-0.15, 0.15]$, $[-0.25, 0.15]$, $[-0.35, 0.1436]$, and $[-0.4098, 0.1098]$, respectively.

winning go up. Second, more extreme right voters approve policy outcomes further to the left. As a result, the left boundary of the equilibrium set of winning policy outcomes shifts to the left, as well. Interestingly, the right endpoint of the equilibrium set of winning policy outcomes, which is determined by left voters, may strictly decrease, as well. That happens when right voters are or become persuadable by policy outcomes left of $\lfloor A_L \rfloor$ (e.g., a change from R_2 to R_3 or R_3 or R_4 in Figure 9).

Voter Welfare

Next, we explore voter welfare in the challenger-preferred equilibrium of the targeted advertising game as right voters become more extreme. Suppose that $\overline{M}_L$ and $\overline{M}_R$ are such that obedience constraints bind for each $v \in \{L, R\}$. Define voter v's welfare as her expected utility:

$$\begin{aligned} W_v(\overline{M}_L, \overline{M}_R) &:= \int\limits_{\overline{M}_v} u_v(\text{approve}, x) d\mu_0(x) + \int\limits_{\overline{M}_v^c} u_v(\text{reject}, x) d\mu_0(x) \\ &= \int\limits_{\overline{M}_v} \alpha_v(x) d\mu_0(x) - v^2 = -v^2. \end{aligned}$$

In words, v's welfare in the challenger-preferred equilibrium is just her payoff from rejection since the challenger's policy outcome is expected to be as good as the status quo when the obedience constraint binds. We immediately conclude that as R increases, right voters' welfare decreases, while left voters' welfare does not change.

To summarize comparative statics, when the electorate becomes more polarized, the challenger is better off while the voters who become more extreme are worse off. This is a novel formulation of the familiar result of Romer and Rosenthal (1978) and subsequent literature on veto bargaining which finds that the proposer is better off and the voter is worse off when the voter's bliss point moves away from the status quo. In that literature, these insights follow from the proposer's increased power of agenda control. In this paper, they follow from the challenger's increased power of persuasion.

5. Discussion and Conclusion

This paper considered a voting model in which (1) there are two candidates, one of whom (the incumbent) is non-strategic, (2) the strategic candidate (the challenger) advertises his policy outcome using verifiable messages, and (3) the voters are expressive. Assuming (1) and (2), I identified elections that are unwinnable for the challenger if the voters hold a common belief (Theorem 1). Assuming (3), I showed that the challenger can win these elections by advertising privately. Below I discuss the role of assumptions (1) and (3) to deliver the main result of the paper, which is that private advertising swings unwinnable elections. For brevity, all modifications of the model that I consider below have two possible bliss points of the voters, $L < 0$ and $R > 0$.

Non-strategic Incumbent

In the model, the incumbent does not advertise and his policy outcome is known and normalized to zero. This assumption can be relaxed in a number of ways.

First, suppose that the status quo is a lottery $\nu_0 \in \Delta[-1, 1]$ (independent of μ_0) and the incumbent is still non-strategic. Then, voter v's expected payoff of rejection is $\int -(v-y)^2 d\nu_0(y)$ (instead of $-v^2$). While there is no such thing as "left" and "right" voters anymore (those were defined relative to 0), we can still define voters v and w as having "diametrically opposing preferences" if $-\int(v-x)^2 d\mu(x) \geq -\int(v-y)^2 d\nu_0(y)$ implies $-\int(w-x)^2 d\mu(x) < -\int(w-y)^2 d\nu_0(y)$ for all $\mu \in \Delta X$, meaning that at most one of these voters prefers to approve given the choice between ν_0 and any μ. Then, an election is unwinnable for the challenger without targeted advertising if every decisive coalition requires convincing such voters. With targeted advertising, the challenger induces different posteriors among different voters and is still able to convince voters with diametrically opposing preferences with a positive probability.

Next, suppose that the incumbent is strategic and can change ν_0 to a common belief ν about the status quo policy outcome, perhaps by publicly advertising it. Assuming that the challenger has time to react, he still benefits from targeted advertising for the same reason as in the above paragraph. In fact, even if the incumbent could choose the status quo policy outcome, the challenger can win as long as not every decisive coalition includes a status quo voter.

Finally, if the candidates are symmetric (e.g. they both use targeted advertising to advertise their own and/or their opponent's policy outcome), then full unraveling of information takes place (Janssen and Teteryatnikova, 2017; Schipper and Woo, 2019). Therefore, the general message of this paper is that having access to a better targeted advertising technology and/or better voter data allows politicians to win otherwise unwinnable elections. Without that advantage, the voters choose the same candidate as under complete information.

Expressive Voting

The expressive voting assumption enters when we assume that voter v's utility function depends on her action (approve/reject) and the policy outcome of the candidate that she votes for (which is x if she approves, 0 if she rejects). Importantly, an expressive voter's utility does not depend on the identity of the *winner* of the election; expressive voters vote to express support for their most preferred candidate, rather than to change the outcome of the election. Instrumental voters, on the other hand, care about who wins the election, which means that they take into consideration the events in which their vote is pivotal. If an election involves a large number of voters, expressive voting is a standard assumption with strong empirical support (see footnote 5): each individual vote is unlikely to change the outcome of the election, which means that voters vote for other reasons, for example, to express their preference.

Now, if voters are instrumental and they know that every decisive coalition includes left and right voters, then targeted advertising is as good as public disclosure. To see this, suppose that there are two voters, L and R, and the challenger wins if and only if he convinces both of them to approve.[20] An instrumental voter's payoff when the challenger's policy outcome is $x \in X$ is $-(v-x)^2$ if both voters approve and $-v^2$ if someone rejects.[21] Suppose that the challenger uses a direct strategy and reveals to voter $v \in \{L, R\}$ whether his policy outcome is in $M_v \subseteq X$, and v approves after message M_v.

[20] Note that if these voters were expressive, then Proposition 2 still describes the challenger-preferred equilibrium of the targeted advertising game.

[21] Note that an instrumental voter's payoff now depends on the policy outcome of the winning candidate; if v approves but the other voter rejects, then v's payoff is $-v^2$ and not $-(v-x)^2$ if v was expressive.

For voter L to approve, we need

$$\underbrace{\int_{M_L \cap M_R} -(L-x)^2 d\mu_0(x)}_{\text{L and R approve}} + \underbrace{\int_{M_L \cap M_R^c} -L^2 d\mu_0(x)}_{\text{L approves, R rejects}} \geq \underbrace{\int_{M_L} -L^2 d\mu_0(x)}_{\text{L rejects}},$$

which is equivalent to $\int_{M_L \cap M_R} \alpha_L(x) d\mu_0(x) \geq 0$. The intuition is simple: when L hears message M_L and takes into account that R approves whenever $x \in M_R$, the event that she is pivotal occurs when the challenger's policy is in $M_L \cap M_R$. For her to approve, her expected net payoff from approval over that region of policy outcomes must be non-negative. Similarly, the right voter prefers to approve whenever $\int_{M_L \cap M_R} \alpha_R(x) d\mu_0(x) \geq 0$. However, both of these inequalities cannot hold at the same time (see the proof of Theorem 1). Therefore, the challenger loses if he uses a direct strategy. The same argument holds for all strategies of the challenger: instrumental voters learn additional information from conditioning on the event of being pivotal and end up with the same belief in that event. Consequently, elections unwinnable with public disclosure remain unwinnable with targeted advertising. This result is reminiscent of the no-trade theorem of Milgrom and Stokey (1982) — risk-averse instrumental voters are not willing to "trade" the Pareto optimal status quo for any other policy outcome. Each voter learns she would be worse off from trading simply from conditioning on the willingness of her counterpart with opposing preferences to trade.

However, if the voters are instrumental but they are uncertain about the preferences of other voters or what the decisive coalitions are, then targeted advertising still swings otherwise unwinnable elections. To see this, suppose that there are two jointly pivotal voters who have one of two possible bliss points, L or R. At the beginning of the game, nature uniformly draws an electorate from the set $\{LL, LR, RL, RR\}$. The challenger learns what the electorate is, while each voter only learns her own bliss point (and finds it equally likely that the other voter is left or right). Suppose that the challenger reveals to voter $v \in \{L, R\}$ whether his policy outcome is in $M_v \subseteq X$, and v approves after message

M_v. For voter with bliss point L to approve, we need

$$\underbrace{\int\limits_{M_L\cap M_R} -(L-x)^2 d\mu_0(x)}_{\text{second voter definitely approves}} + \underbrace{\frac{1}{2}\int\limits_{M_L\cap M_R^c} -(L-x)^2 d\mu_0(x)}_{\substack{\text{second voter is left}\\ \text{and approves}}} + \underbrace{\frac{1}{2}\int\limits_{M_L\cap M_R^c} -L^2 d\mu_0(x)}_{\substack{\text{second voter is right}\\ \text{and rejects}}} \geq \int\limits_{M_L} -L^2 d\mu_0(x),$$

which is equivalent to $\int\limits_{M_L} \alpha_L(x)d\mu_0(x) \geq \frac{1}{2}\int\limits_{M_L\cap M_R^c} \alpha_L(x)d\mu_0(x)$. Two observations are in order. First, using the challenger-preferred messages $\overline{M}_L$ and $\overline{M}_R$ described in Proposition 2 intended for expressive voters would no longer work as they violate the inequality above.[22] That is not surprising – instrumental voters are harder to convince than expressive voters because they infer additional information from the pivotality event. The second observation is that voters with opposing bliss points do not have a common belief after conditioning on the event of being pivotal because of uncertainty about the other voter's preferences. It is easy to show that there exist equilibria of this game in which the challenger convinces these voters with a positive probability.[23]

In summary, targeted advertising is unlikely to influence "small" elections – such as committee votes, corporate board decisions, or small-town council elections – where the number of voters is small and they are well-informed about each other's preferences. In contrast, targeted advertising can be highly effective in "large" elections, such as presidential races, where voters are less informed about others' preferences and the likelihood of any single vote being pivotal is extremely low.

Private Messages

The final key assumption of the model is that there are no information spillovers, meaning that the challenger's targeted ads stay private. If left and right voters observed each others' messages, they would learn the same information, and that would make targeted advertising as good as public disclosure. Therefore, informing voters of all ads transmitted

[22] The LHS of the inequality is $\int\limits_{\overline{M}_v} \alpha_v(x)d\mu_0(x) = 0$ (v's obedience constraint is binding to maximize probability of convincing this voter), while the RHS is $\int\limits_{\overline{M}_L\cap\overline{M}_R^c} \alpha_L(x)d\mu_0(x) > 0$ since $\overline{M}_L \cap \overline{M}_R^c \subset A_L$.

[23] For example, $M_L = [\lfloor A_L \rfloor, \varepsilon]$ and $M_R = [-\varepsilon, \lceil A_R \rceil]$ for $\varepsilon > 0$ that is sufficiently small satisfy the inequality for both voters.

during an electoral campaign is a useful tool to mitigate the negative effects of targeted advertising. In fact, 1433 targeted ads of the Vote Leave campaign were released in the aftermath of the 2016 Brexit referendum except it was done after the vote was finalized.[24]

References

Alesina, Alberto F. and Richard T. Holden (2008), "Ambiguity and Extremism in Elections", *NBER Working Paper*, 14143 (June 2008). (p. 6.)

Alonso, Ricardo and Odilon Câmara (2016), "Persuading Voters", *American Economic Review*, 106, 11 (Nov. 2016), pp. 3590-3605. (pp. 13, 29.)

Aragonès, Enriqueta and Zvika Neeman (2000), "Strategic Ambiguity in Electoral Competition", *Journal of Theoretical Politics*, 12, 2 (Apr. 1, 2000), pp. 183-204. (p. 6.)

Arieli, Itai and Yakov Babichenko (2019), "Private Bayesian Persuasion", *Journal of Economic Theory*, 182 (July 2019), pp. 185-217. (p. 7.)

Artabe, Alaitz and Javier Gardeazabal (2014), "Strategic Votes and Sincere Counterfactuals", *Political Analysis*, 22, 2 (Jan. 2014), pp. 243-257. (p. 8.)

Austen-Smith, D. and J.S. Banks (2000), *Positive Political Theory I: Collective Preference*, Michigan Studies In Political Analysis, University of Michigan Press. (p. 7.)

Bar-Isaac, Heski and Joyee Deb (2014), "(Good and Bad) Reputation for a Servant of Two Masters", *American Economic Journal: Microeconomics*, 6, 4 (Nov. 2014), pp. 293-325. (p. 6.)

Bardhi, Arjada and Yingni Guo (2018), "Modes of Persuasion toward Unanimous Consent", *Theoretical Economics*, 13, 3, pp. 1111-1149. (p. 7.)

Bergemann, Dirk and Stephen Morris (2016), "Information Design, Bayesian Persuasion, and Bayes Correlated Equilibrium", *American Economic Review*, 106, 5 (May 2016), pp. 586-591. (p. 33.)

Brennan, Geoffrey and Loren Lomasky (1993), *Democracy and Decision: The Pure Theory of Electoral Preference*, Cambridge University Press. (p. 8.)

Callander, Steven and Catherine H. Wilson (2008), "Context-Dependent Voting and Political Ambiguity", *Journal of Public Economics*, 92 (Apr. 1, 2008). (p. 6.)

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang (2019), "Pivotal Persuasion", *Journal of Economic Theory*, 180 (Mar. 2019), pp. 178-202. (p. 7.)

[24] In July 2018, Facebook released the ads to UK's Department for Culture, Media and Sport Committee as part of the Committee's inquiry into Fake News.

Druckman, James N., Martin J. Kifer, and Michael Parkin (2009), "Campaign Communications in U.S. Congressional Elections", *American Political Science Review*, 103, 3 (Aug. 2009), pp. 343-366. (p. 2.)

Esteban, Joan-María and Debraj Ray (1994), "On the Measurement of Polarization", *Econometrica*, 62, 4, pp. 819-851. (p. 19.)

Farrell, Joseph and Robert Gibbons (1989), "Cheap Talk with Two Audiences", *The American Economic Review*, 79, 5, pp. 1214-1223. (p. 6.)

Felsenthal, Dan S. and Avraham Brichta (1985), "Sincere and Strategic Voters: An Israeli Study", *Political Behavior*, 7, 4, pp. 311-324. (p. 8.)

Fowler, Erika Franklin, Michael M. Franz, Gregory J. Martin, Zachary Peskowitz, and Travis N. Ridout (2021), "Political Advertising Online and Offline", *American Political Science Review*, 115, 1 (Feb. 2021), pp. 130-149. (p. 2.)

Goltsman, Maria and Gregory Pavlov (2011), "How to Talk to Multiple Audiences", *Games and Economic Behavior*, 72, 1 (May 1, 2011), pp. 100-122. (p. 6.)

Hamlin, Alan and Colin Jennings (2011), "Expressive Political Behaviour: Foundations, Scope and Implications", *British Journal of Political Science*, 41, 3 (July 2011), pp. 645-670. (p. 8.)

Heese, Carl and Stephan Lauermann (2021), "Persuasion and Information Aggregation in Elections", *Working Paper* (Aug. 2021). (p. 7.)

Hillygus, D. Sunshine and Todd G. Shields (2014), *The Persuadable Voter: Wedge Issues in Presidential Campaigns*, Princeton University Press, p. 267. (p. 2.)

Janssen, Maarten C. W. and Mariya Teteryatnikova (2017), "Mystifying but not Misleading: when does Political Ambiguity not Confuse Voters?", *Public Choice*, 172, 3-4 (Sept. 2017), pp. 501-524. (pp. 6, 22.)

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian Persuasion", *American Economic Review*, 101, 6 (Oct. 2011), pp. 2590-2615. (pp. 29, 33.)

Kan, Kamhon and C. C. Yang (2001), "On Expressive Voting: Evidence from the 1988 U.S. Presidential Election", *Public Choice*, 108, 3-4, pp. 295-312. (p. 8.)

Kartik, Navin, SangMok Lee, and Daniel Rappoport (2023), "Single-Crossing Differences in Convex Environments", *The Review of Economic Studies* (Oct. 26, 2023), Forthcoming. (p. 11.)

Kartik, Navin, Richard Van Weelden, and Stephane Wolton (2017), "Electoral Ambiguity and Political Representation", *American Journal of Political Science*, 61, 4, pp. 958-970. (p. 6.)

Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi (2015), "How Do Voters Respond to Information? Evidence from a Randomized Campaign", *American Economic Review*, 105, 1 (Jan. 2015), pp. 322-353. (p. 2.)

Koessler, Frédéric (2008), "Lobbying with Two Audiences: Public vs Private Certification", *Mathematical Social Sciences*, 55, 3 (May 1, 2008), pp. 305-314. (p. 6.)

Le Pennec, Caroline and Vincent Pons (2023), "How do Campaigns Shape Vote Choice? Multicountry Evidence from 62 Elections and 56 TV Debates", *The Quarterly Journal of Economics*, 138, 2 (May 1, 2023), pp. 703-767. (p. 2.)

Meirowitz, Adam (2005), "Informational Party Primaries and Strategic Ambiguity", *Journal of Theoretical Politics*, 17, 1 (Jan. 1, 2005), pp. 107-136. (p. 6.)

Milgrom, Paul and John Roberts (1986), "Relying on the Information of Interested Parties", *The RAND Journal of Economics*, 17, 1, p. 18. (pp. 3, 8.)

Milgrom, Paul and Nancy Stokey (1982), "Information, Trade and Common Knowledge", *Journal of Economic Theory*, 26, 1 (Feb. 1, 1982), pp. 17-27. (p. 23.)

Page, Benjamin I (1978), *Choices and Echoes in Presidential Elections: Rational Man and Electoral Democracy*, University of Chicago Press. (p. 2.)

Persson, Torsten and Guido Tabellini (2002), *Political Economics: Explaining Economic Policy*, MIT Press Books, The MIT Press. (p. 3.)

Romer, Thomas and Howard Rosenthal (1978), "Political Resource Allocation, Controlled Agendas, and the Status Quo", *Public Choice*, 33, 4, pp. 27-43. (p. 20.)

Schipper, Burkhard C. and Hee Yeul Woo (2019), "Political Awareness and Microtargeting of Voters in Electoral Competition", *Quarterly Journal of Political Science*, 14, 1, pp. 41-88. (pp. 6, 22.)

Schnakenberg, Keith E. (2015), "Expert Advice to a Voting Body", *Journal of Economic Theory*, 160 (Dec. 1, 2015), pp. 102-113. (p. 7.)

Shepsle, Kenneth A. (1972), "The Strategy of Ambiguity: Uncertainty and Electoral Competition", *American Political Science Review*, 66, 2 (June 1972), pp. 555-568. (pp. 6, 11.)

Spenkuch, Jörg L and David Toniatti (2018), "Political Advertising and Election Results", *The Quarterly Journal of Economics*, 133, 4 (Nov. 1, 2018), pp. 1981-2036. (p. 2.)

Titova, Maria (2023), "Persuasion with Verifiable Information", *Working Paper*. (p. 12.)

Tolvanen, Juha (2021), "On Political Ambiguity and Anti-Median Platforms", *Working Paper* (Sept. 2021). (p. 6.)

Appendix: Omitted Proofs

Proof of Lemma 1

If $\int_l^r \alpha_v(x) d\mu_0(x) \geq 0$, then $I_v(l,r) = [l,r]$ maximizes the objective. Next, suppose that $\int_l^r \alpha_v(x) d\mu_0(x) < 0$. Then, the constraint binds; otherwise, we could strictly increase the objective while still satisfying that constraint. Next, suppose by contradiction that $W \subseteq [l,r]$ is a solution that is not a.s. characterized by a cutoff value of v's net payoff from approval. Then, there exist two sets $Y, Z \subseteq [l, r]$ such that $\mu_0(Y) = \mu_0(Z)$, $Z \subseteq W$, $Y \cap W = \varnothing$ and $\forall y \in Y$ and $z \in Z$, $\alpha_v(y) < \alpha_v(z)$. Observe that

$$\int_Y \alpha_v(x) d\mu_0(x) < \max_{y \in Y} \alpha_v(y)\mu_0(Y) < \min_{z \in Z} \alpha_v(z)\mu_0(Z) < \int_Z \alpha_v(x) d\mu_0(x).$$

Let $\widetilde{W} := (W \setminus Y) \cup Z$. Observe that $\int_W \alpha_v(x) d\mu_0(x) = 0$ implies that $\int_{\widetilde{W}} \alpha_v(x) d\mu_0(x) > 0$, meaning that the constraint is loose for $\widetilde{W}$. Consequently, $\widetilde{W}$ is not a solution and the maximized objective value must be strictly greater than $\mu_0(\widetilde{W})$. Since $\mu_0(W) = \mu_0(\widetilde{W})$, W is not a solution either, a contradiction.

Proof of Theorem 2

In case (2) of the necessity proof, let $L := \max_{v \in D_{LR},\, v<0} v$ and $R := \min_{v \in D_{LR},\, v>0} v$ be the least extreme left and right voters of the mixed decisive coalitions D_{LR}, respectively. Next, let $M_L := [\lfloor A_L \rfloor, \varepsilon]$ and $M_R := [-\varepsilon, \lceil A_R \rceil]$ for $\varepsilon > 0$ that satisfies $\int_{\lfloor A_L \rfloor}^{\varepsilon} \alpha_L(x) d\mu_0(x) \geq 0$, $\int_{-\varepsilon}^{\lfloor A_R \rfloor} \alpha_R(x) d\mu_0(x) \geq 0$, and $[-\varepsilon, \varepsilon] \subseteq [\lfloor A_L \rfloor, \lceil A_R \rceil]$.[25] Let the challenger's strategy be to send to all left (right) voters message M_L (M_R) when $x \in M_L$ ($x \in M_R$) and message M_L^c (M_R^c) otherwise. When $v \in \mathcal{V}$ hears an off-path message, let her posterior belief be supported on policy outcomes outside of her approval set, whenever possible. Now, the challenger wins (by convincing every voter in D_{LR}) whenever his policy outcome is in $M_L \cap M_R = [-\varepsilon, \varepsilon]$ and his odds of winning $\mu_0([-\varepsilon, \varepsilon])$ are positive since $\varepsilon > 0$. If his policy outcome is outside of that set, he loses but does not have profitable deviations. Indeed, if, say, $x < -\varepsilon$, then any message he may send will not convince the skeptical right voters to approve, but he needs their approval as every decisive coalition without a status quo voter includes a right voter. A similar argument applies to the case when $x > \varepsilon$. We have thus found an equilibrium in which the challenger's odds of winning are positive.

[25] Such ε exists because the function $\phi_L(z) := \int_{\lfloor A_L \rfloor}^{z} \alpha_L(x) d\mu_0(x)$ is continuous and strictly decreasing in z for $z \geq 0$ and $\phi_L(0) > 0$. A similar argument applies to $\phi_R(z) := \int_{\lfloor A_R \rfloor}^{z} \alpha_R(x) d\mu_0(x)$.

Proof of Proposition 1

First, I construct an equilibrium of the PD and TA game in which the challenger's odds of winning are $\mu_0(I_L(-1,1))$. Let the challenger send message $M_L := I_L(-1,1)$ if $x \in M_L$ and message $M_L^c := X \setminus M_L$ otherwise. Off the path, let the left voters' posteriors be supported on policy outcomes outside of A_L, whenever possible. Now, the challenger wins if $x \in M_L$ because M_L satisfies L's obedience constraint. At the same time, L rejects after message M_L^c because $A_L \subseteq M_L$ (From Lemma 1) so that L's net payoff from approval is strictly negative for all $x \in M_L^c$. What voters other than L do on is irrelevant since a coalition is decisive if and only if it includes L. Notice that the challenger does not have profitable deviations. If $x \in M_L$, then he wins the election and gets the highest possible payoff. If $x \notin M_L$, then $x \notin A_L$, so any deviation leads to a rejection by the skeptical left voters. Hence, the described strategies and beliefs form an equilibrium.

To see why the challenger's odds of winning cannot be higher than $\mu_0(I_L(-1,1))$ in any other equilibrium, let us modify the game and allow the challenger to commit to what information each voter gets ahead of learning his policy outcome (as in Kamenica and Gentzkow, 2011). Then, the optimal commitment outcome is that the challenger wins whenever $x \in W$, where W solves Problem (AUX) with $l = -1$ and $r = 1$ and his odds of winning are $\mu_0(I_L(-1,1))$ (see, e.g. Alonso and Câmara, 2016).

Proof of Proposition 2

Step 1: Show that a solution to Problem (1)

$$\max_{(M_L,M_R)\subseteq X^2} \int\limits_{M_L\cap M_R} d\mu_0(x) \quad \text{subject to}$$

$$\int\limits_{M_v} \alpha_v(x)d\mu_0(x) \geq 0 \quad \text{and} \quad \int\limits_{M_v^c} \alpha_v(x)d\mu_0(x) < 0, \quad \forall v \in \{L,R\}$$

is given by

1. $\overline{M}_L = [\lfloor A_L \rfloor, b_L]$ and $\overline{M}_R = [a_R, \lceil A_R \rceil]$ if $a_R \geq \lfloor A_L \rfloor$ and $b_L \leq \lceil A_R \rceil$;
2. $\overline{M}_L = I_L(a_R, 1)$ and $\overline{M}_R = [a_R, \lceil A_R \rceil]$ if $a_R < \lfloor A_L \rfloor$ and $b_L \leq \lceil A_R \rceil$;
3. $\overline{M}_L = [\lfloor A_L \rfloor, b_L]$ and $\overline{M}_R = I_R(-1, b_L)$ if $a_R \geq \lfloor A_L \rfloor$ and $b_L > \lceil A_R \rceil$.

With a slight abuse of notation, let (M_L, M_R) be a solution to Problem (1) and refer to $\mu_0(M_L \cap M_R)$ as the objective value. If a pair (M_L, M_R) is a solution, then the pair $(M_L \cup A_L, M_R \cup A_R)$ is also a solution as it reaches a weakly higher objective value while weakly loosening all constraints. Note that considering a solution (M_L, M_R) such that $A_v \subseteq M_v$ for each v allows us to ignore the less-than-zero constraints: they are automatically satisfied since $\alpha_v(x) < 0$ for all $x \notin A_v^c$. Similarly, if a pair (M_L, M_R) is a solution and M_L includes $Z \subseteq X$ which is not a subset of $M_L \cap M_R$ or A_L, then the pair $(M_L \setminus Z, M_R)$ obtains the same objective value while loosening L's constraint, and is also a solution. Therefore, we will focus on a solution (M_L, M_R) such that $M_v = (M_L \cap M_R) \cup A_v$ for each $v \in \{L, R\}$. The remaining constraints are $\int_{M_R} \alpha_R(x) d\mu_0(x) \geq 0$ (for R) and $\int_{M_L} \alpha_L(x) d\mu_0(x) \geq 0$ (for L). Next, we show that the set $(M_L \cap M_R) \cap [-1, 0]$ is almost surely an interval $[a, 0]$ for some $a \leq 0$.

Claim 1. *Let (M_L, M_R) be a solution to Problem (1) such that $M_v = (M_L \cap M_R) \cup A_v$ for each $v \in \{L, R\}$. Then, for any two sets $Y, Z \subset [-1, 0]$ such that $\lceil Y \rceil \leq \lfloor Z \rfloor$,*

$$Y \subseteq M_L \cap M_R \implies Z \subseteq M_L \cap M_R.$$

Proof. Suppose, by contradiction, that sets Y and Z are as described in the claim, yet $Y \subseteq M_L \cap M_R$ and $Z \nsubseteq M_L \cap M_R$. WLOG suppose that $\mu_0(Y) = \mu_0(Z) > 0$ and $Z \cap (M_L \cap M_R) = \varnothing$. First, observe that since $\lceil Y \rceil \leq \lfloor Z \rfloor$, $\mu_0(Y) = \mu_0(Z)$ and $\alpha_R(x)$ is strictly increasing for $x \leq 0$, we have

$$\int_Y \alpha_R(x) d\mu_0(x) < \alpha_R(\lceil Y \rceil)\mu_0(Y) \leq \alpha_R(\lfloor Z \rfloor)\mu_0(Z) < \int_Z \alpha_R(x) d\mu_0(x).$$

Similarly, noting that $\mu_0(Y \cap A_L^c) \geq \mu_0(Z \cap A_L^c)$ (since Y is left of Z and hence is more likely to be outside of $A_L = [\lfloor A_L \rfloor, 0]$) and $\alpha_L(x)$ is strictly increasing for $x \in [-1, 0] \cap A_L^c$, we have

$$\int_{Y \cap A_L^c} \alpha_L(x) d\mu_0(x) \leq \alpha_L(\lceil Y \rceil)\mu_0(Y \cap A_L^c) \leq \alpha_L(\lfloor Z \rfloor)\mu_0(Z \cap A_L^c) \leq \int_{Z \cap A_L^c} \alpha_L(x) d\mu_0(x).$$

Note that the inequality $\int_{Y \cap A_L^c} \alpha_L(x) d\mu_0(x) \leq \int_{Z \cap A_L^c} \alpha_L(x) d\mu_0(x)$ is strict unless $\mu_0(Y \cap A_L^c) = 0 \iff \mu_0(Y \cap A_L) = \mu_0(Y)$ (Y is almost surely a subset of A_L).

Next, let $\widetilde{M}_L := (M_L \setminus (Y \cap A_L^c)) \cup (Z \setminus A_L^c)$ and $\widetilde{M}_R := (M_R \setminus Y) \cup Z$. We will show that $(\widetilde{M}_L, \widetilde{M}_R)$ cannot be a solution to Problem (1) because the value of the objective can be strictly higher than $\mu_0(\widetilde{M}_L \cap \widetilde{M}_R)$. Since $\mu_0(\widetilde{M}_L \cap \widetilde{M}_R) = \mu_0(M_L \cap M_R)$, that would imply that (M_L, M_R) is not a solution, either.

From the inequalities we derived above, we have for R:

$$\int\limits_{M_R} \alpha_R(x) d\mu_0(x) \geq 0 \implies \int\limits_{\widetilde{M}_R} \alpha_R(x) d\mu_0(x) > 0,$$

meaning that R's constraint for $\widetilde{M}_R$ is loose. For L,

$$\int\limits_{M_L} \alpha_L(x) d\mu_0(x) \geq 0 \implies \int\limits_{\widetilde{M}_L} \alpha_L(x) d\mu_0(x) = \int\limits_{\widetilde{M}_L} \alpha_L(x) d\mu_0(x) \geq 0,$$

the last inequality being strict unless Y is a.s. a subset of A_L.

Now, observe that there exists $\widetilde{Y} \subseteq Y$ such that $\mu_0(\widetilde{Y}) > 0$ and $\widehat{W}_v := \widetilde{M}_v \cup \widetilde{Y}$ satisfies each v's constraint. Indeed, R's constraint is loose for $\widetilde{M}_R$. For L, her constraint is either loose or Y is a.s. a subset of A_L. Either way, we can find a positively-measured $\widetilde{Y} \subseteq Y$ so that $\widehat{W}_L$ satisfies L's constraint. The objective value for $(\widetilde{M}_L, \widehat{W}_R)$ is $\mu_0(\widehat{W}_L \cap \widehat{W}_R) = \mu_0(M_L \cap M_R) + \underbrace{\mu_0(\widetilde{Y})}_{>0}$, so (M_L, M_R) cannot be a solution, a contradiction. ■

Analogously to Claim 1, we can show that the set $(M_L \cap M_R) \cap [0, 1]$ is also almost surely an interval $[0, b]$ for some $b \geq 0$. Consequently, there exists a solution (M_L, M_R) to Problem (1) such that $M_L \cap M_R = [a, b]$ for some $a \leq 0$ and $b \geq 0$. Finally, we consider cases.

- $a_R \geq \lfloor A_L \rfloor$ and $b_L \leq \lceil A_R \rceil$. In this case, the proposed solution to Problem (1) is $\overline{M}_L = [\lfloor A_L \rfloor, b_L]$, $\overline{M}_R = [a_R, \lceil A_R \rceil]$ with $\overline{M}_L \cap \overline{M}_R = [a_R, b_L]$. By contradiction, suppose that the pair $(\overline{M}_L, \overline{M}_R)$ does not solve Problem (1). Then, there exists a solution (M_L, M_R) such that $M_L \cap M_R = [a, b]$ for some $a \leq 0$ and $b \geq 0$. Furthermore, $\mu_0([a, b]) > \mu_0([a_R, b_L])$, implying that $a < a_R$ or $b > b_L$.
  Now, if $a < a_R$, then $M_R = [a, \lceil A_R \rceil]$ satisfies R's constraint and $\mu_0([a, \lceil A_R \rceil]) > \mu_0([a_R, \lceil A_R \rceil])$, which contradicts the definition of $[a_R, \lceil A_R \rceil] =: I_R(-1, \lceil A_R \rceil)$ being a solution to Problem (AUX) for R with $l = -1$ and $r = \lceil A_R \rceil$. Similarly, if $b > b_L$, we obtain a contradiction to the definition of $[\lfloor A_L \rfloor, b_L] := I_L(\lfloor A_L \rfloor, 1)$ being

a solution to Problem (AUX) for L and $l = \lfloor A_L \rfloor$ and $r = 1$. Hence, $(\overline{M}_L, \overline{M}_R)$ is a solution to Problem (1) for the considered values of a_R and b_L.

- $a_R < \lfloor A_L \rfloor$ and $b_L \leq \lceil A_R \rceil$. In this case, the proposed solution is $\overline{M}_L = I_L(a_R, 1)$, $\overline{M}_R = [a_R, \lceil A_R \rceil] = I_R(-1, \lceil A_R \rceil)$ with $\overline{M}_L \cap \overline{M}_R = \overline{M}_L$. In particular, by the definition of $I_L(a_R, 1)$,

$$\overline{M}_L \cap \overline{M}_R = \arg \max_{W \subseteq [a_R, 1]} \int_W d\mu_0(x) \quad \text{subject to} \int_W \alpha_L(x) d\mu_0(x) \geq 0.$$

Therefore, if the pair $(\overline{M}_L, \overline{M}_R)$ does not solve Problem (1), then there exists a solution (M_L, M_R) such that $M_L \cap M_R = [a, b]$ and $a < a_R$, $b \geq 0$. However, we previously showed that $a < a_R$ contradicts the definition of $[a_R, \lceil A_R \rceil]$ being a solution to Problem (AUX) for R with $l = -1$ and $r = \lceil A_R \rceil$.

- $a_R \geq \lfloor A_L \rfloor$ and $b_L > \lceil A_R \rceil$. This case is analogous to the one above.

This completes Step 1.

<u>**Step 2**</u>: Describe the direct equilibrium characterized by $(\overline{M}_L, \overline{M}_R)$.

1. The challenger's strategy is $\sigma : X \to \Delta\left(\{\overline{M}_L, \overline{M}_L^c\} \times \{\overline{M}_R, \overline{M}_R^c\}\right)$ is to send the collection of messages (m_L, m_R) with probability one depending on x, where (m_L, m_R) equals $(\overline{M}_L, \overline{M}_R)$ if $x \in \overline{M}_L \cap \overline{M}_R$; $(\overline{M}_L, \overline{M}_R^c)$ if $x \in \overline{M}_L \cap \overline{M}_R^c$; $(\overline{M}_L^c, \overline{M}_R)$ if $x \in \overline{M}_L^c \cap \overline{M}_R$; $(\overline{M}_L^c, \overline{M}_R^c)$ if $x \in \overline{M}_L^c \cap \overline{M}_R^c$.
2. On the path, v's posterior is $\mu_0(\cdot \mid m)$ for $m \in \{\overline{M}_v, \overline{M}_v^c\}$ (calculated using the Bayes rule). Off the path, v's posterior following message $m \notin \{\overline{M}_v, \overline{M}_v^c\}$ is supported on $A_v^c \cap m$ whenever that set is non-empty.
3. On the path, v approves after message $\overline{M}_v$ and rejects after $\overline{M}_v^c$ because these sets satisfy the constraints of Problem (1). Off the path, voter v approves after message $m \notin \{\overline{M}_v, \overline{M}_v^c\}$ only if $m \subseteq A_v$ due to the assumed skeptical beliefs.

Observe that the challenger has no profitable deviations. Indeed, if $x \in \overline{M}_L \cap \overline{M}_R$, then he wins the election and receives the highest possible payoff. If $x \notin \overline{M}_L \cap \overline{M}_R$, then $x \notin \overline{M}_v \implies x \notin A_v$ for some $v \in \{L, R\}$. Any deviation by the challenger with such policy outcome would lead the skeptical voter v to reject. This completes the equilibrium characterization.

<u>**Step 3**</u>: Establish that the upper bound on the challenger's odds of winning across all

communication protocols is $\mu_0(\overline{M}_L \cap \overline{M}_R)$.

The upper bound on the challenger's ex-ante utility is reached in a setting wherein he has ex-ante commitment power (Kamenica and Gentzkow, 2011). Intuitively, the equilibrium definition of the game considered in this paper requires the challenger to maximize his expected utility for each $x \in X$, whereas there is no such restriction when he has commitment power. Next, let us solve the information design problem.[26] First, the challenger chooses and commits to an experiment, which is a measurable map $\psi : X \to \Delta\{0,1\}^2$. Next, the challenger's policy outcome x is realized according to μ_0 and the signals $s_L \in \{0,1\}$ and $s_R \in \{0,1\}$ are sent to voters with bliss points L and R (resp.) with probability $\psi((s_L, s_R) \mid x)$. Then, voter $v \in \{L, R\}$ privately observes her signal s_v, forms a posterior belief $\mu_v(\cdot \mid s_v) \in \Delta X$ using the Bayes rule and approves after $s_v = 1$ and rejects after $s_v = 0$. Let $\psi_v(s_v \mid x) := \sum_{s_{-v} \in \{0,1\}} \psi((s_v, s_{-v}) \mid x)$ be the marginal probability that v receives signal s_v. For v to approve after signal $s_v = 1$, her net payoff from approval must be non-negative:

$$\int \alpha_v(x) d\mu_v(x \mid 1) \geq 0 \iff \int \alpha_v(x)\psi_v(1 \mid x) d\mu_0(x) \geq 0.$$

Similarly, for v to reject after signal $s_v = 0$, her expected net payoff from approval must be negative, $\int \alpha_v(x)\psi_v(0 \mid x) d\mu_0(x) < 0$. We look for an optimal experiment that maximizes the challenger's odds of winning and solves

$$\max_{\psi} \int \psi((1,1) \mid x) d\mu_0(x) \quad \text{subject to}$$
$$\int \alpha_v(x)\psi_v(1 \mid x) d\mu_0(x) \geq 0 \quad \text{and} \quad \int \alpha_v(x)\psi_v(0 \mid x) d\mu_0(x) < 0, \quad \forall v \in \{L, R\}.$$

Now, this is a linear problem with linear constraints so the solution is an extreme point of the constraint set. Since X is rich and μ_0 is atomless, the extreme points are deterministic. For the optimal deterministic experiment $\psi^* : X \to \{0,1\}^2$, let $M_v := \{x \in X \mid \psi_v^*(1 \mid x) = 1\}$ for each $v \in \{L, R\}$ be the set of policy outcomes that v is recommended to approve. Then, ψ^* now solves Problem (1).

[26] In what follows, we employ the revelation principle Bergemann and Morris (2016) that allows us to restrict attention to action recommendations that are obeyed.

Proof of Lemma 2

Recall from the definition of I_v and Lemma 1 that for any right voter $v > 0$, $\int_{a_v}^{\lceil A_v \rceil} \alpha_v(x) d\mu_0(x) = 0$ unless $a_v = -1$. We prove this lemma for two right voters v and w such that $0 < v < w \leq 1$. The case of left voters is symmetric.

First, observe that $\overline{x}_v := \int_{a_v}^{\lceil A_v \rceil} x d\mu_0(x) > 0$. Indeed, by Jensen's inequality for the strictly concave function α_v, we have $\alpha_v(\overline{x}_v) > \int_{a_v}^{\lceil A_v \rceil} \alpha_v(x) d\mu_0(x) \geq 0$ and $\alpha_v(\overline{x}_v) > 0 \iff \overline{x}_v \in (0, \lceil A_v \rceil)$.

Next, if we evaluate w's obedience constraint for $[a_v, \lceil A_w \rceil]$, we get

$$\int_{a_v}^{\lceil A_w \rceil} \alpha_w(x) d\mu_0(x) = \underbrace{\int_{a_v}^{\lceil A_v \rceil} \alpha_v(x) d\mu_0(x)}_{\geq 0 (\text{ } v\text{'s obedience})} + \underbrace{2(w - v)\overline{x}_v}_{>0} + \int_{\lceil A_v \rceil}^{\lceil A_w \rceil} \underbrace{\alpha_w(x)}_{\geq 0 \text{ for all } x \in A_w} d\mu_0(x) > 0,$$

which means that w's obedience constraint is loose for $[a_v, \lceil A_w \rceil]$. Now, several cases are possible. First, if $\int_{-1}^{\lceil A_v \rceil} \alpha_v(x) d\mu_0(x) \geq 0$, then $a_v = a_w = -1$. Second, if $\int_{-1}^{\lceil A_v \rceil} \alpha_v(x) d\mu_0(x) < 0$ and $\int_{-1}^{\lceil A_w \rceil} \alpha_w(x) d\mu_0(x) \geq 0$, then $a_w = -1 < a_v$. Finally, if $\int_{-1}^{\lceil A_v \rceil} \alpha_v(x) d\mu_0(x) < 0$ and $\int_{-1}^{\lceil A_w \rceil} \alpha_w(x) d\mu_0(x) < 0$, then a_w solves $\int_{a_w}^{\lceil A_w \rceil} \alpha_w(x) d\mu_0(x) = \int_{a_v}^{\lceil A_w \rceil} \alpha_w(x) d\mu_0(x) + \int_{a_w}^{a_v} \underbrace{\alpha_w(x)}_{<0 \text{ for all } x<0} d\mu_0(x) = 0$, which is possible if and only if $a_w < a_v$.

Proof of Proposition 3

Let $\{L, 0, E\}$ be the baseline electorate with the more extreme right voter $E > R$. Denote by $(\widetilde{M}_L, \widetilde{M}_E)$ the challenger-preferred equilibrium sets of policy outcomes approved by left and right voters, respectively. Note that $b_L \leq \lceil A_R \rceil$ and $R < E$ imply that $b_L \leq \lceil A_E \rceil$, so $(\overline{M}_L, \overline{M}_R)$ and $(\widetilde{M}_L, \widetilde{M}_E)$ are both described by Cases 1 or 2 or Proposition 2. Therefore, $\overline{M}_R = [a_R, \lceil A_R \rceil] = I_R$ and $\widetilde{M}_E = [a_E, \lceil A_E \rceil] = I_E$ and, by Lemma 2, $a_E \leq a_R$. Three cases are possible:

1. $\lfloor A_L \rfloor \leq a_E < a_R$. Then, $\overline{M}_L \cap \overline{M}_R = [a_R, b_L]$ and $\widetilde{M}_L \cap \widetilde{M}_E = [a_E, b_L]$. The claim of the proposition holds because $a_E < a_R$;
2. $a_E < \lfloor A_L \rfloor \leq a_R$. Then, $\overline{M}_L \cap \overline{M}_R = [a_R, b_L] \subset I_L(\lfloor A_L \rfloor, 1)$ and $\widetilde{M}_L \cap \widetilde{M}_E = I_L(a_E, 1)$. Clearly, $I_L(\lfloor A_L \rfloor, 1)$ is left of $I_L(a_E, 1)$ and has lower prior measure as both are solutions to Problem (AUX) with $l = \lfloor A_L \rfloor$, $r = 1$ and $l = a_R$, $r = 1$ (resp.) and the latter parametrization allows for policy outcomes left of $\lfloor A_L \rfloor$.
3. $a_E \leq a_R < \lfloor A_L \rfloor$. Then, $\overline{M}_L \cap \overline{M}_R = I_L(a_R, 1)$, $\widetilde{M}_L \cap \widetilde{M}_E = I_L(a_E, 1)$, and the former set is left of the latter set and has a lower prior measure for the same reason as in Case 2.