Targeted Advertising in Elections*

Maria Titova†

August 12, 2019

Abstract

Can targeted messages swing elections? This paper examines a model of optimal information disclosure in which the sender is a candidate who challenges the incumbent, and the receiver is the heterogeneous population of voters. Each voter has an ideal position on a finite number of issues and expressively votes for the candidate who is located closest to it. The challenger's location in the policy space is his private information, and he strategically communicates it to the voters by sending either a public verifiable message, or a targeted verifiable message to each voter based on that voter's ideal position on every issue. I show that the challenger can win the election under public disclosure if and only if a majority of voters agrees on at least one issue. In polarized societies, in which voters disagree on all issues, public disclosure is not effective, but targeting is. I show that the challengers who agree with the voters on at least one issue can still win the election by drawing each voter's attention to the issues that they agree on. The more polarized the society, the more types of challengers can swing elections by targeting.

*I am grateful to my advisors Renee Bowen and Joel Sobel for their continuous guidance and support. I thank Prashant Bharadwaj, Simone Galperti, Remy Levin, Alexander Levkun, Mark Machina, Tara Sullivan, Joel Watson, Alexander Weiss, Michelle White, and the audiences of the 2019 meetings of the Southwest Economic Theory Conference and Young Economists Symposium as well as seminar audiences at UC San Diego. All errors are my own.

†Department of Economics, University of California, San Diego. E-mail: mtitova@ucsd.edu.

1 Introduction

During the 2016 U.S. presidential election, the Democratic and Republican candidates employed vastly different campaign strategies. While Hillary Clinton advertised on traditional media, Donald Trump placed a greater emphasis on targeted online advertising.[1] Specifically, the Trump campaign's strategies on digital platforms like Facebook and Twitter targeted voters based on their positions on issues of LGBT rights, immigration, race, tax policy, gun control, abortion rights, and others, with very high precision. Given the close electoral outcome and divergent advertising practices, one may ask: was the sophisticated targeted advertising campaign the deciding factor? In other words, can targeting swing elections?

To answer this question, I consider a model of optimal information disclosure by a challenging candidate who has private information about his position on issues that the electorate cares about and can send targeted messages to individual voters about it. I employ a multidimensional spatial model of electoral preference based on the canonical models of Hotelling (1929) and Downs (1957), in which a dimension corresponds to an issue, and there are finitely many positions on every issue. Each voter has an ideal position (on each issue) and votes for the candidate (the challenger or the incumbent) who is located closest to it. Voting is expressive: each voter goes to the polls to express her support for a policy position without taking into account that her vote may not be pivotal. The winner of the election is decided by majority rule. Messages are allowed to be quite general, but must satisfy the "grain of truth" restriction: a candidate's message does not have to disclose his position precisely (for instance, he can omit talking about some issues completely, and on others, he can provide a range of positions), but any message he sends must include his true position.[2] In other words, lies of omission are allowed, while lies of commission are not.

To see if targeting is helpful for changing the outcome of an election, I first study the case of public disclosure, in which the challenger is restricted to sending the same public message to each voter (for example, he participates in a public debate). I then proceed to characterize elections that

[1] Williams and Gulati (2018), using the candidate expenditure files collected by the FEC, report that during his presidential campaign, Donald Trump spent US$83.5 million on digital media and US$ 90 million on traditional media, while Hillary Clinton spent US$20.2 million and US$233 million, respectively.

[2] The grain of truth assumption is central in verifiable information disclosure literature after Milgrom (1981), Grossman (1981), and Milgrom and Roberts (1986).

cannot be won by the challenger under public disclosure but become winnable when he is allowed to target individual voters based on their ideal positions.

I show that under public disclosure, the election is decided by a group of (compatible) voters, all of whom agree on at least one issue, which means that all of their positions on this issue are *to the right* or *to the left* of the incumbent. The first main result shows equivalence between *(i)* no position of the challenger being preferred to the incumbent under complete information, *(ii)* no group of compatible voters constituting a majority, and *(iii)* the election being unwinnable under public disclosure. The second main result is that the election becomes winnable under public disclosure when some group of compatible voters does constitute a majority. Since compatible voters agree on some issue, say, to the right, then the position of the challenger that is immediately to the right of the incumbent on this issue, and coincides with the incumbent on all other issues, is preferred to the incumbent by all these voters. Then, this challenger is a "winner" of this election, because him truthfully revealing his position leads to a majority of voters voting for him. The further the mass of compatible voters that constitutes a majority is located from the incumbent, in other words, the more to the right all of their positions are on the issue in agreement, the more "winners" there are. I show that regardless of the prior distribution of beliefs about the challenger's location, there exists an equilibrium in which these "winners" pool with types adjacent to them on any single issue, but not further than that. This equilibrium maximizes the challenger's ex-ante probability of winning the election in the class of ex-post efficient equilibria, in which no voter regrets her choice with probability one. In this sense, the results under public disclosure mimic Milgrom and Roberts (1986)'s famous *unraveling* feature: (almost) only the challengers who can win by fully revealing their type win under public verifiable information disclosure.

I show that targeting is effective in polarized societies, in which no coalition of compatible voters constitutes a majority, so the challenger cannot win under public disclosure. More specifically, I examine elections that are decided by two voters who disagree with each other (which means that they are positioned on the opposite sides of the incumbent) on *every* issue. I show that challenger types who, for each voter, *(i)* agree with this voter on *some* issue and *(ii)* are overall "not too far" from this voter, can win the election under targeted advertising. The "winning" message sent to each voter includes an interval of positions, ranging from this voter's bliss point to the challenger's true position, on every issue. These tailored messages persuade both voters at the same time, and this challenger wins the election. The challengers who are "not too far" from both voters satisfy the ex-post efficiency condition, which means that they are weakly preferred to the incumbent by both voters. As a result, targeting does not lead to a welfare loss. Other desirable features of the constructed equilibrium are: the strategy profiles are robust to changes in prior beliefs, all messages on the path are convex, and it maximizes the challenger's ex-ante utility among all ex-post efficient equilibria. In any polarized election that is decided by two voters, there always exists at least one type of challenger who can swing this election by using targeted advertising. The more polarized the voters are (the larger the range of their positions on issues and the higher the total number of issues), the more challengers agree with each voter on some issue, and hence the more challengers can swing the election by targeting.

Related Literature

This paper contributes to the literature on *voter persuasion*. Jackson and Tan (2013) consider a setup in which a set of voters, before choosing between two alternatives, consult with privately informed experts. They show that the fact that a greater amount of information is disclosed does not necessarily imply that the total utility of the electorate is maximized, and that the optimal choice of the voting rule, in terms of either information revelation or the ex-ante utility of voters, depends on the setup. Schnakenberg (2015) considers a similar setup, but with a multidimensional policy space, and shows that the privately informed expert can make all voters worse off, if his preferred way of communication is cheap talk.[3] In Alonso and Câmara (2016), there are no experts, and the politician himself is trying to persuade the voters to choose him over the incumbent in a Bayesian manner.[4] They show that, under majority rule, all voters are always weakly (and sometimes strictly) worse off. In my paper, I examine existence of ex-post efficient equilibria, in which voters are never strictly worse off, in a multidimensional policy space, under verifiable information disclosure, when the sender can send a private signal based on voter type, as opposed to a public signal to everyone.

[3]The *cheap talk* literature, started by Crawford and Sobel (1982), assumes that the sender can costlessly send any message, regardless of what he knows.

[4]The literature on *Bayesian persuasion*, launched by Kamenica and Gentzkow (2011), makes an ex-ante commitment assumption: the sender does not know the true state, but he can commit a policy experiment that will reveal some information about the true state, the results of which he will truthfully disclose to the receiver.

This paper also contributes to the literature on *discriminatory information disclosure*, in which the sender caters his messages to specific groups of receivers. Farrell and Gibbons (1989) and Goltsman and Pavlov (2011) consider a cheap talk setting, and show that public disclosure may lead to more informative equilibria than private communication when receivers are biased in opposite directions, because in private the sender has a significant incentive to overstate the truth, which prevents each receiver from believing him. In Koessler (2008)'s setting, the sender can send certifiable messages, but again, private communication forces him to reveal more information, and as a consequence, he is better off under public communication. Bar-Isaac and Deb (2014) confirm that public communication is better for the sender if his purpose is to build reputation. Wang (2013) comes to the same conclusion after examining a Bayesian persuasion setting with heterogeneous voters. Overall, this literature implies that if the voters are strategic (that is, if they either cared explicitly about the outcome, or took into account the probability of being pivotal), then they are extremely skeptical when receiving a private message. This, in turn, implies that the sender is always worse off under private disclosure, relative to public disclosure. This paper examines voters who are not strategic, but rather they vote expressively to show support for their favorite position. As a consequence, the challenger is always weakly (and sometimes strictly) better off under private disclosure.

This paper contributes to the literature on *disclosure in electoral competition*. When candidates send verifiable messages about their policy, competition usually results in full unraveling of information. The reason is that the candidates play a zero-sum game, and it pushes them to voluntarily disclose all information, however unfavorable, as shown by Board (2009). Expanding the candidates' options beyond disclosing information about own position does not help, either. Schipper and Woo (2017), like me, examine the effects of allowing candidates to microtarget voters (who may be unaware that some issues even exist), and, separately, allow candidates to "negatively campaign" by sending verifiable messages about each other. In both cases, they find that the voters always end up learning everything and making the same choices they would have made under complete information. Janssen and Teteryatnikova (2017) come to the same conclusion after studying a model of electoral competition in which the candidates are allowed to send (comparative) statements about their position relative to their opponent. In my model, one candidate has a significant advantage over his opponent in that he is the only one who can send (targeted) messages to voters, perhaps because his budget is much higher, or because he has access to better technology. I show that even under such extreme asymmetry, the outcome of the election would still be not too far from the outcome under complete information and that targeting is only useful in polarized societies.

Finally, I provide a theoretical explanation for *empirical literature on voter persuasion*, surveyed by Prat and Strömberg (2013) and DellaVigna and Gentzkow (2010), which documents how catering campaign messages to specific groups of voters affects electoral outcomes and voter turnout. Applicable examples include George and Waldfogel (2006), who argue that candidates who target educated voters who care more about global issues post their ads in the *New York Times*, while candidates with more local agenda post in local newspapers; DellaVigna and Kaplan (2007), who show that the introduction of Fox News in the local markets in the US increased the Republican vote in the 2000 presidential election; Enikolopov, Petrova, and Zhuravskaya (2011), who study the 1999 parliamentary election in Russia and show that introduction of a government-independent TV channel increased the vote for major opposition parties; Oberholzer-Gee and Waldfogel (2009), who document that an increase in availability of Spanish-language news positively affected voter turnout in the US in 1994-2002, suggesting that the reason is that these news outlets are much more likely to report on issues that are of interest to Hispanics. In all these examples, candidates "target" very large groups of voters (college-educated people, conservatives, those unsatisfied with the policy of the ruling party in Russia, Spanish speakers in the US), while this paper focuses on targeted messages of extreme precision based on voter's preferred position on *all* issues. A more relevant example of the type of targeting I study would be the 2016 US presidential election, in which, according to Kim et al. (2018), one of the candidates employed "targeting very specific categories of individuals" using social networks.[5]

In my model, the challenging candidate sends targeted messages as part of his electoral campaign. Alternative setups include citizen-candidate models put forward by Osborne and Slivinski (1996) and Besley and Coate (1997), and competition of media for voter attention, most recently considered by Perego and Yuksel (2018). Examining the role of targeted advertising in these frameworks is left for future research.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents the analysis of the cases of public disclosure and targeted advertising. Section 4 discusses assump-

[5] On September 6, 2017, Facebook CFO Alex Stamos issued a press release titled *An Update On Information Operations On Facebook* stating that "ads appeared to focus on amplifying divisive social and political messages across the ideological spectrum – touching on topics from LGBT matters to race issues to immigration to gun rights".

tions and concludes.

2 Model

This section introduces an *election game* between a challenging candidate and a population of voters in an environment with asymmetric information. The challenger can send verifiable messages to persuade the voters to choose him over the incumbent, and each voter prefers to vote for the candidate whose position is closest to her bliss point.

The policy space is a finite set $\Theta \equiv \times_{k=1}^{K}\{1,\dots,N_k\}$, where $\mathbb{K} \equiv \{1,\dots,K\}$ is the set of issues and N_k is the number of positions on issue $k \in \mathbb{K}$. Each element $\theta = (\theta_1,\dots,\theta_K) \in \Theta$ of the policy space represents a (policy) position that reflects a location on each issue.

The candidates are characterized by their locations in the policy space. The incumbent's position $\theta^{inc} \in \Theta$ is known to everyone. The challenger's position $\theta^{ch} \in \Theta$ (which I will call his *type*) is only known to the challenger, and everyone else shares a common prior $p(\cdot)$ over Θ with full support. Locations of the candidates are given exogenously, in the sense that neither candidate chooses his location strategically: the incumbent's position is fixed and known, while the challenger's position is drawn by nature.

Example. (The 2×2 policy space)

For ease of exposition, I will illustrate my results on a 2×2 policy space $\Theta = \{1,2\} \times \{1,2\}$. That is, there are two issues (for instance, gay marriage and gun control) and two positions on each issue (for instance, "pro" and "against"). The policy space is pictured in Figure 1. The black star denotes the location of the incumbent. The common prior belief that the challenger's location is θ_{ij} is p_{ij}.

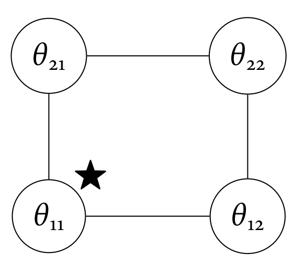


Figure 1: *The 2×2 policy space*

△

There is a unit mass of voters. Each voter can be identified by her location in the policy space, which corresponds to her ideal position on every issue: voter v_θ has ideal position $\theta \in \Theta$. The distribution of voter types over the policy space Θ is given by $g^v(\cdot)$.

I denote the distance between any two positions $\theta, \theta' \in \Theta$ as $d : \Theta \times \Theta \to \mathbb{R}$. I will employ the L_1 metric, which states that the distance between two location points is the sum of the differences between the positions on every issue, over every issue:

$$d(\theta, \theta') = \sum_{k=1}^{K} |\theta_k - \theta'_k|.$$

Voters vote expressively and prefer policies that are closest to their bliss points.[6] To be precise, the utility of voter v_θ under complete information is given by

$$u_\theta(a, \theta^{ch}) = -a \cdot d(\theta, \theta^{ch}) - (1 - a) \cdot d(\theta, \theta^{inc}),$$

where $a = 1$ corresponds to voting for the challenger and $a = 0$ corresponds to voting for the incumbent. Voting is costless and *mandatory*, i.e. each voter chooses an action from a common action space $\mathbb{A} = \{0, 1\}$, and abstaining is not an option. A simple feature of this utility function is that under complete information the voter's utility is maximized when she votes for the challenger if $d(\theta, \theta^{ch}) < d(\theta, \theta^{inc})$, and for the incumbent if $d(\theta, \theta^{ch}) > d(\theta, \theta^{inc})$. I assume that when the voter is indifferent, i.e. when $u_\theta(1, \theta^{ch}) = u_\theta(0, \theta^{ch}) \iff d(\theta, \theta^{ch}) = d(\theta, \theta^{inc})$, ties are broken in favor of the incumbent.

When there is uncertainty about challenger's position and voter v_θ has belief $\mu_\theta(\cdot)$ over Θ, her *expected* utility is

$$\begin{aligned}\mathbb{E}_{\mu_\theta}\Big[u_\theta(a, \theta^{ch})\Big] &= \sum_{\theta^{ch} \in \Theta} \mu_\theta(\theta^{ch})\Big[- a \cdot d(\theta, \theta^{ch}) - (1 - a) \cdot d(\theta, \theta^{inc})\Big] \\ &= -a \cdot \mathbb{E}_{\mu_\theta}\Big[d(\theta, \theta^{ch})\Big] - (1 - a) \cdot d(\theta, \theta^{inc}).\end{aligned}$$

[6] The standard assumption behind expressive voting is that voters derive utility from expressing support for one of the electoral options, and not from the policy that is implemented by the elected candidate. The theory was pioneered by Brennan (1998) and surveyed by Hamlin and Jennings (2011). For costly expressive voting, see Clark and Lee (2016).

This means that under incomplete information, voter v_θ's expected utility is maximized when she votes for the challenger if $\mathbb{E}_{\mu_\theta}\left[d(\theta, \theta^{ch})\right] < d(\theta, \theta^{inc})$ and for the incumbent otherwise. As under complete information, the ties are broken in favor of the incumbent for every $\mu_\theta(\cdot)$.

The challenger aims to win the majority of the votes. That is, given a collection of voters' actions $\{a_\theta\}_{\theta\in\Theta} \in \mathbb{A}^{|\Theta|}$ (which, for every voter v_θ, specifies the action a_θ she takes), his utility is

$$U^{ch}(\{a_\theta\}) = \mathbb{1}\left\{\sum_{\theta\in\Theta} g^v(\theta)\cdot a_\theta > \frac{1}{2}\right\},$$

where $\mathbb{1}\{\cdot\}$ is the indicator function, which equals one if the argument condition is satisfied and zero otherwise. The challenger's payoff is therefore binary: he gets a utility of one if he wins the election (by securing a majority of votes) and zero if he loses.

Example. (Elections in the 2×2 policy space)

The mass of voters whose favorite policy is θ_{ij} is given by g^v_{ij}. When there is no strategic communication (under the prior belief $p(\cdot)$), the election goes as follows:

- each voter calculates her expected distance form the challenger, e.g. for v_{21}:

$$\mathbb{E}_p\left[d(\theta_{21}, \theta^{ch})\right] = 0\cdot p_{21} + 1\cdot p_{22} + 1\cdot p_{11} + 2\cdot p_{12};$$

- each voter votes for the candidate who is located closest to her ideal policy, e.g. for v_{21}:

$$a_{21} = \begin{cases} 1, \text{ if } \mathbb{E}_p\left[d(\theta_{21}, \theta^{ch})\right] < d(\theta_{21}, \bigstar), \\ 0, \text{ otherwise;} \end{cases}$$

- the winner of the election is determined. It is the challenger if

$$\sum_{(i,j)\in\Theta} g^v_{ij}\cdot a_{ij} > \frac{1}{2},$$

  and the incumbent otherwise. $\triangle$

The challenger can send a costless verifiable message from the message space $\mathbb{M} = 2^\Theta \setminus \varnothing$ about his location in the policy space. For instance, message $m = \{\theta^{ch}\}$ is completely informative to the voters because it reveals the true location of the challenger, while $m = \Theta$ is completely uninformative, since it does not reveal anything, and the challenger could be located anywhere. The verifiability of a message, also known as the *grain of truth* assumption, restricts type θ^{ch} of the challenger to only send messages $m \in \mathbb{M}$ that include his true type, i.e. $\theta^{ch} \in m$. Notice that the challenger's utility does not explicitly depend on either his type θ^{ch} or on the message he sends. Messages serve as a persuasion tool to manipulate voters' perception of the challenger's position in favor of the challenger.

I consider equilibria that are sequentially rational (Kreps and Wilson (1982)). The challenger sends a collection of messages $\{m_\theta\}_{\theta\in\Theta} \in \mathbb{M}^{|\Theta|}$ to the voters, where message m_θ is intended for voter v_θ, for every $\theta \in \Theta$. The challenger's *behavior strategy* is a function $\sigma : \Theta \times \mathbb{M}^{|\Theta|} \to [0,1]$, such that $\sum_{\{m_\theta\}\in\mathbb{M}^{|\Theta|}} \sigma(\theta^{ch}, \{m_\theta\}) = 1$ for every $\theta^{ch} \in \Theta$. That is, $\sigma(\theta^{ch}, \{m_\theta\})$ is the probability that type θ^{ch} sends the collection of messages $\{m_\theta\}_{\theta\in\Theta}$. For every distribution of message collections σ, one can calculate $\sigma_\theta : \Theta \times \mathbb{M} \to [0,1]$, the marginal distribution of messages heard by each individual voter v_θ, as

$$\sigma_\theta(\theta^{ch}, m_\theta) = \sum_{\{m_{-\theta}\}\in\mathbb{M}^{|\Theta|-1}} \sigma\Big(\theta^{ch}, \Big\{m_\theta, \{m_{-\theta}\}\Big\}\Big),$$

where $\{m_{-\theta}\}$ is a collection of messages intended for every voter, except for v_θ. Upon hearing message m_θ, voter v_θ forms a posterior belief $\mu_\theta(\cdot, m_\theta) \in \Delta\Theta$ and votes according to her voting rule $a_\theta : \mathbb{M} \to \mathbb{A}$. I consider equilibria in which voters do not use mixed strategies and vote for the incumbent when indifferent.

Definition 2.1. *Given the policy space Θ, the incumbent location θ^{inc}, the distribution of voters $g^v(\cdot)$, and the distribution of priors $p(\cdot)$, the equilibrium $\Big(\sigma, \boldsymbol{a}, \boldsymbol{\mu}\Big)$ of the election game consists of the signaling rule $\sigma(\theta^{ch}, \cdot)$ for the challenger, a collection of voting rules $\boldsymbol{a}(\{m_\theta\}) \equiv \{a_\theta(m_\theta)\}_{\theta\in\Theta}$ and posterior beliefs $\boldsymbol{\mu}(\{m_\theta\}) \equiv \{\mu_\theta(\cdot, m_\theta)\}_{\theta\in\Theta}$ for the voters, such that*

(i) $\forall \theta^{ch} \in \Theta$, $\sigma(\theta^{ch}, \{m_\theta\}) > 0$ *implies*

$$\{m_\theta\} = \arg \max_{\{m'_\theta\}\in\mathbb{M}^{|\Theta|}} U^{ch}\Big(\boldsymbol{a}(\{m'_\theta\})\Big), \text{ subject to } \theta^{ch} \in m'_\theta.$$

The following conditions must hold for every voter v_θ:

(ii) $\forall m \in \mathbb{M}$,

$$a_\theta(m) = \arg\max_{a\in\mathbb{A}} \sum_{\theta^{ch}\in\Theta} u_\theta(a, \theta^{ch}) \cdot \mu_\theta(\theta^{ch}, m);$$

(iii) $\forall m \in \mathbb{M}$ *such that* $\sum_{\theta^{ch}\in\Theta} \sigma_\theta(\theta^{ch}, m) > 0$,

$$\mu_\theta(\theta^{ch}, m) = \frac{\sigma_\theta(\theta^{ch}, m) \cdot p(\theta^{ch})}{\sum_{\theta'\in\Theta} \sigma_\theta(\theta', m) \cdot p(\theta')}.$$

In words, *(i)* states that every message collection $\{m_\theta\}$ sent by θ^{ch} with positive probability in equilibrium maximizes the challenger's utility given the collection of voters' voting rules; *(ii)* states that each voter is sequentially rational, and always maximizes her expected utility given her belief $\mu_\theta(\cdot, m)$; *(iii)* asserts that voter's belief $\mu_\theta(\cdot, m)$ following challenger's message m is Bayes-rational when message m is heard in equilibrium with positive probability. Since the "grain of truth" assumption effectively restricts the challenger's message space, it trivially holds that $\forall m \in \mathbb{M}$, $\mu_\theta(\theta^{ch}, m) = 0$ for all $\theta^{ch} \notin m$.

3 Analysis

To see whether targeting makes the challenger better off relative to public disclosure, I first introduce some definitions that will be useful in characterizing the set of equilibria for different elections. An election is defined by *(1)* what the incumbent position is, *(2)* what the distribution of voters' political beliefs is, and *(3)* what the prior belief of the voters regarding the challenger's position is.

Definition 3.1. *An* election $\mathcal{E}$ *is a triple* $\Big(\theta^{inc}, g^v(\cdot), p(\cdot)\Big)$.

I will select among equilibria on the basis of the challenger's ex-ante utility, which coincides with his *ex-ante probability of winning* the election. Given election $\mathcal{E}$ and equilibrium $\Big(\sigma, \boldsymbol{a}, \boldsymbol{\mu}\Big)$, it is given by

$$\hat{Pr}^{ch}\Big(\sigma, \boldsymbol{a}, \boldsymbol{\mu}\Big) = \sum_{\theta^{ch}\in\Theta} p(\theta^{ch}) \cdot \sum_{\{m_\theta\}\in\mathbb{M}^{|\Theta|}} \sigma\Big(\theta^{ch}, \{m_\theta\}\Big) \cdot U^{ch}\Big(\boldsymbol{a}(\{m_\theta\})\Big). \tag{1}$$

Definition 3.2. *Election* $\mathcal{E}$ *is*

- *unwinnable if* $\hat{Pr}^{ch}(\cdot) = 0$ *in every equilibrium;*
- *winnable if there exists an equilibrium in which* $\hat{Pr}^{ch}(\cdot) > 0$.

Example. (Unwinnable elections in the 2×2 policy space)

Suppose $g_{11} > 1/2$. Recall that $\bigstar = \theta_{11}$. Then, for voter v_{11}

$$d(\theta_{11}, \bigstar) = 0 \leq \mathbb{E}_{\mu}\Big[d(\theta_{11}, \theta^{ch})\Big]$$

for any belief μ. In other words, voter of type v_{11} always votes for the incumbent, because the incumbent is located at this voter's favorite position. Since the mass of these voters constitutes a majority, the incumbent always wins, and this election is unwinnable for the challenger. △

This example illustrates elections that are trivially unwinnable for the challenger, simply because the bliss point of the majority of voters coincides with the incumbent's position. I will later show that the elections may be unwinnable for less obvious reasons. For instance, when the challenger is restricted to sending public messages and the population of voters is sufficiently polarized, the election is unwinnable because a public message that attracts some voters offends the others.

I will say the equilibrium is *challenger-preferred* if the challenger's ex-ante probability of winning in this election is maximized over all equilibria for this election.

Ex-post efficient equilibria

Under complete information, if the challenger's position is fully revealed, the voters are able to make choices that are not only ex-ante, but also ex-post, efficient. Under incomplete information, I define ex-post efficiency as voters not making the "wrong" choices with probability one:

Definition 3.3. *In election $\mathcal{E}$, equilibrium $\big(\sigma, \boldsymbol{a}, \boldsymbol{\mu}\big)$ is* <u>*ex-post efficient*</u> *if for every voter v_θ such that $g^v(\theta) > 0$:*

$$\arg\max_{a \in \mathbb{A}} \mathbb{E}_{\mu_\theta}\Big[u_\theta(a, \theta^{ch})\Big] = 1 \iff d(\theta, \theta^{ch}) \leq d(\theta, \theta^{inc}) \text{ for all } \theta^{ch} \in supp\ \mu_\theta.$$

In an ex-post inefficient equilibrium, voters sometimes regret their choices. To see this, suppose voter v_θ votes for the challenger after hearing message m that includes some $\tilde{\theta}^{ch}$ that is strictly further away from her than the incumbent. Then, this voter regrets her choice with the positive probability of $\tilde{\theta}^{ch}$ getting elected. In an ex-post efficient equilibrium, no voter would ever vote for the challenger if there is a non-zero chance that the challenger is located strictly further away from this voter than the incumbent. In other words, in an ex-post efficient equilibrium the choices of voters are not too distorted by uncertainty, and there is no welfare loss.

Ex-post efficient equilibria possess a nice property in that the equilibrium strategies of the challenger and the voters are *robust* to changes in voters' prior beliefs: if $(\sigma, \boldsymbol{a}, \boldsymbol{\mu})$ is an ex-post efficient equilibrium of election $\left(\theta^{inc}, g^v(\cdot), p(\cdot)\right)$, then there exists an equilibrium $(\sigma, \boldsymbol{a}, \boldsymbol{\mu}')$ of election $\left(\theta^{inc}, g^v(\cdot), p'(\cdot)\right)$ for any other prior belief $p'(\cdot) > 0$. The reason for this it is that in an ex-post efficient equilibrium, any message that leads to the voter v voting for the challenger must be sent exclusively by θ^{ch} that are weakly closer to v than the incumbent (and at least one that is strictly closer to break the indifference). A change in the prior belief would lead to a change in the posterior belief, but not to a change in the voter's choice, since all that it does is shift the probability mass among the challengers, all of whom are weakly preferred to the incumbent.

Before I proceed to analyzing elections, I introduce some useful mathematical concepts associated with the metric space (Θ, d) that will be useful for analyzing elections.

Mathematical preliminaries

Definition 3.4. *$S \subseteq \Theta$ is convex if for any pair of positions $\tilde{\theta}, \hat{\theta} \in S$, every position $\theta \in \Theta$ such that $\theta_k \in \left[\min\{\tilde{\theta}_k, \hat{\theta}_k\}, \max\{\tilde{\theta}_k, \hat{\theta}_k\}\right]$ for every $k \in \mathbb{K}$ is also in S.*

Definition 3.5. *A convex hull of positions $\tilde{\theta}, \hat{\theta} \in \Theta$ is a set*

$$conv(\tilde{\theta}, \hat{\theta}) \equiv \left\{\theta \in \Theta \,\middle|\, \theta_k \in \left[\min\{\tilde{\theta}_k, \hat{\theta}_k\}, \max\{\tilde{\theta}_k, \hat{\theta}_k\}\right], \forall k \in \mathbb{K}\right\}.$$

Definition 3.6. *A path $\left(\theta^0 \to \theta^T\right)$ between positions $\theta^0, \theta^T \in \Theta$ is a sequence $\{\theta^t\}_{t=0}^T$, $\theta^t \in \Theta$ such that $d(\theta^t, \theta^{t+1}) = 1\ \forall t = 1 \dots T-1$, and $\sum_{t=0}^{T-1} d(\theta^t, \theta^{t+1}) = d(\theta^0, \theta^T)$. The path goes through $\theta \in \Theta$ if $\theta^t = \theta$ for some $t = 1 \dots T-1$.*

One may think of a path as a shortest walk that starts at θ^0, ends at θ^T, and each "step" is an edge between two positions θ^t and θ^{t+1} that are adjacent on one issue. The walk is shortest in the sense that the least possible number of steps is taken: by the triangle inequality, $\sum_{t=0}^{T-1} d(\theta^t, \theta^{t+1})$ must always be greater or equal than $d(\theta^0, \theta^T)$, and the fact that it holds as an equality means that no steps are taken in the "wrong" direction.

Below I list several intuitive properties of paths that will be useful in the rest of the paper.

Lemma 3.1 (Properties of Paths).

(1) *A path between* $\theta^0 \neq \theta^T$ *always exists; if* $\theta^0_k \neq \theta^T_k$ *for only one issue* $k \in \mathbb{K}$*, then the path is unique;*

(2) $\forall \theta^t \in \left(\theta^0 \to \theta^T\right)$, $\theta^t \in conv(\theta^0, \theta^T)$,

(3) *if* $\left(\theta^0 \to \theta^T\right) = \{\theta_t\}_{t=0}^T$*, then* $\forall \tau = 1 \dots T-1$, $\{\theta^t\}_{t=0}^{\tau} = \left(\theta^0 \to \theta^{\tau}\right)$ *and* $\{\theta^t\}_{t=\tau}^{T} = \left(\theta^{\tau} \to \theta^T\right)$*;*

(4) $\theta \in \left(\theta^0 \to \theta^T\right) \iff d(\theta^0, \theta) + d(\theta, \theta^T) = d(\theta^0, \theta^T)$.

In words, *(1)* states that some path between distinct policies always exists, and if the policies differ on at least two issues, there could be many of them; *(2)* states that all policies on the path must be in the convex hull of where you start and where you finish: when you travel from θ^0 to θ^T, you never have to go further on any issue than the difference on this issue between the beginning and the ending points; according to *(3)*, any path can be broken into a sequence of shorter paths; *(4)* states that a policy can be on a path if and only if it is not located too far from the start and the finish. The proof of Lemma 3.1 and all remaining proofs can be found in the appendix unless stated otherwise.

3.1 Public Disclosure

I begin analyzing elections by considering a special case, in which the challenger can only send the same public message to all voters. I show that there exists a wide class of elections that are unwinnable for the challenger in any equilibrium. For those that are winnable, I show which challenger types can win and what their messages will look like.

Under public disclosure (**PD**), the challenger maximizes his expected utility (equilibrium condition 1) subject to an additional constraint that requires that all message collections $\{m_\theta\} \in \mathbb{M}^{|\Theta|}$ are such that $m_\theta = m \in \mathbb{M}$ for all $\theta \in \Theta$. It follows directly that the challenger's behavior strategy will simplify to $\sigma : \Theta \times \mathbb{M} \to [0, 1]$, with $\sigma(\theta^{ch}, m)$ being the probability that message m is *publicly* sent by type θ^{ch}. It further follows that on the voters' side, the marginal probability that voter v_θ hears message m becomes independent of voter type, i.e. $\sigma_\theta(\theta^{ch}, m) = \sigma(\theta^{ch}, m)$, and all voters will have identical posterior beliefs (on and off the path) $\mu_\theta(\cdot, m) = \mu(\cdot, m)$. It is useful to characterize the voters the following way.

Definition 3.7. *Voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ agree on issue $\kappa \in \mathbb{K}$ if $\tilde{\theta}_\kappa, \hat{\theta}_\kappa > \theta^{inc}_\kappa$ or $\tilde{\theta}_\kappa, \hat{\theta}_\kappa < \theta^{inc}_\kappa$.*

Definition 3.8. *Voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are incompatible if they do not agree on any issue $k \in \mathbb{K}$, and are compatible otherwise.*

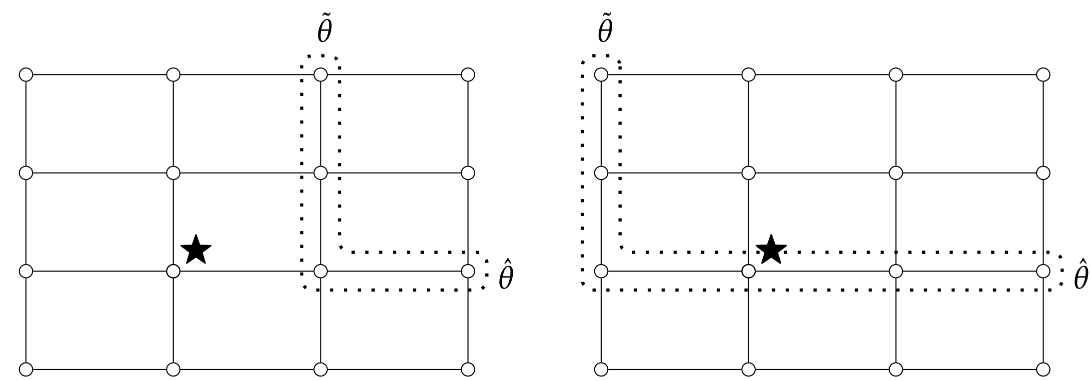
Compatibility of voters will play an important role under public disclosure. Lemma 3.2 provides alternative ways of determining incompatibility, and Figure 2 illustrates them.

Lemma 3.2. *The following conditions for voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are equivalent:*

(1) they are incompatible;

(2) $d(\tilde{\theta}, \hat{\theta}) = d(\tilde{\theta}, \theta^{inc}) + d(\hat{\theta}, \theta^{inc})$;[7]

(3) there exists a path $\left(\tilde{\theta} \to \hat{\theta}\right)$ that goes through θ^{inc}.



(a) *$v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are compatible: they agree on issue 1; $d(\tilde{\theta}, \hat{\theta}) < d(\tilde{\theta}, \bigstar) + d(\hat{\theta}, \bigstar)$; no path goes through $\bigstar$.*

(b) *$v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are incompatible: they do not agree on any issue; $d(\tilde{\theta}, \hat{\theta}) = d(\tilde{\theta}, \bigstar) + d(\hat{\theta}, \bigstar)$; drawn path goes through $\bigstar$.*

Figure 2: *Policy space Θ, $\bigstar$ is the incumbent's policy.*

In other words, the favorite positions of incompatible voters are on the opposite sides of the incumbent: for no issue will these voters agree that one of their positions is strictly better than the incumbent's. As a result, they both prefer the incumbent to their counterpart's ideal positions, or, according to Proposition 3.1, *any other position*.

[7] By the triangle inequality, $d(\tilde{\theta}, \hat{\theta}) \leq d(\tilde{\theta}, \theta^{inc}) + d(\hat{\theta}, \theta^{inc})$ for all $\tilde{\theta}, \hat{\theta}, \theta^{inc} \in \Theta$, which means $\tilde{\theta}$ and $\hat{\theta}$ are as far from each other, as the properties of $d(\cdot, \cdot)$ allow.

Proposition 3.1. *Incompatible voters never vote for the challenger at the same time under common belief $\mu(\cdot)$ over Ω. In other words, $\forall \tilde{\theta}, \hat{\theta} \in \Theta$, if $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are incompatible, then*

$$\mathbb{E}_{\mu}\Big[d(\tilde{\theta}, \theta^{ch})\Big] < d(\tilde{\theta}, \theta^{inc}) \Rightarrow \mathbb{E}_{\mu}\Big[d(\hat{\theta}, \theta^{ch})\Big] \geq d(\hat{\theta}, \theta^{inc}).$$

It follows directly that the challenger's electorate can only include pairwise compatible voters.

Definition 3.9. *Coalition of compatible voters $\mathcal{CV}$ is a subset of Θ such that for any $\tilde{\theta}$, $\hat{\theta} \in \mathcal{CV}$, voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are compatible.*

There are at most $2K$ "largest" coalitions of compatible voters, two for every issue in $\mathbb{K}$. Each one is characterized by a single issue κ that *all* voters within this coalition agree on, and whether they agree *to the left* ($\theta_{\kappa} < \theta^{inc}_{\kappa}$ for all v_{θ}), or *to the right* ($\theta_{\kappa} > \theta^{inc}_{\kappa}$ for all v_{θ}). These coalitions are referred to as the largest because every other $\mathcal{CV}$ will be a subset of one of them.

Definition 3.10. *$\mathcal{CV}$ is a (minimal) winning coalition of compatible voters if*

$$\sum_{\theta \in \mathcal{CV}} g^{v}(\theta) > \frac{1}{2} \text{ and } \sum_{\theta \in \mathcal{CV} \setminus \hat{\theta}} g^{v}(\theta) < \frac{1}{2} \text{ for all } \hat{\theta} \in \mathcal{CV}.$$

To win, the challenger would have to persuade every voter in some winning $\mathcal{CV}$ to vote for him. Before I proceed to the results on when the challenger can win an election, and how well he can do, I introduce the following definition for the challenger who wins under *complete information*.

Definition 3.11. *Challenger of type θ^{ch}_{w} is a winner under complete information of election $\mathcal{E}$ if the majority of voters prefers voting for him over the incumbent under complete information, i.e.*

$$\sum_{\theta \in \Theta} g^{v}(\theta) \cdot \mathbb{1}\Big\{d(\theta, \theta^{ch}_{w}) < d(\theta, \theta^{inc})\Big\} > \frac{1}{2}.$$

Note that if θ^{ch}_{w} is a *winner* of election $\mathcal{E}$, he *can always win* the election under public disclosure by announcing his true type, i.e. sending message $m = \{\theta^{ch}_{w}\}$ with probability one: $\sigma(\theta^{ch}_{w}, m) = 1$. Voters' belief after this message is always $\mu(\theta^{ch}_{w}, \{\theta^{ch}_{w}\}) = 1$ since no other type can send the same message under the grain of truth assumption. Consequently, every voter v_{θ} will best respond the way she would under complete information, since $\mathbb{E}_{\mu}\Big[u_{\theta}(a, \theta^{ch}_{w})\Big] = u_{\theta}(a, \theta^{ch}_{w})$, and the winner will get elected, i.e. $U^{ch}\Big(\theta^{ch}_w, \{a_\theta(m)\}\Big) = 1$. By (1), the ex-ante probability of the challenger winning this election will be *at least* $p(\theta^{ch}_w)$, and $\mathcal{E}$ will be winnable with probability of at least $p(\theta^{ch}_w) > 0$. Furthermore, since I require the grain of truth condition to also hold off the equilibrium path, the winner will win the election *in every equilibrium*, since he can always deviate to the fully revealing message $m = \{\theta^{ch}\}$.

When no challenger wins under full information, the following set of *negative* results holds.

Proposition 3.2. *The following statements are equivalent:*

1. *election $\mathcal{E}$ is unwinnable under public disclosure;*
2. *there is no winning coalition of compatible voters $\mathcal{CV}$;*
3. *no type of challenger $\theta^{ch} \in \Theta$ is a winner under complete information.*

This proposition not only connects impossibility results under complete and incomplete information, but also provides a link to the degree of heterogeneity of the population, since the presence of a winning $\mathcal{CV}$ is a property of $g^v(\cdot)$. Moreover, since there are at most $2K$ largest winning coalitions of compatible voters, it would take checking at most $2K$ linear inequalities to see if at least one of them is winning.[8] This provides a simple test for whether an election is winnable under public disclosure, and whether any challenger types could have won under complete information.

Corollary 3.1. *Proposition 3.2 has a particularly interesting implication if the number of issues is $K = 1$ (the policy space is one-dimensional). If there is no winning $\mathcal{CV}$, the masses of voters to the left and to the right of θ^{inc} are both less than one half, in which case θ^{inc} is the favorite policy of the median voter. Analogously to the* ***median voter theorem****, which states that θ^{inc} could not lose if his opponent could choose his location, Proposition 3.2 states that the incumbent also cannot lose if his opponent had an unknown position anywhere in the policy space and is allowed to send a public verifiable message about it.*[9]

Next I show which types of the challengers are guaranteed to win if there exists a winning $\mathcal{CV}$.

[8] For example, if there are two issues and θ^{inc} is in the interior of Θ, the four inequalities to check would be whether $\sum_\theta g^v(\theta) > 1/2$ for $\theta \in \Theta$ such that *(1)* $\theta_1 > \theta^{inc}_1$, *(2)* $\theta_1 < \theta^{inc}_1$, *(3)* $\theta_2 > \theta^{inc}_2$, and *(4)* $\theta_2 < \theta^{inc}_2$.

[9] Black (1948) states the median voter theorem as "If Θ is a single-dimensional issue and all voters have single-peaked preferences defined over Θ, then θ_m, the median position, could not lose under majority rule."

Proposition 3.3. *Consider election $\mathcal{E}$ and a winning coalition of compatible voters $\mathcal{CV}$.*

(1) *Every $\tilde{\theta}^{ch}$, such that $d(\theta, \tilde{\theta}^{ch}) < d(\theta, \theta^{inc})$ for every $\theta \in \mathcal{CV}$, is a winner. There is at least one such type.*

(2) *Every $\hat{\theta}^{ch}$, such that $d(\theta, \hat{\theta}^{ch}) \leq d(\theta, \theta^{inc})$ for every $\theta \in \mathcal{CV}$, can win the election by pooling with a winner, for any prior distribution $p(\cdot)$, in the challenger-preferred ex-post efficient equilibrium.*

For simplicity, suppose the election is such that all voters who are not in the winning $\mathcal{CV}$ are of type θ^{inc}. In $\mathcal{CV}$, all voters must agree on at least one issue; suppose they agree on issue $\kappa \in \mathbb{K}$ to the right. Then θ^{ch} such that $\theta^{ch}_k = \theta^{inc}_k$ for all $k \neq \kappa$ and $\theta^{ch}_\kappa = \theta^{inc}_\kappa + 1$ is trivially a winner, since he is closer to every v_θ in $\mathcal{CV}$ on the issue that they agree on and coincides with the incumbent on every other issue. Then, for every challenger type that is *adjacent* to θ^{ch} ($d(\theta^{ch}, \theta^{adj}) = 1$), it is the case that $d(\theta, \theta^{adj}) \leq d(\theta, \theta^{inc})$ for every $\theta \in \mathcal{CV}$, which also satisfies condition *(1)* or *(2)* of the proposition. By that logic, all challenger types that are winners, or that can win by pooling with the challenger, will be adjacent to each other. Moreover, the final set of types who win in equilibrium will be a connected set with winners under complete information in the interior, and types that win by pooling with winners (for example, ones they are adjacent to) on the boundary. A comprehensive example for the election game under public disclosure in a 2×2 policy space follows.

Example. (Elections in the 2×2 policy space, <u>*public disclosure*</u>)

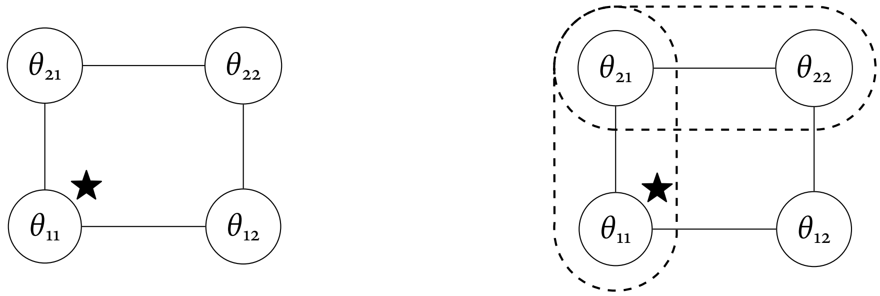
Election $\mathcal{E} = \left(\theta^{inc}, \{g^v_{ij}\}, \{p_{ij}\}\right)$ is winnable if $g^v_{21} + g^v_{22} > 1/2$ (then θ_{21} is a winner), or $g^v_{12} + g^v_{22} > 1/2$ (then θ_{12} is a winner). If both inequalities are true, then $g^v_{22} > 1/2$, and $\theta_{21}, \theta_{22}, \theta_{12}$ are all winners.

Consider election $\mathcal{E}$ in which θ_{21} is the sole winner under complete information. In terms of model primitives, this translates to

$$g^v_{21} + g^v_{22} > \frac{1}{2} \quad (\theta_{21} \text{ is a winner}),$$

$$g^v_{22} + g^v_{12} < \frac{1}{2} \quad (\theta_{22} \text{ and } \theta_{12} \text{ are not winners}).$$

These conditions imply that $\mathcal{CV} = \{v_{21}, v_{22}\}$ is the only winning coalition of compatible voters in this election: these voters agree on issue 2 and they constitute a majority. Proposition 3.3 states that in the challenger-preferred ex-post efficient equilibrium, θ_{21}, θ_{22}, and θ_{11} all win the election.



***(a)** The 2×2 policy space.*
If $g_{21} + g_{22} > 1/2$, then θ_{21} is a winner.

***(b)** Winning messages on the equilibrium path.*
θ_{11}, θ_{22} pool with the winner θ_{21}.

***Figure 3:** Public disclosure, illustrated.*

They do so by sending the following profile of public messages:

$$\theta_{21} \text{ sends } \frac{1}{2}\{\theta_{21}, \theta_{22}\} + \frac{1}{2}\{\theta_{21}, \theta_{11}\}$$

$$\theta_{22} \text{ sends } \{\theta_{21}, \theta_{22}\}$$

$$\theta_{11} \text{ sends } \{\theta_{21}, \theta_{11}\}$$

$$\theta_{12} \text{ sends any } m \in \mathbb{M} \text{ such that } \theta_{12} \in m.$$

Both messages $\{\theta_{21}, \theta_{22}\}$ and $\{\theta_{21}, \theta_{11}\}$ lead to the challenger winning the election, because he is perceived to be strictly closer than the incumbent by both v_{21} and v_{22}. This means that θ_{21}, θ_{22}, and θ_{11} all win. On the equilibrium path, any m that is sent by θ_{12} will generate posterior belief $\mu(\theta_{12}, m) = 1$, which will lead to the incumbent winning the election by securing votes of v_{21} and v_{11} (they constitute a majority). Off the path, the belief must be such that $\mu(\theta_{12}, m) > \mu(\theta_{21}, m)$ for $\forall m \in \mathbb{M}$ such that $\theta_{12} \in m$ to ensure that v_{21} does not want to vote for θ_{12} (ex-post efficiency). The off-the-path belief following any other message may be arbitrary (as long as the grain of truth assumption is satisfied), because θ_{21}, θ_{22}, θ_{11} win the election and have no profitable deviations. △

Proposition 3.3 characterizes the challenger-preferred ex-post efficient equilibrium under **PD**. Ex-post efficiency ensures that every coalition of compatible voters elects the challenger only if he is located not further away from every member of this coalition than the incumbent. The equilibrium is challenger-preferred because every type of challenger who is weakly preferred to the incumbent by *some* winning coalition of compatible voters wins the election.

There may exist other ex-post inefficient equilibria, in which even the challenger who is located on the fringe of the policy space may be winning with positive probability. Such equilibria are only possible provided that the election is already winnable; they strongly rely on restrictions on voters' prior beliefs (every voter in some winning $\mathcal{CV}$ must have a relatively high prior on positions of the challenger close to her favorite) and are left for future research.

3.2 Targeted Advertising

To illustrate that the challenger can do strictly better under targeted advertising than under public disclosure, I examine a simple class of elections that are unwinnable under the latter, but winnable under the former. For the rest of this section, suppose that any majority requires two types of incompatible voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$.[10] Such elections are trivially unwinnable under **PD**, because these voters never vote for the challenger under common belief. I show that there exists an equilibrium in which the challenger's profile of (targeted) messages induces heterogeneous beliefs of the incompatible voters in a way that persuades them both to vote for him at the same time. In other words, this section illustrates that the challenger *can persuade incompatible voters* under targeted advertising, which means he can win elections that are decided by incompatible voters.

Which types of the challenger can win the votes of incompatible voters? I focus on ex-post efficient equilibria, which require that any θ^{ch}, who $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ vote for, is weakly closer than the incumbent to both of them. This implies

$$\begin{array}{c} d(\tilde{\theta},\theta^{ch}) \le d(\tilde{\theta},\theta^{inc}), \\ d(\hat{\theta},\theta^{ch}) \le d(\hat{\theta},\theta^{inc}) \end{array} \implies d(\tilde{\theta},\theta^{ch}) + d(\hat{\theta},\theta^{ch}) = d(\tilde{\theta},\hat{\theta}), \tag{2}$$

which follows from *(i)* adding the inequalities on the left together, *(ii)* using property *(2)* of incompatible voters, and *(iii)* applying the triangle inequality. In other words, ex-post efficiency restricts our attention to the challenger types that are on the path $\left(\tilde{\theta} \to \hat{\theta}\right)$, located just as far from voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ as the incumbent.

[10] Since voters of type θ^{inc} always vote for the incumbent, any majority would require votes of both $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ if $g^v(\tilde{\theta}), g^v(\hat{\theta}) < 1/2$, $g^v(\tilde{\theta}) + g^v(\hat{\theta}) > 1/2$, and $g(\theta^{inc}) = 1 - g^v(\tilde{\theta}) - g^v(\hat{\theta})$.

How do they win? The "winning" collection of messages of each θ^{ch} includes $\tilde{m} \equiv conv(\theta^{ch}, \tilde{\theta})$ sent to $v_{\tilde{\theta}}$, and $\hat{m} \equiv conv(\theta^{ch}, \hat{\theta})$ sent to $v_{\hat{\theta}}$. In other words, $v_{\tilde{\theta}}$ receives a message, in which every single type, except θ^{ch}, is strictly preferred to the incumbent (most notably, her ideal policy $\tilde{\theta} \in \tilde{m}$), while θ^{ch} is just as good as the incumbent. In order for $v_{\tilde{\theta}}$ to strictly prefer voting for the challenger after hearing $\tilde{m}$, other challenger types $\tilde{\theta}^{ch} \in \tilde{m}$ must also be sending this message (i.e. pool with θ^{ch}) with positive probability. This is arranged by specifying of the off-the path beliefs such that $v_{\hat{\theta}}$ never votes for $\tilde{\theta}^{ch}$ (because he is further away from her than the incumbent), which implies that $\tilde{\theta}^{ch}$ never wins the majority of votes, and hence he is indifferent between all feasible messages.

Intuitively, by pooling with all the types that the voter strictly prefers to the incumbent, θ^{ch} puts more focus on issues that he and the voter agree on, and less focus on the issues in disagreement. Because incompatible voters disagree on all issues (which makes public disclosure ineffective), the challenger who agrees with the first voter on the first issue and with the second voter on the second issue can win both of their votes by drawing more attention to his position on issue k in his message to kth voter. This idea is formalized in the following proposition.

Proposition 3.4. *Consider election $\mathcal{E}$ such that*

$$g^v(\tilde{\theta}), g^v(\hat{\theta}) < \frac{1}{2}, \quad g^v(\tilde{\theta}) + g^v(\hat{\theta}) > \frac{1}{2}, \quad g^v(\theta^{inc}) = 1 - g^v(\tilde{\theta}) - g^v(\hat{\theta}),$$

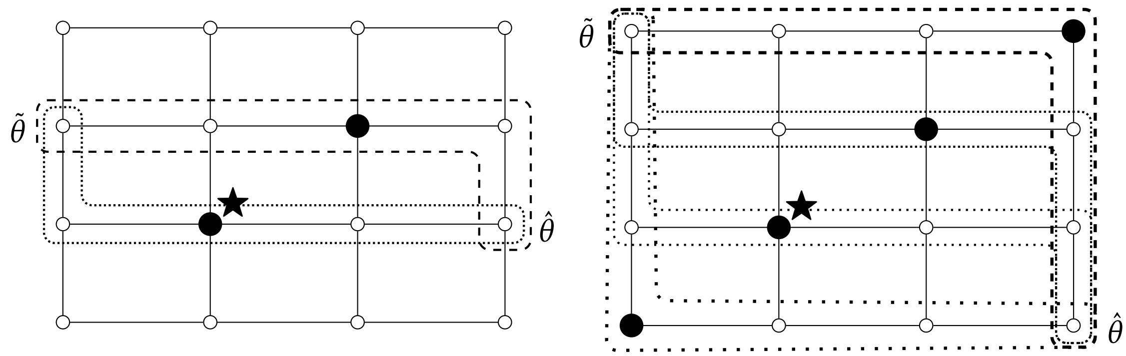
where $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are incompatible voters. Then, there exists an equilibrium, in which every θ^{ch} such that $d(\tilde{\theta}, \theta^{ch}) = d(\tilde{\theta}, \theta^{inc})$, $d(\hat{\theta}, \theta^{ch}) = d(\hat{\theta}, \theta^{inc})$ wins with probability one by sending message $\tilde{m}(\theta^{ch}) = conv(\theta^{ch}, \tilde{\theta})$ to $v_{\tilde{\theta}}$ and $\hat{m}(\theta^{ch}) = conv(\theta^{ch}, \hat{\theta})$ to $v_{\hat{\theta}}$.

Note that the described equilibrium is challenger-preferred in the class of ex-post efficient equilibria under targeted advertising. By construction, the voters who vote for the challenger weakly prefer him to the incumbent. At the same time, all challenger types that are weakly preferred to the incumbent by a majority of voters are winning at the same time.

Corollary 3.2. *The number of challenger types θ^{ch} who win the election by securing votes of the incompatible voters increases with the degree of polarization of said voters.*

To see this, recall that to satisfy (2), type θ^{ch} must lie of a path $\left(\tilde{\theta} \to \hat{\theta}\right)$, as far from both voters as the incumbent.[11] The number of paths between $\tilde{\theta}$ and $\hat{\theta}$ grows as *(i)* the range of positions $|\tilde{\theta}_k - \hat{\theta}_k|$ increases for all issues $k \in \mathbb{K}$, and *(ii)* the total number of issues (that the voters disagree on) grows.[12]

An illustration of how the degree of polarization of two voters affects the number of challengers who win both their votes under **TA** is presented in Figure 4.



(a) *$v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are not that polarized → there are few paths between them → not many challengers win.*

(b) *$v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ are very polarized → there are many paths between them → many challengers win.*

Figure 4: *Challenger types (in bold) who persuade incompatible voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$, depending on how polarized these voters are.*

The following example illustrates the challenger-preferred ex-post efficient equilibrium of an election that is decided by incompatible voters in the 2*x*2 policy space.

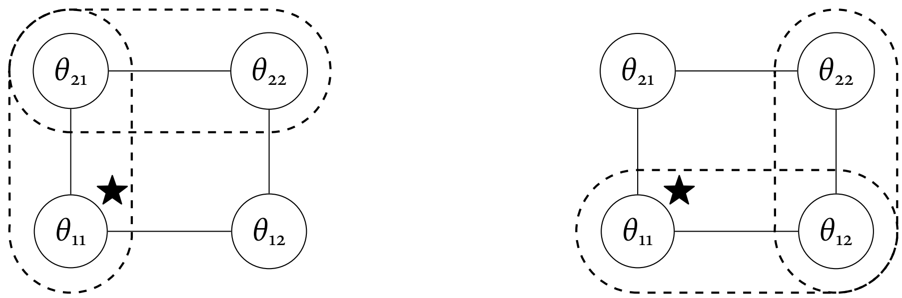
Example. (Elections in the 2×2 policy space, *targeted advertising*)

Suppose that the challenger wins the election if and only if he receives the votes incompatible voters v_{21} and v_{12}. In terms of model primitives, this translates to *(i)* their votes together, but not separately, constitute a majority: $g^v_{21}, g^v_{12} < \frac{1}{2}$, $g^v_{21} + g^v_{12} > \frac{1}{2}$, and *(ii)* all other voters always vote for the incumbent: $g^v_{11} = 1 - g^v_{21} - g^v_{12}$.

According to Proposition 3.4, there exists an equilibrium, in which challenger types θ_{11} and θ_{22} win this election, even though it is unwinnable under **PD**, with probability one. The equilibrium

[11] It trivially follows that if the policy space is one-dimensional, then there only exists one path between $\tilde{\theta}$ and $\hat{\theta}$, and this path goes through the incumbent. Proposition 3.4 then states that $\theta^{ch} = \theta^{inc}$ will be the only type of challenger to persuade $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$. It is hence more interesting to consider multidimensional policy spaces.

[12] For example, if Θ is two-dimensional, then the number of challenger types who can persuade incompatible voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ is $\min\left\{|\tilde{\theta}_1 - \hat{\theta}_1|, |\tilde{\theta}_2 - \hat{\theta}_2|\right\} + 1$ (see Figure 4). Clearly, it will increase if the range of positions on both issues goes up. If more dimensions are added, then the number of paths between the voters will grow, because there will be more directions to walk in.



(a) *v_{21} votes for the challenger after messages $\{\theta_{21}, \theta_{22}\}$ and $\{\theta_{21}, \theta_{11}\}$.*

(b) *v_{12} votes for the challenger after messages $\{\theta_{22}, \theta_{12}\}$ and $\{\theta_{11}, \theta_{12}\}$.*

Figure 5: *Winning messages under **TA**.*

profile of collections of messages is

$$\theta_{21} \text{ sends } \frac{1}{2}\begin{bmatrix} \{\theta_{21}, \theta_{11}\} \text{ to } v_{21} \\ \{\theta_{21}\} \text{ to others} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \{\theta_{21}, \theta_{22}\} \text{ to } v_{21} \\ \{\theta_{21}\} \text{ to others} \end{bmatrix},$$

$$\theta_{22} \text{ sends } \begin{matrix} \{\theta_{21}, \theta_{22}\} \text{ to } v_{21} \\ \{\theta_{22}, \theta_{12}\} \text{ to } v_{12}, \\ \{\theta_{22}\} \text{ to others} \end{matrix} \qquad \theta_{11} \text{ sends } \begin{matrix} \{\theta_{21}, \theta_{11}\} \text{ to } v_{21} \\ \{\theta_{11}, \theta_{12}\} \text{ to } v_{12}, \\ \{\theta_{11}\} \text{ to others} \end{matrix}$$

$$\theta_{12} \text{ sends } \frac{1}{2}\begin{bmatrix} \{\theta_{11}, \theta_{12}\} \text{ to } v_{12} \\ \{\theta_{12}\} \text{ to others} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \{\theta_{22}, \theta_{12}\} \text{ to } v_{12} \\ \{\theta_{12}\} \text{ to others} \end{bmatrix}.$$

In words, θ_{11} and θ_{22} pool with θ_{21} to persuade v_{21}, and with θ_{12} to persuade v_{12}. The mixed strategy by θ_{21} and θ_{12} is required in order to accommodate both θ_{11} and θ_{22} at the same time. When v_{21} hears $\{\theta_{21}, \theta_{11}\}$ or $\{\theta_{21}, \theta_{22}\}$, she puts positive probability of both types in each message. Her expected distance from the challenger becomes strictly less than one, which is the distance from the incumbent, so she votes for the challenger. By the same logic, so does v_{12}, and the challenger wins. Figure 5 summarizes the messages that persuade voters v_{21} and v_{12} and lead to θ_{11} and θ_{22} winning the election with probability one.

Any other message heard on the equilibrium path by any voter fully reveals the type of the sender, leads to a point belief by the grain of truth assumption, and results in the challenger acquiring the vote of at most one of v_{21} and v_{12} and losing the election. Note that ex-post efficiency is trivially satisfied on the equilibrium path, since v_{21} (v_{12}) never votes for the challenger of type θ_{12} (θ_{21}).

Off-the-path, it is sufficient to let $\mu_{12}(\theta_{21}, m_{21}) > \mu_{12}(\theta_{12}, m_{21})$ for any $m_{21} \in \mathbb{M}$ such that $\theta_{21} \in m_{21}$ (this ensures that v_{12} votes for the incumbent for any possible deviation of θ_{21}), and analogously, $\mu_{21}(\theta_{12}, m_{12}) > \mu_{21}(\theta_{21}, m_{12})$ for any $m_{12} \in \mathbb{M}$ such that $\theta_{12} \in m_{12}$. All other off-the-path beliefs are unrestricted (as long as the grain of truth condition is satisfied), since θ_{11} and θ_{22} are winning the election and do not have any profitable deviations. $\triangle$

The proposition illustrates the simplest election for which the set of equilibria under **TA** is larger than under **PD**. These elections are obviously unwinnable under public disclosure, since there is only one majority that consists of a pair of incompatible voters $v_{\tilde{\theta}}$ and $v_{\hat{\theta}}$ ($v_{\theta^{inc}}$ votes for the incumbent no matter what). Under targeted advertising, however, an equilibrium not only exists, but also has the following desirable properties:

- it is ex-post efficient, so no voter regrets her choice;
- equilibrium strategy profiles are robust to changes in prior beliefs;
- all messages heard on the equilibrium path are convex;[13]
- the ex-ante utility of the challenger is maximized in the class of ex-post efficient equilibria (all θ^{ch} that satisfy ex-post efficiency are winning at the same time).

One may observe that every θ^{ch} who wins under **TA** but loses under public disclosure is just as good for both voters as the incumbent (which will be true in any ex-post efficient equilibrium). This means that under a different tie-breaking rule, θ^{ch} may win under complete information or by disclosing his type truthfully. That said, the described equilibrium exists regardless of the tie-breaker, since each of the incompatible voters strictly prefers the challenger following a personalized message she receives, and would likely still exist if the voters were to face small costs of voting or have a slight preference for the incumbent.

---

[13] *Convex messages* can be thought of as a Cartesian product of intervals of policy positions on every issue. One can imagine that when the challenger sends a convex message, he simply lists a range of positions on every issue.

4 Conclusion

This paper examines the incentives of a political candidate, who is challenging the incumbent, to reveal his true position on relevant issues to uninformed voters. These incentives are studied for the case of public disclosure, in which the candidate advertises his position in the form of a public message, and the case of targeted advertising, in which each voter receives a private message tailored to this voter's political views. The main purpose of this paper is to examine whether targeting can swing elections. To answer this question, I characterize the set of equilibria under public disclosure and then show that some elections that cannot be won with a public message become winnable with targeting.

I find that the outcomes of elections under public disclosure are surprisingly easy to characterize. It is sufficient to check whether any group of voters who agree on some issue (that is, their position on this issue is *to the right* or *to the left* of the incumbent) constitutes a majority. If it does, then the challenger who is immediately to the right (or to the left) of the incumbent would win under complete information and hence by truthfully publicly revealing his position. If no such majority exists, then the electorate is polarized, because no majority agrees on any issue, and the challenger cannot win by sending a public message.

To understand how the challenger can change the outcome of a polarized election, I consider a simple set of elections which are decided by just two voters who do not agree on any issue. I find that these elections can be won by the challengers who agree with each voter on at least one issue, and are overall not further away from each voter than the incumbent. The message sent to each voter includes an interval of positions, ranging from this voter's preferred position to the challenger's true position, on every issue. These tailored messages persuade both voters at the same time, and this challenger wins the election. The more polarized the voters are, the more challengers agree with each voter on some issue, and hence the more challengers can swing elections.

References

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

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.

Besley, Timothy and Stephen Coate (1997), "An Economic Model of Representative Democracy", *The Quarterly Journal of Economics*, 112, 1 (Feb. 1997), pp. 85-114.

Black, Duncan (1948), "On the Rationale of Group Decision-making", *Journal of Political Economy*, 56, 1, pp. 23-34.

Board, Oliver (2009), "Competition and Disclosure", *The Journal of Industrial Economics*, 57, 1 (Feb. 2009), pp. 197-213.

Brennan, Geoffrey (1998), "Expressive Voting and Electoral Equilibrium", *Public Choice*, 95, pp. 149-175.

Clark, J. R. and Dwight R. Lee (2016), "Higher Costs Appeal to Voters: Implications of Expressive Voting", *Public Choice*, 167, 1-2, pp. 37-45.

Crawford, Vincent P. and Joel Sobel (1982), "Strategic Information Transmission", *Econometrica*, 50, 6 (Nov. 1982), p. 1431.

DellaVigna, Stefano and Matthew Gentzkow (2010), "Persuasion: Empirical Evidence", *Annual Review of Economics*, 2, 1 (Sept. 2010), pp. 643-669.

DellaVigna, Stefano and Ethan Kaplan (2007), "The Fox News Effect: Media Bias and Voting", *The Quarterly Journal of Economics*, 122, 3 (Aug. 2007), pp. 1187-1234.

Downs, Anthony (1957), "An Economic Theory of Political Action in a Democracy", *Journal of Political Economy*, 65, 2 (Apr. 1957), pp. 135-150.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya (2011), "Media and Political Persuasion: Evidence from Russia", *American Economic Review*, 101, 7 (Dec. 2011), pp. 3253-3285.

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

George, Lisa M and Joel Waldfogel (2006), "The New York Times and the Market for Local Newspapers", *American Economic Review*, 96, 1 (Feb. 2006), pp. 435-447.

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

Grossman, Sanford J. (1981), "The Informational Role of Warranties and Private Disclosure about Product Quality", *The Journal of Law and Economics*, 24, 3 (Dec. 1981), pp. 461-483.

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

Hotelling, Harold (1929), "Stability in Competition", *The Economic Journal*, 39, 153 (Mar. 1929), p. 41.

Jackson, Matthew O. and Xu Tan (2013), "Deliberation, Disclosure of Information, and Voting", *Journal of Economic Theory*, 148, 1 (Jan. 2013), pp. 2-30.

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.

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

Kim, Young Mie, Jordan Hsu, David Neiman, Colin Kou, Levi Bankston, Soo Yun Kim, Richard Heinrich, Robyn Baragwanath, and Garvesh Raskutti (2018), “The Stealth Media? Groups and Targets behind Divisive Issue Campaigns on Facebook”, *Political Communication*, pp. 1-29.

Koessler, Frédéric (2008), “Lobbying with two Audiences: Public vs Private Certification”, *Mathematical Social Sciences*, 55, 3 (May 2008), pp. 305-314.

Kreps, David M. and Robert Wilson (1982), “Sequential Equilibria”, *Econometrica*, 50, 4 (July 1982), p. 863.

Milgrom, Paul (1981), “Good News and Bad News: Representation Theorems and Applications”, *The Bell Journal of Economics*, 12, 2, p. 380.

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

Oberholzer-Gee, Felix and Joel Waldfogel (2009), “Media Markets and Localism: Does Local News en Español Boost Hispanic Voter Turnout?”, *American Economic Review*, 99, 5 (Dec. 2009), pp. 2120-2128.

Osborne, Martin J. and Al Slivinski (1996), “A Model of Political Competition with Citizen-Candidates”, *The Quarterly Journal of Economics*, 111, 1 (Feb. 1996), pp. 65-96.

Perego, Jacopo and Sevgi Yuksel (2018), “Media Competition and Social Disagreement”, *Mimeo*.

Prat, Andrea and David Strömberg (2013), “The Political Economy of Mass Media”, *Advances in Economics and Econometrics*, ed. by Daron Acemoglu, Manuel Arellano, and Eddie Dekel, Cambridge University Press, Cambridge, pp. 135-187.

Schipper, Burkhard C. and Hee Yeul Woo (2017), “Political Awareness and Microtargeting of Voters in Electoral Competition”, *SSRN Electronic Journal*.

Schnakenberg, Keith E (2015), “Expert advice to a voting body”, *Journal of Economic Theory*, 160 (Dec. 2015), pp. 102-113.

Wang, Yun (2013), “Bayesian Persuasion with Multiple Receivers”, *SSRN Electronic Journal*.

Williams, Christine B. and Girish J. Gulati (2018), “Digital Advertising Expenditures in the 2016 Presidential Election”, *Social Science Computer Review*, 36, 4 (Aug. 2018), pp. 406-421.

Appendix

Proof of Lemma 3.1

(1) Suppose that θ° and θ^T are different on issue $\kappa \in \mathbb{K}$, $\theta^\circ_\kappa \neq \theta^T_\kappa$, but $\forall k \in \mathbb{K} \setminus \{\kappa\}$ we have $\theta^\circ_k = \theta^T_k$. Then there exists a unique path $\left(\theta^\circ \to \theta^T\right) = \{\theta_t\}_{t=0}^T$, where $T = d(\theta^\circ, \theta^T) = |\theta^\circ_\kappa - \theta^T_\kappa|$, and

the sequence $\{\theta_t\}$ starts at θ_0 and with each step gets closer to θ_T, in the only direction that is available.

If θ^0 and θ^T differ on more than one issue, there always exists a path in which the distance is traveled issue-by-issue, i.e.

$$\left(\left(\theta_1^0, \theta_2^0, \theta_3^0, \ldots, \theta_K^0\right) \to \left(\theta_1^T, \theta_2^0, \theta_3^0, \ldots, \theta_K^0\right)\right), \text{ path of length } |\theta_1^0 - \theta_1^T|,$$

$$\left(\left(\theta_1^T, \theta_2^0, \theta_3^0, \ldots, \theta_K^0\right) \to \left(\theta_1^T, \theta_2^T, \theta_3^0, \ldots, \theta_K^0\right)\right), \text{ path of length } |\theta_2^0 - \theta_2^T|,$$

$$\ldots$$

$$\left(\left(\theta_1^T, \theta_2^T, \theta_3^T, \ldots, \theta_K^0\right) \to \left(\theta_1^T, \theta_2^T, \theta_3^T, \ldots, \theta_K^T\right)\right), \text{ path of length } |\theta_K^0 - \theta_K^T|,$$

and the total distance traveled will be $\sum_{k=1}^{K} |\theta_k^0 - \theta_k^T| = d(\theta^0, \theta^T)$, by definition of the L_1 distance.

(2) Suppose, on the contrary, that $\exists \kappa \in \mathbb{K}$ such that $\theta_\kappa^\tau \notin \left[\min\{\theta_\kappa^0, \theta_\kappa^T\}, \max\{\theta_\kappa^0, \theta_\kappa^T\}\right]$ for some $\tau = 1 \ldots T-1$. Without loss of generality, assume that $\theta_\kappa^\tau > \theta_\kappa^T > \theta_\kappa^0$. We have

$$\sum_{t=0}^{T-1} d(\theta^t, \theta^{t+1}) = d(\theta^0, \theta^T) \iff \forall k \in \mathbb{K} \ \sum_{t=0}^{T-1} |\theta_k^t - \theta_k^{t+1}| = |\theta_k^0 - \theta_k^T|,$$

since otherwise the triangle inequality will be violated. Then, applying the triangle inequality again, we get

$$\sum_{t=0}^{T-1} |\theta_\kappa^t - \theta_\kappa^{t+1}| \geq |\theta_\kappa^0 - \theta_\kappa^\tau| + |\theta_\kappa^\tau - \theta_\kappa^T| = \theta_\kappa^\tau - \theta_\kappa^0 + \theta_\kappa^\tau - \theta_\kappa^T > \theta_\kappa^T - \theta_\kappa^0 = |\theta_\kappa^T - \theta_\kappa^0|,$$

which is a contradiction.

(3) Since $d(\theta^t, \theta^{t+1}) = 1$ for all $t = 0 \ldots T-1$, it is only left to show that $\sum_{t=0}^{\tau-1} d(\theta^t, \theta^{t+1}) = d(\theta^0, \theta^\tau)$, and $\sum_{t=\tau}^{T-1} d(\theta^t, \theta^{t+1}) = d(\theta^\tau, \theta^T)$. This must be true since

$$d(\theta^0, \theta^T) = \sum_{t=0}^{\tau-1} d(\theta^t, \theta^{t+1}) + \sum_{t=\tau}^{T-1} d(\theta^t, \theta^{t+1}) \geq d(\theta^0, \theta^\tau) + d(\theta^\tau, \theta^T) \geq d(\theta^0, \theta^T).$$

(4) If $\theta \in \left(\theta^0 \to \theta^T\right)$, then by *(3)* we can split the path in two, letting $\theta^\tau = \theta$. It will follow directly

that $d(\theta^0, \theta) + d(\theta, \theta^T) = d(\theta^0, \theta^T)$.

If $d(\theta^0, \theta) + d(\theta, \theta^T) = d(\theta^0, \theta^T)$ for some $\theta \in \Theta$, define paths $\left(\theta^0 \to \theta\right) = \{\theta^t\}_{t=0}^{\tau}$ and $\left(\theta \to \theta^T\right) = \{\theta^t\}_{t=\tau}^{T}$. Then,

$$\sum_{t=0}^{\tau-1} d(\theta^t, \theta^{t+1}) + \sum_{t=\tau}^{T-1} d(\theta^t, \theta^{t+1}) = d(\theta^0, \theta) + d(\theta, \theta^T) = d(\theta^0, \theta^T),$$

so $\left(\theta^0 \to \theta^T\right) = \{\theta^t\}_{t=0}^{T}$ is also a path, and $\theta \in \left(\theta^0 \to \theta^T\right)$.

Proof of Lemma 3.2

(1) $\iff$ ***(2)***: using the definition of the L_1 distance,

$$d(\tilde{\theta}, \theta^{inc}) + d(\hat{\theta}, \theta^{inc}) = d(\tilde{\theta}, \hat{\theta}) \iff \sum_{k=1}^{K} |\tilde{\theta}_k - \theta_k^{inc}| + |\hat{\theta}_k - \theta_k^{inc}| = \sum_{k=1}^{K} |\tilde{\theta}_k - \hat{\theta}_k|.$$

This equality is possible if and only if $\nexists \kappa \in \mathbb{K}$ such that $|\tilde{\theta}_\kappa - \theta_\kappa^{inc}| + |\hat{\theta}_\kappa - \theta_\kappa^{inc}| > |\tilde{\theta}_\kappa - \hat{\theta}_\kappa|$, since $|\tilde{\theta}_k - \theta_k^{inc}| + |\hat{\theta}_k - \theta_k^{inc}| \geq |\tilde{\theta}_k - \hat{\theta}_k|$ for every $k \in \mathbb{K}$. This, in turn, is possible if and only if $\nexists \kappa \in \mathbb{K}$ such that $\tilde{\theta}_\kappa, \hat{\theta}_\kappa > \theta_\kappa^{inc}$ or $\tilde{\theta}_\kappa, \hat{\theta}_\kappa < \theta_\kappa^{inc}$.

(2) $\iff$ ***(3)***: holds by property *(4)* of path $\left(\tilde{\theta} \to \hat{\theta}\right)$.

Proof of Proposition 3.1

Let v_{θ_i} and v_{θ_j} be incompatible voters. Let $x = d(\theta_i, \theta_j)$, $x_i = d(\theta_i, \theta^{inc})$, $x_j = d(\theta_j, \theta^{inc})$. It follows from the definition that $x_i + x_j = x$.

Voter v_{θ_i} votes for the challenger if the expected distance from the challenger under belief $\mu(\cdot)$ is strictly less than the distance from the incumbent x_i:

$$\underbrace{0}_{=d(\theta_i,\theta_i)} \cdot \mu_i + \underbrace{x}_{=d(\theta_i,\theta_j)} \cdot \mu_j + \sum_{k \neq i,j} d(\theta_i, \theta_k) \cdot \mu_k < x_i = x_i \cdot \underbrace{\left[\mu_i + \mu_j + \sum_{k \neq i,j} \mu_k\right]}_{=1 \text{ since } \mu(\cdot) \text{ is a prob. distr.}},$$

which can be rearranged to

$$\mu_j < \frac{1}{x - x_i}\left[x_i\mu_i + x_i \sum_{k \neq i,j} \mu_k - \sum_{k \neq i,j} d(\theta_i, \theta_k) \cdot \mu_k\right]. \tag{3}$$

Similarly, v_{θ_j} votes for the challenger if

$$0 \cdot \mu_j + x \cdot \mu_i + \sum_{l \neq i,j} d(\theta_j, \theta_l) \cdot \mu_l < x_j = x_j \cdot \left[\mu_j + \mu_i + \sum_{l \neq i,j} \mu_l\right],$$

which is equivalent to

$$\mu_j > \frac{1}{x_j}\left[x \cdot \mu_i + \sum_{l \neq i,j} d(\theta_j, \theta_l) \cdot \mu_l - x_j\mu_i - x_j \cdot \sum_{l \neq i,j} \mu_l\right]. \tag{4}$$

If v_i and v_j both vote for the challenger, conditions (3) and (4) must hold simultaneously. Combining these inequalities, we get[14]

$$x \cdot \mu_i + \sum_{l \neq i,j} d(\theta_j, \theta_l) \cdot \mu_l - x_j\mu_i - x_j \cdot \sum_{l \neq i,j} \mu_l < x_i\mu_i + x_i \sum_{k \neq i,j} \mu_k - \sum_{k \neq i,j} d(\theta_i, \theta_k) \cdot \mu_k$$

$$\iff \sum_{k \neq i,j} d(\theta_i, \theta_k) \cdot \mu_k + \sum_{l \neq i,j} d(\theta_j, \theta_l) \cdot \mu_l < x_i \sum_{k \neq i,j} \mu_k + x_j \cdot \sum_{l \neq i,j} \mu_l$$

$$\iff \sum_{k \neq i,j} \mu_k \cdot \left[d(\theta_i, \theta_k) + d(\theta_j, \theta_k)\right] < \sum_{k \neq i,j} \mu_k \cdot d(\theta_i, \theta_j),$$

which contradicts the triangle inequality. Thus, the voters v_{θ_i} and v_{θ_j} cannot both vote for the challenger under common belief.

Proof of Proposition 3.2

(2) $\Rightarrow$ ***(1), (3)***: according to Proposition 3.1, only compatible voters vote for the challenger under common belief. Since no mass of compatible voters constitutes a majority, under no common belief (including any common posterior that the challenger may be able to induce using behavior strategies, and complete information) would a majority of voters vote for the

[14] Since $x - x_i = x_j$, we will multiply both sides of both inequalities by x_j.

challenger.

(1) $\Rightarrow$ ***(2)*** by contradiction: suppose $\mathcal{E}$ is winnable. Then there must exist a message $m \in \mathbb{M}$ on the equilibrium path which induces posterior belief $\mu(\cdot, m)$ under which a majority of voters prefers to vote for the challenger, which means

$$\exists \mathcal{V} \subseteq \Theta \text{ s.t. } \forall \theta \in \mathcal{V}: \ \mathbb{E}_{\mu}\Big[d(\theta, \theta^{ch})\Big] < d(\theta, \theta^{inc}) \text{ and } \sum_{\theta \in \mathcal{V}} g^{v}(\theta) > \frac{1}{2}.$$

Any two voters $v_{\tilde{\theta}}, v_{\hat{\theta}} \in \mathcal{V}$ must be compatible, since they both voted for the challenger under common belief $\mu(\cdot, m)$. Hence, $\exists \mathcal{CV} \subseteq \mathcal{V}$ that is winning, which contradicts the premise.

(3) $\Rightarrow$ ***(2)*** by contraposition, is logically equivalent to proving that if there exists a winning $\mathcal{CV}$ then some $\theta^{ch}_{w} \in \Theta$ must be a winner. This statement is true according to Proposition 3.3.

Proof of Proposition 3.3

(1) All voters in winning $\mathcal{CV}$ must agree on at least one issue. Without loss of generality, suppose that they all agree *to the right* on issue $\kappa \in \mathbb{K}$, so $\theta_{\kappa} > \theta^{inc}_{\kappa}$ for every voter v_{θ}. Then, type θ^{ch} such that $\theta^{ch}_{\kappa} = \theta^{inc}_{\kappa} + 1$ and $\theta^{ch}_{k} = \theta^{inc}_{k}$ for all $k \in \mathbb{K} \setminus \{\kappa\}$ must be a winner, since $\forall \theta \in \mathcal{CV}$

$$d(\theta, \theta^{ch}) = \sum_{k=1}^{K} |\theta_k - \theta^{ch}_k| = \sum_{k \in \mathbb{K}, \neq \kappa} |\theta_k - \theta^{ch}_k| + |\theta_{\kappa} - \theta^{ch}_{\kappa}| = \sum_{k \neq \kappa} |\theta_k - \theta^{inc}_k| + \theta_{\kappa} - \theta^{inc}_{\kappa} - 1 < d(\theta, \theta^{inc}).$$

Furthermore, any other $\tilde{\theta}^{ch}$ s.t., such that $d(\theta, \tilde{\theta}^{ch}) < d(\theta, \theta^{inc})$ for every $\theta \in \mathcal{CV}$, is a winner by definition (this set may be empty).

(2) Suppose $\hat{\theta}^{ch}$ pools with winner $\tilde{\theta}^{ch}$, that is, he sends message $m = \{\hat{\theta}^{ch}, \tilde{\theta}^{ch}\}$ all the time, i.e. $\sigma(\hat{\theta}^{ch}, m) = 1$. The winner, in turn, sends this message only with probability $\alpha \in (0,1]$, i.e. $\sigma(\tilde{\theta}^{ch}, m) = \alpha$. The common posterior belief of the voters following this message is

$$\tilde{\mu} \equiv \mu(\tilde{\theta}^{ch}, m) = \frac{\alpha \cdot p(\tilde{\theta}^{ch})}{\alpha \cdot p(\tilde{\theta}^{ch}) + p(\hat{\theta}^{ch})} > 0 \text{ and } \hat{\mu} \equiv \mu(\hat{\theta}^{ch}, m) = \frac{p(\hat{\theta}^{ch})}{\alpha \cdot p(\tilde{\theta}^{ch}) + p(\hat{\theta}^{ch})} > 0.$$

Then, for every $\theta \in \mathcal{CV}$

$$\mathbb{E}_{\mu}\Big[d(\theta, \theta^{ch})\Big] = \tilde{\mu} \cdot d(\theta, \tilde{\theta}^{ch}) + \hat{\mu} \cdot d(\theta, \hat{\theta}^{ch}) < d(\theta, \theta^{inc}),$$

no matter how small $\alpha > 0$ is. Then the winner $\tilde{\theta}^{ch}$ can pool with any finite number of $\hat{\theta}^{ch}$ for any prior distribution $p(\cdot)$ with a full support.

Proof of Theorem 3.4

__Messages__: every θ^{ch} sends message $\tilde{m}(\theta^{ch}) = conv(\theta^{ch}, \tilde{\theta})$ to voter $v_{\tilde{\theta}}$. Every other type contained in this message sends the same message with a strictly positive probability; in other words, $\sigma\Big(\tilde{\theta}^{ch}, \tilde{m}(\theta^{ch})\Big) = \tilde{\alpha}(\tilde{\theta}^{ch}, \theta^{ch}) \in (0,1]$ for every $\tilde{\theta}^{ch} \in conv(\theta^{ch}, \tilde{\theta})$, with $\sigma\Big(\theta^{ch}, \tilde{m}(\theta^{ch})\Big) = \tilde{\alpha}(\theta^{ch}, \theta^{ch}) = 1$. Analogously, θ^{ch} sends message $\hat{m}(\theta^{ch}) = conv(\theta^{ch}, \hat{\theta})$ to voter $v_{\hat{\theta}}$, and other types $\hat{\theta}^{ch}$ in that message send it with probability $\hat{\alpha}(\hat{\theta}^{ch}, \theta^{ch}) \in (0,1]$.

__Voting__: suppose voter $v_{\tilde{\theta}}$ hears message $\tilde{m}(\theta^{ch})$. This message is a convex hull of θ^{ch}, such that $d(\tilde{\theta}, \theta^{ch}) = d(\tilde{\theta}, \theta^{inc})$, and $\tilde{\theta}$, which is the voter's favorite position, which is strictly closer than the incumbent's. Then,

$$\mathbb{E}_{\mu_{\tilde{\theta}}}\Big[d(\tilde{\theta}, \theta^{ch})\Big] = \sum_{\tilde{\theta}^{ch} \in \Theta} \mu_{\tilde{\theta}}\Big(\tilde{\theta}^{ch}, \tilde{m}(\theta^{ch})\Big) \cdot d(\tilde{\theta}, \tilde{\theta}^{ch})$$

$$= \sum_{\tilde{\theta}^{ch} \in \Theta} \frac{\tilde{\alpha}\Big(\tilde{\theta}^{ch}, \tilde{m}(\theta^{ch})\Big) \cdot p(\tilde{\theta}^{ch})}{\sum_{\theta' \in \Theta} \tilde{\alpha}\Big(\theta', \tilde{m}(\theta^{ch})\Big) \cdot p(\theta')} \cdot d(\tilde{\theta}, \tilde{\theta}^{ch}) < d(\tilde{\theta}, \theta^{inc})$$

since $\tilde{\alpha}(\cdot, \theta^{ch})$ has a full support as long as $p(\cdot) > 0$, $d(\tilde{\theta}, \tilde{\theta}^{ch}) \le d(\tilde{\theta}, \theta^{inc})$ for all $\tilde{\theta}^{ch}$, and $d(\tilde{\theta}, \tilde{\theta}^{ch}) = d(\tilde{\theta}, \theta^{inc})$ for $\tilde{\theta}^{ch} = \tilde{\theta}$. Hence, $v_{\tilde{\theta}}$ will vote for the challenger after hearing $\tilde{m}(\theta^{ch})$. By the same logic, $\hat{\theta}$ will vote for the challenger after message $\hat{m}(\theta^{ch})$.

__Equilibrium conditions__: here I show that every $\tilde{\theta}^{ch} \in conv(\theta^{ch}, \tilde{\theta}) \setminus \{\theta^{ch}\}$ may never win the vote of $v_{\hat{\theta}}$ (and hence the election) in an ex-post efficient equilibrium (or have any profitable deviations), and so he will be indifferent between all messages $m \in \mathbb{M}$ that satisfy the grain of truth condition. That is, I show that $d(\hat{\theta}, \tilde{\theta}^{ch}) > d(\hat{\theta}, \theta^{inc})$. Suppose, on the contrary, that $\exists \tilde{\theta}^{ch} \in conv(\theta^{ch}, \tilde{\theta}) \setminus \{\theta^{ch}\}$ such that $d(\hat{\theta}, \tilde{\theta}^{ch}) \le d(\hat{\theta}, \theta^{inc})$. Since $\tilde{\theta}^{ch} \in conv(\theta^{ch}, \tilde{\theta}) \setminus \{\theta^{ch}\}$, it must also be true that $d(\tilde{\theta}, \tilde{\theta}^{ch}) < d(\tilde{\theta}, \theta^{inc})$, since the equality holds for θ^{ch}, and every other element of the convex hull is relatively closer to $\tilde{\theta}$ than the incumbent. But then, if we add these inequalities together,

$$d(\hat{\theta}, \tilde{\theta}^{ch}) + d(\tilde{\theta}, \tilde{\theta}^{ch}) < d(\hat{\theta}, \theta^{inc}) + d(\tilde{\theta}, \theta^{inc}) = d(\hat{\theta}, \tilde{\theta}),$$

which violates the triangle inequality.

Voter $v_{\tilde{\theta}}$'s off-the-equilibrium-path beliefs should be skeptical enough, in a sense that if he hears a message that is not $\tilde{m}(\theta^{ch})$ for some θ^{ch}, then she puts probability one on the type that is furthest away from her. This will guarantee that no challenger types have profitable deviations. Now, since every $\tilde{\theta}^{ch} \neq \theta^{ch}$ does not have a profitable deviation, he will not have an incentive to deviate from any strategy, including $\tilde{\alpha}(\tilde{\theta}^{ch}, \theta^{ch})$.

Maximizing the ex-ante utility of the challenger: it is possible for every θ^{ch} to win the election in the same ex-post efficient equilibrium with probability one (which would maximize his ex-ante probability of winning). Suppose there are $|X|$ total types θ^{ch}_x ($x \in X$), each of whom can win the election by persuading incompatible voters $v_{\hat{\theta}}$ and $v_{\tilde{\theta}}$ by sending message $conv(\theta^{ch}_x, \tilde{\theta})$ to $v_{\tilde{\theta}}$ and $conv(\theta^{ch}_x, \hat{\theta})$ to $v_{\hat{\theta}}$, which other types in these message also send with some positive probability, since they lose the election anyway.

Since there are multiple types θ^{ch}_x (one of which is always θ^{inc}), it is possible that some type $\tilde{\theta}^{ch}$ may be in $\tilde{m}(\theta^{ch}_x) = conv(\theta^{ch}_x, \tilde{\theta})$ for more than one θ^{ch}_x. The most obvious example of this is $\tilde{\theta}$, who will be a part of every $\tilde{m}(\theta^{ch})$. The way to deal with this is by letting $\sigma(\tilde{\theta}, \tilde{m}(\theta^{ch})) = \tilde{\alpha}(\tilde{\theta}, \theta^{ch}) = \frac{1}{|X|}$, and for every other $\tilde{\theta}^{ch}$ who needs to send more than one message, to uniformly randomize between these messages. Since we proved the existence for any $\tilde{\alpha} > 0$ and Θ is finite, and because $\tilde{\theta}^{ch}$ is indifferent between all feasible messages, the equilibrium conditions will still hold.