Political Disagreement and Information in Elections*

RICARDO ALONSO† ODILON CÂMARA‡

London School of Economics *University of Southern California*

October 14, 2015

Preliminary and Incomplete Draft

Abstract

We study a probabilistic voting model in which candidates representing two different groups of voters compete for office. In equilibrium, the candidate representing the majority group wins with a probability that increases in the degree of political disagreement — the difference in expected payoffs from the policies supported by the candidates. Prior to the election, the incumbent party (IP) is able to influence voters' behavior by designing a policy experiment, i.e. a public signal about a payoff-relevant state. We show that if the IP supports the majority candidate, then it uses this experiment to increase political disagreement and hence her victory probability. We then define conditions such that (i) the IP optimally chooses an upper-censoring experiment, which fully reveals low disagreement states and pools high disagreement states, and (ii) the experiment's informativeness decreases with the majority candidate's competence. Finally, we show how the IP uses information to increase disagreement even when all voters share the same payoff function, so that political disagreement is solely due to belief disagreement.

JEL classification: D72, D83.

Keywords: Disagreement, Bayesian persuasion, strategic experimentation, voting.

---

*We thank Patrick Le Bihan and Galina Zudenkova for their helpful discussions, as well as the audiences at the 2015 MPSA Conference, 2015 Princeton-Warwick Conference on Political Economy, 2015 ESSET Conference, and 2015 SAET Conference for their suggestions.

†LSE, Houghton Street, London WC2A 2AE, United Kingdom. `R.Alonso@lse.ac.uk`

‡USC FBE Dept, 3670 Trousdale Parkway Ste. 308, BRI-308 MC-0804, Los Angeles, CA 90089-0804. `ocamara@marshall.usc.edu`.

1 Introduction

In his classic paper, Stokes (1963) highlights some empirical regularities that should be present in electoral models. First, voters care about multiple issues, and hence may have to trade-off the candidate that offers the highest payoff in one issue for the competing candidate that offers the highest payoff in a different, more relevant issue. Second, it is important to differentiate position-issues and valence-issues.[1] Third, voter behavior only depends on voter's *perception* of parties' ideological positions. Fourth, information may change this perception, thus leading a voter to revise the relative importance of the different issues. To the extent that voters' perceptions can be shaped by information, interested parties have an incentive to affect voters' learning, and control the salience of position- and valence-issues (see Iyengar and Simon 2000 for a survey).

In this paper we incorporate these empirical regularities into a persuasion model. We focus on how information shapes political disagreement, and how disagreement steers politics. We consider an incumbent party (IP) who supports one candidate and wants to maximize her victory probability. Through its control of the government, the IP is able to influence voters' learning about a particular position-issue. To easy exposition, we focus on the IP's ability to strategically design a *policy experiment*. The experiment's results serve as a public signal about the payoff consequences of different policies.[2] In equilibrium, the IP designs the policy experiment to change voters' perception of the degree of political disagreement between the candidates, which endogenously shifts the relative salience of policy and valence issues. We show that when a majority of voters share a similar policy view as the candidate supported by the IP, the IP designs the experiment with the sole purpose of *increasing* political disagreement, which benefits its supported candidate.[3]

[1]In position-issues, parties advocate certain actions, and voters might have heterogenous preferences over actions. In valence-issues, parties are linked to some condition that is positively or negatively valued by the electorate.

[2]Alternatively, the IP could influence voter's learning by controlling which information is publicly generated by government agencies, as part of their activities.

[3]Our IP resembles the persuaders in Downs (1957, pg. 83), who "are not interested *per se* in helping people who are uncertain become less so; they want certainty to produce a decision which aids their cause." The strategic use of information to shift the salience of issues is highlighted by Stokes (1963, pg. 372), who says that "the skills of political leaders who must maneuver for public support in a democracy consist partly

To illustrate the roles of political disagreement and information, consider the following example. Voters who care about a position-issue (policy) and a valence-issue (competence) are divided into two groups: a majority A that has a lower per capita income and a minority B that has a higher per capita income. One citizen-candidate from each group runs for office. The elected official must choose a policy: the rate of a proportional income tax. The government uses all tax revenues to finance a novel government program.[4] Political disagreement arrises since voters in group A prefer a higher proportional tax than voters in B. Candidates cannot commit to future policies, but voters can predict their behavior: candidate A will implement a higher tax than candidate B. Candidate A has a *policy advantage* — a majority of voters believe that she will choose a better policy. Voters in majority A are nonetheless willing to vote for the minority candidate if during the campaign they learn that she is sufficiently more competent.[5] A key observation is that voters are less willing to trade off policy for valence if voters' perception of the degree of political disagreement is higher. That is, if the difference in voters' expected payoffs from the policies supported by the candidates is higher. Suppose voters are uncertain about the marginal payoff derived from the government program. Political disagreement is low if this expected marginal payoff is low, in which case both groups want similar low taxes. However, disagreement is high if this expectation is high, in which case group A wants much higher taxes than B (see Section 4.1 for details). Therefore, new information that leads players to belief that the program's payoff is higher increases political disagreement. The majority group is then more likely to vote for candidate A, increasing her victory probability. Consequently, information can be strategically used to shape political disagreement and steer the election in favor of a candidate.

To study the supply of information in elections, we consider a Bayesian persuasion game (see Kamenica and Gentzkow 2011, KG henceforth). Voters are uncertain about an underlying state that describes how different policies map into payoffs. One "sender" can sway voters' decision by designing what a group of "receivers" (voters) can learn from a public signal, i.e. by specifying the statistical relation of the signal with the underlying state. In

in knowing what issue dimensions [...] can be made salient by suitable propaganda."

[4]E.g., a novel after-school program to be adopted by public schools, aimed to help low-income students.

[5]Voters for which the final vote goes in consonance with valence preferences, rather than policy preferences, are dubbed "Stokes voters" by Groseclose (2001).

our benchmark model we interpret the sender to be the office-motivate party that currently controls the government, and wants to influence the outcome of the next election. We interpret the public signal to be a policy experiment. For example, the incumbent party can commission a trial on selected public schools to study the effects of different educational policies. While the IP may not be in control of the study's final conclusion, it can nevertheless shape public learning by dictating which questions the report should answer or which angle it should consider.

Section 2 presents the basic model, which features the following ingredients: (i) Electorate: Uninformed voters are divided into two groups, majority A and minority B, with differing preferences about the optimal policy. (ii) Parties and Candidates: Two parties compete for office. Party $\mathcal{A}$ runs with the incumbent candidate A, who will implement the preferred policy of group A if elected; while candidate B from party $\mathcal{B}$ defends the preferred policy of group B. Besides their supported policies, candidates also differ in a second dimension, competence. In this scenario candidate A holds an initial policy advantage, since a majority of voters prefer her policies over that of candidate B. (iii) Policy Experiment: Party $\mathcal{A}$ currently controls the government, and hence can design a policy experiment, which reveals information about a unknown state that is relevant for voters' policy payoffs. Party leaders (or bureaucrats) are purely office-motivated, thus Party $\mathcal{A}$ chooses the experiment that maximizes the victory probability of its candidate. (iv) Election: After observing this experiment and its results, candidates revise their beliefs, and hence their implemented policies if elected, while voters update their evaluation of the candidates' policies. Voters already know the valence of incumbent A. Before the election, voters also observe a noisy signal about the valence of untried candidate B. Each voter then chooses candidate A if she is expected to deliver a higher total payoff (valence+policy) than B.

We start our analysis in Section 3 by characterizing the informativeness of optimal experiments as a function of the competence η^A of incumbent A. In the benchmark model, we assume that the valence distribution of undried candidate B has a log-concave probability density function, such as a Normal Distribution. We show that the following single-crossing result holds for every policy experiment. If an experiment does not increase the victory probability of incumbent A when her valence is η^A, then this experiment does not increase her victory probability if her valence is higher than η^A (Lemma 1). This result implies that there are two cutoffs η_1^A and η_2^A in the extended real line such that the following holds. The IP finds it optimal to implement a fully informative experiment if $\eta^A < \eta_1^A$; a partially informative experiment if $\eta_1^A < \eta^A < \eta_2^A$; and a completely uninformative experiment if $\eta_2^A < \eta^A$ (Proposition 1 and Corollary 1). Consequently, the IP prefers to be completely transparent about policies when the majority candidate is sufficiently incompetent, partially transparent for intermediate levels of competence, and completely opaque when the majority candidate is sufficiently competent.

In Section 4 we focus on cases where political disagreement *endogenously* becomes a strictly increasing function of voters' expectation of the state, as in the tax example we described earlier. In this case, political disagreement increases if voters learn that the realized state is "high", which benefits the IP, and disagreement decreases if they learn that the state is "low". One could conjecture that the IP would always prefer to hide information about low disagreement states, and to fully reveal information about high states. Proposition 2 defines conditions such that the opposite is true: it is optimal to use an upper-censoring experiment. The experiment defines a cutoff state and works as follows. If the realized state is below the cutoff, then voters learn the true state. Otherwise, voters only learn that the state is above the cutoff. That is, the experiment fully reveals low disagreement states and pools high disagreement states. The intuition for the result is rooted on the impact of marginal increases in political disagreement on the marginal victory probability of majority candidate A. The marginal increase in the victory probability is small when disagreement is already high, and large when disagreement is low.

The above results fundamentally depend on the assumption that the p.d.f. of the challenger's valence distribution is log-concave. The results are reversed if this p.d.f. is log-convex. In the log-convex case, the single-crossing property goes in the opposite direction: lower values of incumbent's competence η^A induce less experimentation, while higher competence induces more experimentation. Moreover, if in the log-convex case we also have that disagreement is a function of the expected state, then we define conditions such that the optimal experiment is lower-censoring. The reason for the sharp change in results is rooted on the change in the curvature of the incumbent's victory probability, as a function of political disagreement and valence η^A. In the log-concave case, it is as if the IP features increasing absolute risk-aversion (IARA). When disagreement and incumbent's valence are low, the IP benefits from gambling on disagreement. That is, the IP benefits from implementing a risky experiment that might increase or decrease disagreement. When disagreement and valence are high, the IP prefers to avoid these gambles. In the log-convex case, it is as if the IP features DARA, and the reverse results hold.

In Section 5 we extend the basic model to allow for heterogeneous prior beliefs. We then restrict attention to cases in which voters share the same payoff function, so that political disagreement stems solely from belief disagreement. That is, in the absence of uncertainty, all voters would agree on the optimal policy and candidates would be judged solely on their valence. One may conjecture that if public information creates consensus among voters, then the IP will seldom benefit from persuasion and thus belief disagreement will foster opaqueness. We show that this view is flawed. For example, if political disagreement is increasing in the distance between each group's expected state, then the IP generically benefits from releasing some information, irrespective of the expected competence of the candidates (Proposition 3).

We relate our paper to the literature in Section 6. Section 7 extends the model and Section 8 concludes. All proofs are in Appendix A, and additional results are available on-line in Appendix B.

2 Model

Overview: There are two parties and two groups of voters. Party $\mathcal{A}$ represents voters in group A and party $\mathcal{B}$ represents voters in group B, where group A is larger than B. In our benchmark model, party $\mathcal{A}$ holds office at the beginning of the game (the opposite case is presented in Section 7.1). The incumbent party (IP) strategically designs a policy experiment to influence the next election. Voters observe the experiment's results and update their beliefs about policy-payoffs. Voters then observe a (possibly noisy) signal about the valence of untried candidate B — voters already know the valence of incumbent A. The election takes place, the elected candidate implements a policy, payoffs are realized and the game ends.

Voters' Preferences: Voters care about the policy choice and the valence (i.e. competence) of the elected official. If elected, the candidate has to choose one policy x from the compact, convex set $X \subset \mathbb{R}^d$. For example, X can represent the set of feasible governmental budget allocations across d projects, the government's policy on a left-right Downsian model, or a proportional income tax rate. Each citizen's payoff from policy x depends on an unknown state $\theta \in \Theta \equiv \{\theta_1, \ldots, \theta_N\}$, with a finite $N \geq 2$. To simplify presentation, let $\theta \in \mathbb{R}$ and $\theta_1 < \ldots < \theta_N$. Players share a common prior belief p in the interior of the simplex $\Delta(\Theta)$ — Section 5 extends the model to heterogenous prior beliefs. Citizens within each group are homogeneous, but groups differ in their policy preferences. Formally, each citizen in group $i \in \{A, B\}$ has preferences over policies characterized by the von Neumann-Morgenstern utility function $u^i(x, \theta)$, where u^i is a differentiable, strictly quasi-concave function of x. Each candidate is also endowed with a valence $\eta \in \mathbb{R}$, which we discuss momentarily. For a voter in group i, the total payoff from electing a politician with valence η who implements policy x when state θ is realized is

$$U^i(\eta, x, \theta) = \eta + u^i(x, \theta).$$

Political Parties: We model each party as a primarily office motivated institution (or similarly, party leaders and bureaucrats as purely office-motivated individuals), with ties to the policy interest of a particular group of voters. Formally, each party receives payoff one if its candidate is elected and zero otherwise. If elected, party $\mathcal{A}$ implements the policy that maximizes the expected payoff of voters in group A, while party $\mathcal{B}$ implements the best policy for voters in group B.[6] Consequently, the preferences of each party and the voters it represents are only partially aligned. Party $\mathcal{A}$ always strictly prefers to elect its own candidate, independently of policies and valences. However, given parties' policies, voters in group A prefer to elected the candidate from party $\mathcal{B}$ if he is sufficiently more competent than the candidate from party $\mathcal{A}$.

[6] We are implicitly assuming that, prior to the election, each party cannot commit to a specific policy. However, each party has long-run ties to the group of voters it represents and once in office will implement the policy favored by this group. For example, suppose group A consists of poor voters and B consists of rich voters. Although parties cannot commit to a particular tax rate, party $\mathcal{A}$ is expected to implement higher taxes than $\mathcal{B}$. The actual implemented tax will depend on the particular information available to parties about the state of the economy. See Section 4.1 for a formal optimal tax model.

Strategic Policy Experimentation: The IP controls the government and has the monopoly over a *policy experiment* (a public signal that is correlated with the state). By strategically designing this experiment the party can influence voters' beliefs and electoral outcomes. Formally, prior to the election the IP chooses a policy experiment π, consisting of a finite realization space S and a family of distributions over S, $\{\pi(\cdot|\theta)\}_{\theta\in\Theta}$, with $\pi(\cdot|\theta) \in \Delta(S)$. Experiment π is "commonly understood": π is observed by all players who agree on the likelihood functions $\pi(\cdot|\theta), \theta \in \Theta$. Players process information according to Bayes rule, so that $q(s|\pi, p)$ is the updated posterior belief of voters after observing realization s of π.[7] To simplify notation, we use q or $q(s)$ as a shorthand for $q(s|\pi, p)$.

Our learning technology follows important assumptions from KG: the IP has the monopoly over the experiment, it has no private information, it can choose any experiment that is correlated with the state, and experiments are costless to the IP. As in our model, Callander (2011) and Callander and Hummel (2014) consider a learning technology where the incumbent party has the monopoly over the policy experiment and has no private information. However, they consider a different learning technology, related to a Brownian process. In order to learn the incumbent must implement a new policy and all players (including the IP) incur the resulting policy payoff of this experiment. Thus, we interpret these as models of "full scale" policy experimentation. In our setup, we view the experiment as a small scale policy trial, that does not directly affect the payoff of the IP.[8]

Candidate's Policy: We refer to the candidate from party $\mathcal{A}$ and $\mathcal{B}$ as candidate A and B, respectively. There are no exogenous commitment devices available to politicians. However, since party filiation of candidates and the experiment's results are common knowledge, in equilibrium voters can correctly anticipate the policy that would be chosen by each candidate. If elected, candidate $i \in \{A, B\}$ will implement policy $x^{i*}(q) \equiv \arg\max_{x\in X} \sum_{\theta\in\Theta} q_\theta u^i(x, \theta)$. We refer to $x^{i*}(q)$ as the "preferred policy" of candidate i.

Candidate's Valence: Besides this policy dimension, candidates also differ in a valence

[7] We use "posterior belief" to indicate the players' belief about the state after observing s but before voting.

[8] For example, to study the effects of public schools policies, the IP can run a policy experiment on a few public schools, and choose which questions will be addressed and evaluated during the experiment. In Section 7.2 we consider costs that increase in the experiment's informational content.

dimension. All players already know the valence η^A of incumbent A, since they observe his performance in office. After the IP chooses its experiment, but before the election, voters observe valence η^B of untried candidate B. Our timing assumption is rooted on the fact that it takes time to setup and implement policy experiments, while the identity (and hence the actual valence) of the challenger is only defined much closer to the election. Hence, at the time the IP chooses an experiment, there is significant uncertainty over the valence of the next challenger.

We assume that η^B is a random variable distributed according to the twice differentiable cumulative distribution function F, with probability density function f. Throughout the paper we maintain the following assumption:

(A1) η^B *has full support on the real numbers and follows a log-concave probability measure.*

Note that η^B follows a log-concave probability measure if and only if its density function f is log-concave (see Prékopa 1971). Condition **(A1)** holds, for example, for the normal, logistic, and extreme value distributions. See Bagnoli and Bergstrom (2005) for a discussion on the properties of log-concave density functions. In Section 6.1 we discuss the case in which f is log-convex.

The model is easily extended to the case in which the incumbent politician is not running for reelection. The incumbent party $\mathcal{A}$ then runs with an untried candidate and voters simultaneously observe valences η^A and η^B of the untried candidates. Although we say that voters observe the "true" valences of candidates, the model can be easily reinterpreted as voters observe a noisy, exogenous signal about the valence of each candidate (e.g., information from media coverage during the campaign). In this case, variables η^A and η^B are interpreted as the new expected valence of each candidate, after voters observe the implicit realization of the signals about valence.[9] See Boleslavsky and Cotton 2015 for a model of noisy information about valence.

Election: At the time of the election, voters can predict candidates' policies $x^{*A}(q)$ and $x^{*B}(q)$. Voters also observe the realized valences η^A and η^B. Thus, for a citizen in group i,

[9] Defining the new random variable $\xi \equiv \eta^B - \eta^A$, our assumption **(A1)** refers to the distribution of ξ. In this case, our results on changes in η^A would then refer to location shifts of the distribution of ξ.

the total expected payoff of electing candidate j is

$$\mathcal{U}^{ij}(q, \eta^A, \eta^B) = \eta^j + \sum_{\theta\in\Theta} q_\theta u^i(x^{*j}(q), \theta). \tag{1}$$

To rule out uninteresting equilibria, we eliminate weakly dominated voting strategies. This implies that each voter votes for the candidate who delivers to him the highest expected utility[10]. The candidate who wins the majority of votes is elected and then implements his preferred policy. Voters in group A are decisive, since the group encompasses a majority of voters. That is, a candidate wins if and only if he receives the support of the majority group.

2.1 Political Disagreement

The previous discussion implies that a voter from group i votes for the candidate from group A if and only if[11]

$$\begin{aligned} \mathcal{U}^{iA}(q, \eta^A, \eta^B) &\geq \mathcal{U}^{iB}(q, \eta^A, \eta^B) \\ \iff \sum_{\theta\in\Theta} q_\theta \left[u^i(x^{*A}(q), \theta) - u^i(x^{*B}(q), \theta)\right] &\geq -(\eta^A - \eta^B). \end{aligned} \tag{2}$$

The RHS of (2) captures the realized valence differential. The LHS of (2) captures the *degree of political disagreement* between the two groups. That is, it captures, from the point of view of a voter in group i, the expected policy-payoff difference from electing the different candidates. Define the political disagreement from the point of view of voters in group A as

$$D(q) \equiv \sum_{\theta\in\Theta} q_\theta \left[u^A(x^{*A}(q), \theta) - u^A(x^{*B}(q), \theta)\right]. \tag{3}$$

Majority group A is decisive: after a signal realization that induces q, candidate A wins the election if and only if he receives the support of voters in group A,

$$D(q) \geq -\eta^A + \eta^B.$$

[10] In our model voters have no private information about the state, so there is no information aggregation problem. Hence, the strategic voting considerations related to the probability of being pivotal are not relevant in our setup.

[11] We abstract from abstentions. One could extend our model so that a citizen is less likely to abstain if his expected payoff difference between the candidates is higher, similar to Matsusaka (1995).

If the realized η^B is sufficiently high, then even voters from group A vote for candidate B, and vice-versa. Since $\eta^B \sim F$, given η^A, the majority candidate wins with probability

$$v(q; \eta^A) \equiv F(D(q) + \eta^A). \tag{4}$$

Therefore, candidate A wins the election with a probability that increases in the degree of political disagreement — candidate A has a “policy advantage” because a majority of voters believe she has the “correct” preference, and hence he will implement the “correct” policy.

In order to guarantee the existence of an optimal experiment and simplify notation, throughout the paper we maintain the following assumption:

(A2) *Political disagreement D is upper semicontinuous in $\Delta(\Theta)$, and differentiable at the prior belief.*

Condition **(A2)** holds for a large class of models, including the applications we study throughout this paper. Differentiability of F and **(A2)** imply that v is upper semicontinuous in $\Delta(\Theta)$, and differentiable at the prior belief.

2.2 Notational Conventions

For vectors $q, w \in \mathbb{R}^J$, we denote by $\langle q, w \rangle$ the standard inner product in $\mathbb{R}^J$, i.e. $\langle q, w \rangle = \sum_{j=1}^{J} q_j w_j$, and we denote by qw the component-wise product of vectors q and w, i.e. $(qw)_j = q_j w_j$.

For an arbitrary real-valued function g define $\widetilde{g}$ as the concave closure of g,

$$\widetilde{g}(q) = \sup \{y | (q, y) \in co(g)\},$$

where $co(g)$ is the convex hull of the graph of g.

We use $\pi \succ \pi'$ to denote that experiment π is Blackwell more informative than experiment π'. Finally, $card(S)$ denotes the cardinality of the set S.

2.3 Party's Expected Payoff

The incumbent party's problem is to choose an experiment π that maximizes, from her point of view, the expected victory probability $E_\pi[v(q; \eta^A)]$. Upper semicontinuity of v ensures the existence of an optimal experiment, and choosing an optimal experiment is equivalent to choosing a probability distribution σ over q that maximizes $E_\pi[v(q;\eta^A)]$, subject to the constraint $E_\sigma[q] = p$ (see KG). That is, the supremum of the expected victory probability is

$$V = \sup_\sigma E_\sigma[v(q;\eta^A)], \qquad \text{s.t. } E_\sigma[q] = p.$$

The following remarks follow immediately from KG:

(R1) An optimal experiment exists;

(R2) There exists an optimal experiment with $card(S) \leq N$;[12]

(R3) The IP's maximum expected payoff is

$$V = \widetilde{v}(p;\eta^A); \tag{5}$$

(R4) The value of persuasion is

$$V - v(p;\eta^A) = \tilde{v}(p;\eta^A) - v(p;\eta^A). \tag{6}$$

2.4 Application: Spatial Policy Model

Although we prove our main results using the general setup described above, for concreteness throughout the paper we illustrate our results using the following application.

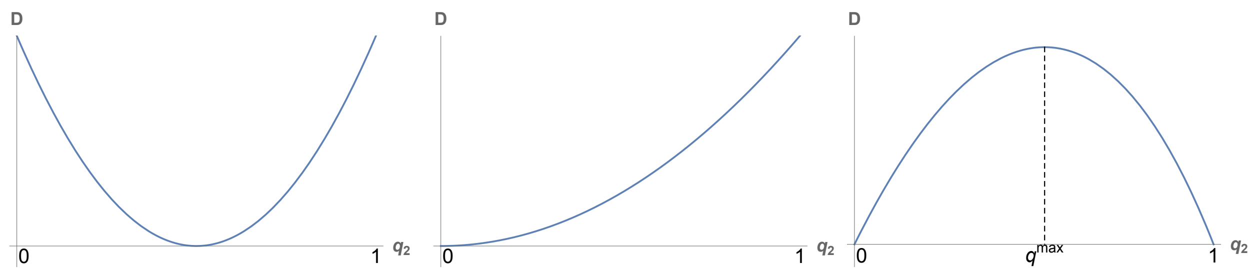
Consider a spatial policy model where the state $\theta \in \Theta \subset \mathbb{R}$ captures voters' uncertainty over the optimal policy in a left-right dimension. Let $\theta_1 < \cdots < \theta_N$ and $X = [-\overline{x}, +\overline{x}]$, with $\overline{x}$ sufficiently large. Voters $i \in \{A, B\}$ have a quadratic policy payoff $u^i(x,\theta) = -(x - \beta^i\theta)^2$, where $\beta^A \neq \beta^B$ are preference parameters. From the point of view of voter i with belief q, the optimal policy is linear on the expected value of the state, $x^{i*}(q) = \beta^i E[\theta|q]$. Political

[12]Note that in the original setup of KG, there exists an optimal straightforward signal that directly recommends an action to the receiver. In our setup, the pivotal majority voter has a binary action space: vote for candidate A or B. However, when $N > 2$ in our model, an optimal experiment might require more than two realizations. This is so because from the point of view of the IP, before the valence shock is realized, the voting behavior is probabilistic rather than binary. That is, voting behavior can be interpreted ex ante as a continuous "action" (probability of electing A) in the interval $[0, 1]$ rather than a binary choice.

disagreement (3) is

$$
\begin{aligned}
D(q) &= \sum_{\theta\in\Theta} q_\theta \left[u^A(x^{*A}(q),\theta) - u^A(x^{*B}(q),\theta)\right] \\
&= \sum_{\theta\in\Theta} q_\theta \left[-(\beta^A E[\theta|q] - \beta^A\theta)^2 + (\beta^B E[\theta|q] - \beta^A\theta)^2\right] \\
&= (\beta^A E[\theta|q] - \beta^B E[\theta|q])^2 \qquad (7) \\
&= (\beta^A - \beta^B)^2 E[\theta|q]^2. \qquad (8)
\end{aligned}
$$

From (7), political disagreement translates naturally into the degree of disagreement over optimal policies, $D(q) = (x^{*A}(q) - x^{*B}(q))^2$. From (8), disagreement is zero if $E[\theta|q] = 0$ and strictly positive otherwise. Figure 1(a) uses a binary-state example to illustrate D, with $\Theta = \{-1, +1\}$. In the graph, q_2 corresponds to the probability that the state is $\theta = +1$. In Figure 1(b) we change the state space to $\Theta = \{0, +1\}$.



(a) Spatial Model, $\Theta = \{-1, +1\}$ (b) Spatial Model, $\Theta = \{0, +1\}$ (c) Budget Allocation Model

Figure 1: Political disagreement as a function of belief, with $\Theta = \{\theta_1, \theta_2\}$, $q_2 = Pr(\theta = \theta_2)$.

In this spatial policy model political disagreement D is always a convex function of belief q. Therefore, any informative experiment π increases the expected political disagreement, $E_\pi[D(q)] \geq D(p)$. However, in Section 5.3 we present a budget allocation model where political disagreement is concave, hence any informative experiment decreases the expected political disagreement. Figure 1(c) illustrates disagreement in that model. To streamline the paper, we postpone the formal presentation of that application to Section 5.3.

3 Valence and Information

In this section we first show that the incumbent party's gain from *any given experiment* π has a single crossing property with respect to the incumbent's valence. This property leads to a monotone behavior of the informativeness of optimal experiments: as we increase the incumbent's competence η^A, his party does not benefit from providing a more informative experiment.

3.1 Single-Crossing

In our model, the incumbent party seeks to maximize the chances of reelection of its candidate. Following (4), the likelihood that candidate A wins the election increases in the degree of political disagreement — a larger D implies that, in the eyes of group A voters, the minority candidate B is expected to implement a much "worse policy" than A. As the outcome of the experiment can change the policy championed by each candidate, as well as voters' expected payoff from these policies, it follows that policy experimentation can change the degree of political disagreement. As a result, the IP's choice of an experiment is driven by her desire to uncover information that increases political disagreement.

As the underlying state θ is independent of the valences of both candidates, the IP's choice of experiment cannot affect the distribution of the challenger's valence. Nevertheless, if the IP has access to an experiment that on average increases disagreement, as in the example in Figure 1(b), it is then not clear why she would not gain from this experiment independently of η^A. The next lemma shows that for any experiment π this gain actually satisfies a single-crossing condition: If the IP prefers not to experiment rather than provide experiment π when its candidate's valence is η^A, the IP continues to find no experimentation better than experiment π for any higher valence $\eta^{A\prime} > \eta^A$. Conversely, if when the incumbent's valence is η^A the IP finds that providing experiment π is better than no information disclosure, then this experiment continues to be better than no information for every lower valence.

Lemma 1 *Consider any experiment π and incumbent's valence η^A.*
(i) If $E_\pi[v(q;\eta^A)] \leq v(p;\eta^A)$, then $E_\pi[v(q;\eta^{A\prime})] \leq v(p;\eta^{A\prime})$ for all $\eta^{A\prime} > \eta^A$.
(ii) If $E_\pi[v(q;\eta^A)] \geq v(p;\eta^A)$, then $E_\pi[v(q;\eta^{A\prime})] \geq v(p;\eta^{A\prime})$ for all $\eta^{A\prime} < \eta^A$.

To understand Lemma 1, note that the effect of changing disagreement by an amount Δ is that it changes victory probability by $F(x+\Delta) - F(x)$, with $x = D(p) + \eta^A$. Thus, if $\Delta > 0$, the benefit in increasing victory probability relative to the likelihood that the challenger's valence is x, is given by

$$\frac{F(x+\Delta)-F(x)}{f(x)}=\int_0^{\Delta}\frac{f(x+s)}{f(x)}ds,$$

while, if $\Delta<0$, the cost of decreasing victory probability relative to $f(x)$ is

$$\frac{F(x)-F(x+\Delta)}{f(x)}=\int_0^{|\Delta|}\frac{f(x-s)}{f(x)}ds.$$

Lemma 1 then follows from the fact that, for log-concave probability measures, the ratio $f(x+\Delta)/f(x)$ decreases in x if $\Delta>0$, but increases in x if $\Delta<0$. That is, the relative benefit of increasing victory probability decreases in x — hence, in the incumbent's competence η^A — while the relative cost increases in x. Integrating over all possible realizations of Δ generated by experiment π, we then have that the relative gain from an experiment π (weakly) decreases in the incumbent's competence. In other words, if the IP does not gain from experiment π when the incumbent's valence is η^A, this is still true for an incumbent candidate of higher valence. Notice, in particular, that this property is satisfied irrespective of whether the incumbent is expected to win the election ($F(x)>1/2$) or the minority candidate is the frontrunner ($F(x)<1/2$) in the absence of the IP's experiment.

The next Proposition builds upon Lemma 1 to show that, if we increase the competence of the majority candidate, then the IP does not benefit from providing a more informative experiment.

Proposition 1 *Suppose π^* is an optimal experiment given incumbent's valence η^A. Then for any higher valence, experiment π^* is weakly better than any Blackwell more informative experiment. That is, for every $\eta^{A\prime}>\eta^A$ and every $\pi'\succ\pi^*$ we have*

$$E_{\pi^*}[v(q;\eta^{A\prime})]\geq E_{\pi'}[v(q;\eta^{A\prime})].$$

In the proof of the Proposition, we first rewrite the Blackwell more informative experiment π' as a payoff equivalent grand experiment. In this grand experiment, voters first observe realization s of π^*, then they observe an additional experiment π_s conditional on s. When the incumbent's valence is η^A, optimality of π^* implies that the IP does not benefit from disclosing any additional information π_s after each realization s of π^*. We then apply Lemma 1(i) to each posterior belief q^* in the support of π^*: if the IP does not benefit from disclosing information in addition to π^* when the incumbent's valence is η^A, then the IP does not benefit from disclosing any information in addition to π^* when the incumbent's valence is higher.

Next we apply Proposition 1 to characterize the relationship between the IP's optimal level of transparency and the incumbent's valence.

Corollary 1 *There are cutoffs η_1^A and η_2^A in the extended real line, with $\eta_1^A \leq \eta_2^A$, such that:*
(i) a fully informative experiment is optimal if $\eta^A < \eta_1^A$;
(ii) a partially informative experiment is optimal if $\eta_1^A < \eta^A < \eta_2^A$;
(iii) an uninformative experiment is optimal if $\eta_2^A < \eta^A$.

Corollary 1 defines partitions on the expected competence of the majority candidate. The incumbent party prefers to be completely transparent about policies, and engages in fully informative experimentation, when her candidate is sufficiently incompetent, partially transparent for intermediate levels of competence, and completely opaque, and foregoes experimentation, when her candidate is sufficiently competent.

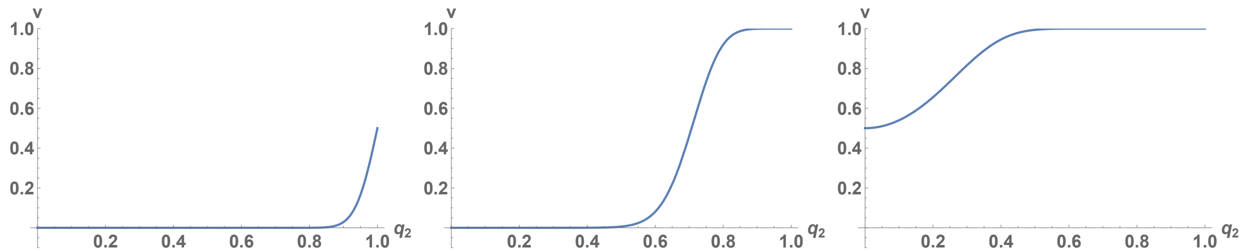
Corollary 1 does not guarantee that cutoffs η_1^A and η_2^A are finite. For example, in Figure 1(a) a fully informative experiment is always optimal, $\eta_1^A = \eta_2^A = +\infty$. In Figure 1(c), if the prior belief is $p_2 = q^{max}$, then no information disclosure is always optimal, $\eta_1^A = \eta_2^A = -\infty$. Proposition B.1 in the on-line Appendix B provides sufficient conditions so that η_1^A or η_2^A are finite.

3.2 Examples

We next provide some examples to illustrate the effects of η^A on the IP's payoff function v and on the optimal experiment.

Recall that $v(q; \eta^A) = F(D(q) + \eta^A)$. Figure 2 illustrates how a higher η^A increases v for each q and changes the overall curvature of v. It assumes F follows a Normal Distribution and uses the political disagreement D from the spatial policy model in Figure 1(b).

Recall that we can derive the optimal experiment from the concave closure of v (see KG for details). In particular, whether v is concave or convex is important to define whether or not the IP benefits from implementing an informative experiment. Although in this example disagreement D is strictly convex, the resulting payoff v might be locally concave or

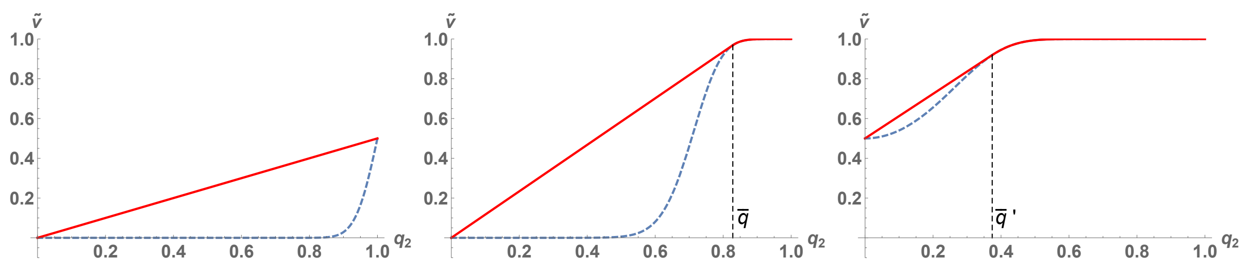


(a) Low Values of η^A (b) Intermediate Values of η^A (c) High Values of η^A

Figure 2: Effects of η^A on victory probability v, using disagreement D from Figure 1(b)

locally convex, depending on belief q_2 and on valence η^A. Log-concavity of F implies that $F(D(q) + \eta^A)$ is locally concave for sufficiently high values of $D(q) + \eta^A$, and locally convex for sufficiently low values. The red solid lines in Figure 3 depict the concave closure of v. We next use Figure 3 to derive an optimal experiment.

First, suppose η^A is sufficiently low, as in Figure 3(a). The IP's payoff v is everywhere strictly convex, hence any optimal experiment must be fully informative, independently of the prior belief.



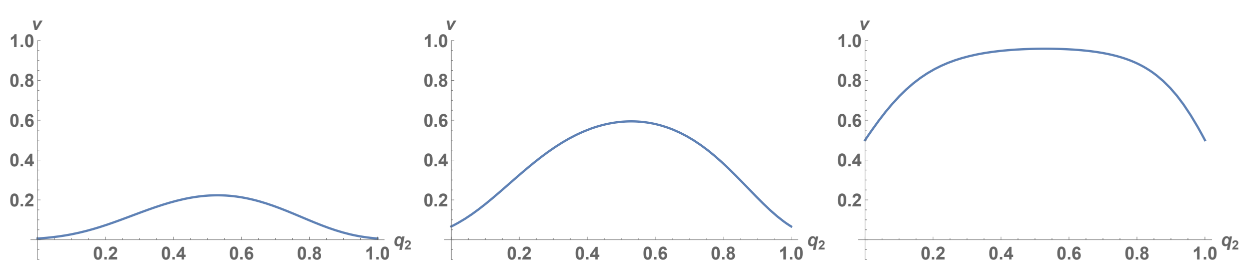
(a) Low Values of η^A (b) Intermediate Values of η^A (c) High Values of η^A

Figure 3: Concave closure of v from Figure 2

Now suppose η^A is intermediate, as in Figure 3(b). The concave closure $\tilde{v}$ is given by a straight line in the set of beliefs $q_2 \leq \bar{q}$, and by v itself for $q_2 \geq \bar{q}$. Consequently, no experimentation is optimal for all priors $p_2 \geq \bar{q}$. When $p_2 \geq \bar{q}$, although any informative experiment increases average disagreement (D is strictly convex), any informative experiment is strictly worse for the IP than no information disclosure. Signal realizations that increase *political disagreement* only increase the *victory probability* by a small amount, while signal realizations that decrease political disagreement decrease the victory probability by a relatively large amount. Now suppose $p_2 \leq \bar{q}$. Since in this set $\tilde{v}(q; \eta^A) > v(q; \eta^A)$, policy experimentation is valuable. Every optimal experiment is partially informative and induces exactly two posterior beliefs, $q_2 = 0$ and $q_2 = \bar{q}$. Finally, for each prior belief $p_2 \in (0, 1)$, optimal experiments are less informative in Figure 3(b) than in Figure 3(a).

As we further increase η^A the cutoff $\bar{q}$ decreases to $\bar{q}'$, see Figure 3(c). Therefore, no experimentation is optimal for a larger set of prior beliefs. Moreover, for the prior beliefs in the set $p_2 \leq \bar{q}'$, every optimal experiment is only supported on the posterior beliefs $q_2 = 0$ and $q_2 = \bar{q}'$. Consequently, the partially informative experiment in Figure 3(c) is less informative than the partially informative experiment in Figure 3(b).

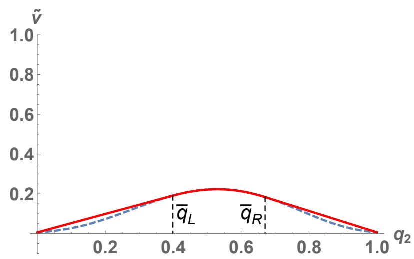
What if political disagreement is concave? Consider the political disagreement D from Figure 1(c). Figure 4 illustrates how increasing η^A can change the overall curvature of v, assuming F follows a Normal Distribution. Figure 5 depicts the concave closure of v for different values of η^A. Consider the case where η^A is low, given by Figure 5(a). The concave closure $\tilde{v}$ is a straight line in the set of beliefs $q_2 \leq \bar{q}_L$, it is v itself in $q_2 \in [\bar{q}_L, \bar{q}_R]$, and it is a straight line in the set $q_2 \geq \bar{q}_R$. Consequently, if the prior belief is in the set $p_2 \leq \bar{q}$, then the optimal experiment induces posterior beliefs $q_2 = 0$ and $q_2 = \bar{q}_L$. If the prior belief is sufficiently close to the maximum feasible disagreement, $p_2 \in [\bar{q}_L, \bar{q}_R]$, then no experimentation is optimal. If the prior belief is in the set $p_2 \geq \bar{q}_R$, then the optimal experiment induces posterior beliefs $q_2 = \bar{q}_R$ and $q_2 = 1$. Optimal experiments become less informative as we increase η^A. In Figure 5(c), no information disclosure is optimal for all prior beliefs. Finally, note that a fully informative experiment is never optimal, since it minimizes disagreement with certainty.



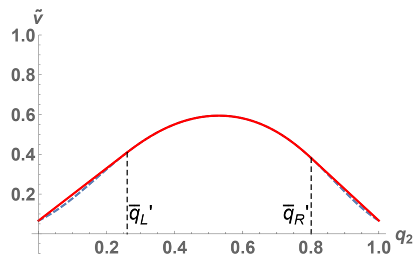
(a) Low Values of η^A (b) Intermediate Values of η^A (c) High Values of η^A

Figure 4: Effects of η^A on victory probability v, using disagreement D from Figure 1(c)

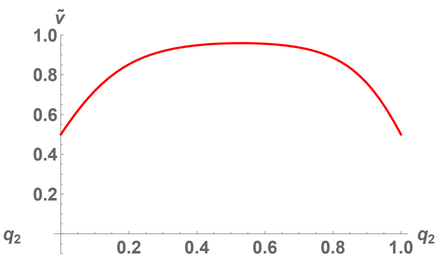
We conclude by highlighting that the incumbent party might find it optimal to experiment even when her candidate is the frontrunner, and might find it optimal not to experiment



(a) Low Values of η^A



(b) Intermediate Values of η^A



(c) High Values of η^A

Figure 5: Concave closure of v from Figure 4

even when her candidate is the underdog. For example, in Figure 3(b), if the prior belief is $p_2 = 0.8$, then without experimentation the majority candidate wins with a very high probability, above 90%. Nevertheless, it is optimal for the incumbent party to provide a partially informative experiment, because it increases even further her candidate's expected victory probability. In Figure 5(a), if the prior belief is $p_2 = 0.45$, then without information disclosure the majority candidate wins with a very low probability, around 26%. Nevertheless, *any* informative experiment decreases even further the candidate's expected victory probability.

4 Disagreement as a Function of the Expected State

To derive a sharper characterization of optimal experiments, in this section we focus on models where political disagreement is a strictly increasing function of the expected value of some unknown state. Formally, we assume:

(A2′) *Political disagreement takes the form $D(q) = H(E[\theta|q])$, where H is twice differentiable and strictly increasing, with a log-concave derivative h.*

Assumption **(A2′)** holds in many important cases. For example, it holds if disagreement is a power function of expectation $D(q) = \gamma E[\theta|q]^\rho$, with $\gamma > 0$, $\theta_1 \geq 0$ and $\rho \geq 1$. The spatial policy model of Section 2.4 satisfies **(A2′)** when $\theta_1 \geq 0$, since disagreement (8) is proportional to the square of the expectation of the state. Later in this section we study two other relevant applications in which **(A2′)** holds (optimal tax and the relative importance of policy dimensions).

Given **(A2′)**, political disagreement increases if voters learn that the realized state is "high", which benefits the incumbent, and disagreement decreases if they learn that the state is "low". One could then conjecture that the incumbent party would prefer to hide information about low disagreement states, and to fully disclose information about high states. However, Proposition 2 shows that the opposite is true. Borrowing from the statistics literature, we define an upper-censoring experiment (or right-censoring experiment) as an experiment that fully reveals low disagreement states ($\theta < \theta_k$) and pools high disagreement states. Formally,

Definition: Experiment π is *upper-censoring* at cutoff state θ_k if it has a realization space $S = \{s_1, \ldots, s_k, s_{\text{pooling}}\}$ and the following holds. For each $n < k$, state θ_n induces signal realization s_n with probability one. For each $n > k$, state θ_n induces signal realization s_{pooling} with probability one. State θ_k induces realization s_{pooling} with some probability $\alpha_k \in [0, 1]$ and induces realization s_k with probability $1 - \alpha_k$.

Proposition 2 *Suppose **(A2′)** holds. Then there exists an optimal experiment π^* that is upper-censoring at some cutoff state θ_k. Moreover, cutoff state θ_k weakly decreases with incumbent's valence η^A.*

Assumption **(A2′)** holds in many important cases. For example, it holds if disagreement is a power function of expectation $D(q) = \gamma E[\theta|q]^\rho$, with $\gamma > 0$ and $\rho \geq 1$. The spatial policy model of Section 2.4 satisfies **(A2′)** when $\theta_1 \geq 0$, since disagreement (8) is proportional to the square of the expectation of the state. Later in this section we study two other relevant applications in which **(A2′)** holds (optimal tax and the relative importance of policy dimensions).

In the proof of Proposition 2 we show that for each optimal experiment π^* there exists a payoff equivalent upper-censoring experiment. The intuition behind the result is the following. Under our assumptions, given η^A, the IP's payoff $v(q; \eta^A) = F(H(E[\theta|q]) + \eta^A)$ is concave if $E[\theta|q]$ is high and strictly convex if $E[\theta|q]$ is low. Strict convexity implies that the incumbent party always strictly benefits from providing additional information if the initial experiment yielded a non-degenerate belief corresponding to a low expected state. Therefore, outcomes under optimal experiments that indicate the state to be low must be fully revealing. Conversely, concavity of the incumbent's payoffs implies that she cannot be made worse off by an experiment that pools all outcomes corresponding to high expected states into a single realization. That is, the incumbent then (weakly) gains from bundling all states in the concave (high disagreement) region: they all induce signal s_{pooling} with probability one, resulting in a single posterior belief q^+ and a high expectation $E[\theta|q^+]$.

While the IP does not gain from designing an experiment that pools together only states in the convex region, she may gain from "hiding" some low disagreement states, such that these states induce signal s_{pooling} with positive probability. Of course, pooling low disagreement states would make s_{pooling} more likely but would reduce expected disagreement if s_{pooling} occurs. Still, the incumbent must decide which disagreement states should be pulled in s_{pooling}. Suppose θ_l and θ_h are in the convex region, with $\theta_l < \theta_h$. Should θ_l or θ_h be the incumbent's first choice to be mixed with the high disagreement signal s_{pooling}? She now faces an important trade off. In one hand, pooling θ_h leads to a lower reduction in posterior disagreement resulting from s_{pooling}. In another hand, disclosing θ_l to voters is worse than disclosing θ_h, hence "hiding" θ_l by pooling it with s_{pooling} is more important than hiding θ_h. The proof of Proposition 2 shows that, given **(A1)** and **(A2′)**, the first effect always dominates: the IP's optimal decision must be a cutoff on θ, independently of prior beliefs, incumbent's valence, and the other parameters of the model -these values are only relevant to define the actual cutoff state.

Finally, the cutoff state defined by Proposition 2 monotonically decreases with incumbent's valence η^A. This implies that the set of optimal upper-censoring experiments that we construct are Blackwell ordered: they become less Blackwell-informative as the majority candidate becomes more competent.

We next present two additional applications of our model, where political disagreement endogenously becomes a function of the expected state and Proposition 2 applies.

4.1 Application: Optimal Tax

Consider the following model where the elected politician must choose a proportional income tax $x \in [0, 1]$. Voters care about the consumption of a private good and a public good. Each voter in group $i \in \{A, B\}$ is endowed with income $\beta^i > 0$, where $\beta^A \neq \beta^B$. Given the implemented tax rate x, voter i consumes $(1-x)\beta^i$ units of the private good. The Government uses all tax revenues to produce the public good. The production technology is such that the Government produces x^ψ units of the public good, where $\psi \in (0,1)$ is a known technology parameter.[13] Voters' policy payoff is $u^i(x,\theta) = (1-x)\beta^i + \theta x^\psi$, where state θ represents the unknown marginal value of the public good, with $0 \leq \theta_1 < \ldots < \theta_N < \max\{\beta^A, \beta^B\}\frac{1}{\psi}$. Given belief q, the optimal tax rate of voter i is $x^{i*}(q) = \left(\frac{\psi E[\theta|q]}{\beta^i}\right)^{\frac{1}{1-\psi}}$. Both groups want higher taxes if the marginal value of the public good is higher. However, voters agree on the optimal tax if and only if the expected marginal value of the public good is zero. Political disagreement increases with $E[\theta|q]$:

$$\begin{aligned} D(q) &= (1-x^{A*}(q))\beta^A + E[\theta|q](x^{A*}(q))^\psi - (1-x^{B*}(q))\beta^A - E[\theta|q](x^{B*}(q))^\psi \\ &= \gamma E[\theta|q]^\rho, \end{aligned}$$

where $\gamma \equiv \psi^{\frac{\psi}{1-\psi}}\left\{(1-\psi)(\beta^A)^{\frac{-\psi}{1-\psi}} + \beta^A\psi(\beta^B)^{\frac{-1}{1-\psi}} - (\beta^B)^{\frac{-\psi}{1-\psi}}\right\} > 0$ and $\rho \equiv \frac{1}{1-\psi} > 1$.

Independently of whether the majority group is richer or poorer than the minority group (β^A is higher or lower than β^B), disagreement is a power function of the expected state and satisfies the conditions of Proposition 2. To maximize the victory probability of the majority group, the IP's optimal experiment either partially reveals that the public good is sufficiently important, or fully reveals that the public good has a low marginal value.

4.2 Application: The Relative Importance of Policy Dimensions

There is a single policy-issue in the spatial policy model of Section 2.4 and in the optimal tax model of Section 4.1. Moreover, information about θ induces politicians to reevaluate their beliefs and choose a new policy. In other important cases, the policy-issue is multidimensional and voters and politicians are convinced about what is the optimal policy, but are uncertain about the relative importance of different policy dimensions.

To study these cases, consider the following model. There are $d \geq 2$ policy dimensions (e.g., public education, public health, national defense, etc.). A policy is a d-dimensional vector $x = (x_1, \ldots, x_d)$. The preferences of voter $i \in \{A, B\}$ are captured by the preference

[13] Without loss of generality, let $m^A\beta^A + m^B\beta^B = 1$ where m^i is the number of voters in group i. Total tax revenue is then x.

vector $\beta^i = (\beta^i_1, \ldots, \beta^i_d)$ and by the loss function l, with $l(0) = 0$ and $l' > 0$. Voter's policy payoff is

$$u^i(x, \theta) = \sum_{j=1}^{d} -\lambda_j(\theta) l(|x_j - \beta^i_j|),$$

where each function $\lambda_j(\theta)$ captures the relative importance of policy-dimension j given state θ, with $\sum_{j=1}^{d} \lambda_j(\theta) = 1$ and $\lambda_j(\theta) > 0$. Note that voters' preferred policies are independent of beliefs about θ, $x^{i*}_j(q) = \beta^i_j$. Although voters know what is their preferred policy in education and national defense, they are uncertain about which policy issue will be more important during the next term.

The degree of political disagreement, from the point of view of voters in group A, is simply the expected (weighted) loss from the policy of candidate B,

$$\begin{aligned} D(q) &= E\left[\sum_{j=1}^{d} -\lambda_j(\theta) l(|\beta^A_j - \beta^A_j|)] - \sum_{j=1}^{d} -\lambda_j(\theta) l(|\beta^B_j - \beta^A_j|)\right] \\ &= E\left[\sum_{j=1}^{d} \lambda_j(\theta) l(|\beta^B_j - \beta^A_j|)\right]. \end{aligned}$$

To apply Proposition 2, rewrite the unknown state as follows. For each $\theta \in \Theta$ compute $\theta' \equiv \sum_{j=1}^{d} \lambda_j(\theta) l(|\beta^B_j - \beta^A_j|)$. Define a new state space Θ' as the collection of θ'. We can then rewrite disagreement simply as the expected value of θ' and apply Proposition 2.

In summary, voters have a fundamental disagreement over the optimal policy, but are uncertain about how important each policy dimension will be. For instance, suppose there are only two issues. Voters disagree relatively more on national defense and less on education (that is, $|\beta^B_j - \beta^A_j|$ is larger for the national defense dimension). The incumbent's optimal experiment pools together states that attached more weight to national defense, and fully reveals states that attach more weight to education. That is, the optimal experiment either reveals that the controversial national defense issue will be "sufficiently important" in the upcoming years, or fully reveals that the more agreeable education issue will be more important.

5 The Role of Belief Disagreement

Heterogeneous prior beliefs play an important role in politics — see Millner, Ollivier, and Simon (2014) for a recent review of the literature on heterogeneous priors in politics. In this Section we extend our model to the case where voters have heterogeneous prior beliefs about the state, and show that the results from Section 3 continue to hold. We then focus on cases where all voters share the same preferences, so that political disagreement is zero when voters share a common belief. That is, voters would agree on the optimal policy if they all had the same beliefs about the state. We define conditions such that, even though voters have the same payoff function, the incumbent can design an experiment that increases political disagreement with probability one, independently of the incumbent's valence.

5.1 Heterogeneous Prior Beliefs

Consider the general model from Section 2. Assume that voters in the same group share a common prior belief, but voters in opposite groups openly disagree over the likelihood of state θ. That is, voters in group i have a common prior belief $p^i = (p^i_1, \ldots, p^i_N)$ in the interior of the simplex $\Delta(\Theta)$, but prior beliefs differ across groups, $p^A \neq p^B$. To simplify presentation, assume that each party shares the beliefs of its affiliates. Preferences and prior beliefs are common knowledge — voters "agree to disagree." If we interpret θ as describing the mapping between policy x and outcomes, then different prior beliefs represent differences in voters' views of which outcomes are produced by the different government policies.

Given priors p^A and p^B, policy experiment π and signal realization s, let q^A and q^B be the respective posterior beliefs of each group. Since party filiation of candidates, prior beliefs and the experiment are common knowledge, voters can correctly infer the policy each candidate would implement if elected, $x^{i*}(q^i) \equiv \arg\max_{x \in X} \sum_{\theta \in \Theta} q^i_\theta u^i(x, \theta)$. So we can rewrite (1), the expected payoff of voter i if candidate j wins, as

$$\mathcal{U}^{ij}(q^A, q^B, \eta^A, \eta^B) = \eta^j + \sum_{\theta \in \Theta} q^i_\theta u^i(x^{*j}(q^j), \theta).$$

A voter from group i votes for the candidate from group A if and only if

$$\begin{aligned} \mathcal{U}^{iA}(q^A, q^B, \eta^A, \eta^B) \quad &\geq \quad \mathcal{U}^{iB}(q^A, q^B, \eta^A, \eta^B) \\ \iff \sum_{\theta \in \Theta} q^i_\theta \left[u^i(x^{*A}(q^A), \theta) - u^i(x^{*B}(q^B), \theta) \right] \quad &\geq \quad -(\eta^A - \eta^B). \end{aligned} \tag{9}$$

The LHS of (9) captures the degree of political disagreement. Disagreement from the point of view of voters in group A is

$$\mathcal{D}(q^A, q^B) \equiv \sum_{\theta \in \Theta} q^A_\theta \left[u^A(x^{*A}(q^A), \theta) - u^A(x^{*B}(q^B), \theta) \right]. \tag{10}$$

Since group A forms a majority, they are decisive: after a signal realization that induces posterior beliefs q^A and q^B, candidate A wins the election if and only if

$$\mathcal{D}(q^A, q^B) \geq -\eta^A + \eta^B.$$

Candidate A wins the election with a probability that increases in the degree of political disagreement — candidate A has a "policy advantage" because a majority of voters believe he has not only the "correct" preference, but also the "correct" belief, and hence he will implement the "correct" policy.

We now rewrite $\mathcal{D}$. Let $r_\theta \equiv \frac{p^B_\theta}{p^A_\theta}$ and $r \equiv \{r_\theta\}_{\theta \in \Theta}$ capture the likelihood ratio of prior beliefs. Alonso and Câmara (2015a, Proposition 1) show that independently of the experiment π and its realization s, we can rewrite q^B solely as a function of the belief of voters in group A,

$$q^B_\theta = \frac{q^A_\theta r_\theta}{\langle q^A, r \rangle}. \tag{11}$$

Therefore, we can express $\mathcal{D}(q^A, q^B)$ as a function of q^A only,

$$D(q^A) \equiv \mathcal{D}\left(q^A, q^A \frac{r}{\langle q^A, r \rangle} \right). \tag{12}$$

Since $\eta^B \sim F$, the majority candidate wins with probability

$$v(q^A; \eta^A) \equiv F(D(q^A) + \eta^A). \tag{13}$$

We can then replace (4) with (13) and show that all the results in Section 3 continue to hold.

5.2 Increasing Belief Disagreement

In many important cases, political disagreement stems solely from belief disagreement. As Callander (2011) points out, a large part of the difficulty in policy making is that policy makers may be uncertain about which policies produce which outcomes, and much political disagreement is over beliefs about this mapping.

Consider the spatial policy model from Section 2.4. Suppose voters share a common preference parameter $\beta^A = \beta^B \equiv \beta$, and recall that the optimal policy is $x^{i*}(q^i) = \beta E[\theta|q^i]$. Political disagreement (10) translates naturally into the degree of belief disagreement over expectations,

$$\begin{aligned} \mathcal{D}(q^A, q^B) &= \sum_{\theta\in\Theta} q^A_\theta \left[u^A(x^{*A}(q^A),\theta) - u^A(x^{*B}(q^B),\theta)\right] \\ &= \sum_{\theta\in\Theta} q^A_\theta \left[-(\beta E[\theta|q^A] - \beta\theta)^2 + (\beta E[\theta|q^B] - \beta\theta)^2\right] \\ &= \beta^2(E[\theta|q^A] - E[\theta|q^B])^2. \end{aligned} \tag{14}$$

Similarly, consider the tax model from Section 4.1. Suppose voters have heterogenous priors, but they have the same income normalized to one, $\beta^A = \beta^B = 1$. In this case, disagreement over the optimal tax derives solely from the belief disagreement over the marginal value of the public good. Suppose the production technology of the public good is $\psi = \frac{1}{2}$. The optimal tax rate becomes $x^{i*}(q^i) = \frac{E[\theta|q^i]^2}{4}$. Political disagreement takes the simple form $\mathcal{D}(q^A, q^B) = \frac{1}{4}(E[\theta|q^A] - E[\theta|q^B])^2$.

When voters have the same payoff function but different beliefs, can the IP increase political disagreement? As in these two applications, suppose political disagreement is a strictly increasing function of the difference between voters' expectation over the state. Although voters share a common payoff function, we show in the next proposition that if the state space is rich enough, then the IP can generically design an experiment that increases political disagreement with probability one.

Proposition 3 *Suppose political disagreement strictly increases in degree of belief disagreement over expectations, $\mathcal{D}(q^A, q^B) = R(|E[\theta|q^A] - E[\theta|q^B]|)$, where $R \geq 0$ and $R' > 0$. If $N \geq 4$, then the IP can generically*[14] *design an experiment that increases political disagreement with probability one. Consequently, the value of persuasion is positive for each finite incumbent's valence η^A.*

The following example illustrates Proposition 3 and shows how the IP can guarantee a higher disagreement.

[14]Genericity is interpreted over the space of pairs of prior beliefs.

Example 1 — Increasing Belief Disagreement: Let $\Theta = \{1, 2, 3, 4\}$. Consider priors $p^A = (.05, .45, .45, .05)$ and $p^B = (.45, .05, .05, .45)$, so that $E[\theta|p^A] = E[\theta|p^B] = 2.5$. Although prior beliefs are different, initial political disagreement is zero. The following binary signal $S = \{s_1, s_2\}$ is optimal for the IP. States 1 and 3 induce signal s_1 with probability one, while states 2 and 4 induce signal s_2 with probability one. After observing signal s_1 beliefs become $E[\theta|q^A] = 2.8$ and $E[\theta|q^B] = 1.2$, while s_2 induces $E[\theta|q^A] = 2.2$ and $E[\theta|q^B] = 3.8$. That is, from the point of view of voters A, candidate B not only "overreacts" to the information (updates his policy "too much"), but also moves the policy in the wrong direction. Therefore, the signal induces a strictly higher belief disagreement for any of its realizations. This signal is optimal independently of valence η^A and distribution F because each signal realization induces the maximum feasible disagreement, given the prior beliefs.$\square$

5.3 Application: Budget Allocation

We conclude this section by presenting an application where the degree of political disagreement is endogenously given by the degree of belief disagreement, as measured by the relative entropy. Although any informative signal decreases the expected political disagreement, the IP still finds it optimal to implement a partially informative signal when the expected valence of candidate A is sufficiently low.

Consider the following budget allocation model. The government has one dollar to allocate among $N \geq 2$ different government projects. Let $x_n \geq 0$ represent the amount of money allocated to project n, such that the budget balances. Thus $X = \{x \in [0, 1]^N | \sum_{n=1}^{N} x_n = 1\}$ and the vector $x = (x_1, \ldots, x_N) \in X$ represents a complete government budget.

There is uncertainty about the payoff derived from investing in each project. To simplify presentation, we consider the case where only one project is beneficial to voters — only one project can increase voters' payoff — while investment in any other project delivers a payoff of zero. Formally, there are N possible states, $\theta \in \Theta \equiv \{1, \ldots, N\}$, and citizens share a common payoff function: if the realized state is $\theta = n$ then voters receive a logarithmic payoff $\ln(x_n)$. In other words, $u(\theta, x) = \sum_{n=1}^{N} \mathbb{1}(n, \theta) \ln(x_n)$, where $\mathbb{1}(n, \theta) = 1$ if $\theta = n$, and $\mathbb{1}(n, \theta) = 0$ if $\theta \neq n$. If voter i has belief $q^i = (q^i_1, \ldots, q^i_N)$, then budget x delivers an expected policy payoff $\sum_{n=1}^{N} q^i_n \ln(x_n)$, where we apply the convention $0 \ln(0) = 0$. The logarithmic utility implies that each voter prefers the budget to be allocated proportionally to his own beliefs — the preferred budget x^{i*} of voter i is simply $x^{i*}_n = q^i_n$ for all n.

Political disagreement (10) becomes

$$\mathcal{D}^A(q^A, q^B) = \sum_{n=1}^{N} q^A_n \left[\ln(q^A_n) - \ln(q^B_n)\right] = \sum_{n=1}^{N} q^A_n \ln\left(\frac{q^A_n}{q^B_n}\right) \equiv D_{KL}(q^A||q^B).$$

The relative entropy $D_{KL}(q^A||q^B)$, or Kullback-Leibler distance[15] between probability distributions q^A and q^B, is a measure of the belief disagreement between the two groups. Therefore, in our electoral model, the degree of of *political disagreement* is given directly by the level of *belief disagreement* as measured by the relative entropy: from the point of view of the majority group, D_{KL} measures the difference in the expected payoff derived from the different policies favored by each group. Political disagreement is zero if and only if both groups share common beliefs, and it is increasing in the extent of belief disagreement between the groups. Figure 1(c) presents a binary-state example of disagreement as a function of the posterior belief of voter A, given a particular pair of prior beliefs.

It is a known fact that any informative experiment, on average, decreases the relative entropy. Consequently, information *always* decreases average political disagreement in this budget allocation model. Nevertheless, Lemma B.1 in the on-line Appendix B shows that it is possible to design an experiment with at least one realization that strictly increases disagreement under the following condition:

Condition C1: *Interior prior beliefs* (p^A, p^B) *are such that* $r_\theta - \ln(r_\theta) \neq r_{\theta'} - \ln(r_{\theta'})$ *for at least one pair of states* $\theta, \theta' \in \Theta$.

Condition **(C1)** is violated if priors are common, in which case $r_\theta - \ln(r_\theta) = 1$ for all states. Nevertheless, condition **(C1)** holds generically, where genericity is interpreted over the space of pairs of prior beliefs.

The policy advantage of candidate A derives solely from the belief disagreement among voters, thus full information disclosure is never optimal, $\eta^A_1 = -\infty$. Moreover, any information disclosure always decreases the expected disagreement. Does the IP ever benefit from

[15] Although it is not formally a distance measure, the relative entropy is a measure of the inefficiency of assuming that the probability distribution is q^B when the true distribution is q^A. See Cover and Thomas (2006, Chapter 2) for a discussion.

disclosing some information? The answer is yes if the incumbent politician is sufficiently incompetent.

Proposition 4 *In the budget allocation model, if* $\lim_{\eta^B \to -\infty} \frac{f(\eta^B)}{F(\eta^B)} = \infty$ *and condition* ***(C1)*** *holds, then there exists a finite cutoff* η_2^A *such that a partially informative is optimal iff* $\eta^A < \eta_2^A$.

Loosely speaking, condition $\lim_{\eta^B \to -\infty} \frac{f(\eta^B)}{F(\eta^B)} = \infty$ (which holds, for instance, in F is a Normal Distribution) implies that the IP's payoff v becomes a "very convex" function of beliefs when η^A is low. Consequently, although information always decreases average *political disagreement*, information can increase average *victory probability* if the majority candidate is sufficiently incompetent.

Figures 4 and 5 illustrate the IP's payoff in a binary-state example of our budget allocation model. See Section 3.2 for a discussion about the optimal experiment in that example.

6 Discussion

6.1 Log-convex Valence Distribution

Our results fundamentally depend on the assumption that the p.d.f. of the challenge's valence distribution is log-concave. The results are reversed if this p.d.f. is log-convex. In the log-convex case, the single-crossing property goes in the opposite direction: lower values of incumbent's competence η^A induce less experimentation, while higher competence induces more experimentation. Moreover, suppose that political disagreement is a strictly increasing function of the expected state, as in Section 4. If we change assumption **(A2′)** so that h and f are log-convex, then the optimal experiment is lower-censoring at some cutoff state θ_k. Furthermore, this cutoff state *decreases* with the incumbent's valence η^A, hence the experiment becomes more informative.

The reason for the sharp change in results is rooted on the change in the curvature of the incumbent's victory probability, as a function of political disagreement and valence η^A. Loosely speaking, in the log-concave case it is as if the IP features increasing absolute risk-aversion (IARA). This is so because the coefficient of absolute risk aversion $\frac{-u''}{u'}$ in our setup takes the form $-\frac{f'}{f}$, which is increasing if f is log-concave. When disagreement and incumbent's valence are low, the IP benefits from gambling on disagreement. That is, the IP benefits from implementing a risky experiment that might increase or decrease disagreement. When disagreement and valence are high, the IP prefers to avoid these gambles. In the log-convex case, it is as if the IP features DARA, and the reverse results hold.

6.2 Information and Voter Welfare

In many voting models, information is strategically provided to voters by interested parties. In some cases, this information can adversely affect voters' equilibrium welfare — voters' payoff would be higher if they made uninformed choices. For instance, in Alonso and Câmara (2015b) the information provided by the IP always weakly decreases the expected payoff of a majority of voters under a simple majority voting rule. This is so because the optimal experiment has signal realizations targeting different winning coalitions of voters.

In our model, the IP cannot target different winning coalitions because voters in group A are representative. Nevertheless, the next Proposition shows that the IP's optimal experiment may hurt all voters in majority group A.

Proposition 5 *Consider the spatial policy model from Section 2.4. Suppose voters have a common prior belief but different preferences. If either $\beta^A\beta^B < 0$ or $|\beta^B| > 2|\beta^A|$, then there exists a finite cutoff $\bar{\eta}^A$ such that, for any $\eta^A < \bar{\eta}^A$, the IP's optimal experiment strictly decreases the expected payoff of all voters in majority group A.*

The result follows from the different interests of IP and voters A, and the fact that the IP benefits from promoting disagreement. The IP's goal is to elect candidate A, not to ensure that the elected candidate implements a good policy for voters A. Candidate A is more likely to win when information leads candidate B to adopt a new policy that is worse from the point of view of voters A. This worse policy by B benefits the IP, but it hurts voters' A when candidate B is elected.

Under the conditions of the proposition, the IP implements a fully informative experiment, and the information (on average) leads candidate B to implement a worse policy for voters A. This is the case here as preference misalignment is sufficiently severe: If $\beta^A\beta^B < 0$

(candidates want policies with opposing signals) or $|\beta^B| > 2|\beta^A| \geq 0$ (candidate B "overreacts" to information from the point of view of A), then voter A strictly prefers an elected candidate B not to have access to any informative experiment. When η^A is sufficiently low, candidate A is unlikely to win (the benefit of providing information to candidate A is small) while candidate B is likely to win (the loss of providing information to candidate B is large). Hence, voters in group A are strictly worse off because of the experiment provided by the IP. They would prefer no experiment over the IP's equilibrium experiment.

Interestingly, the information strategically provided by the IP can decrease the expected payoff of a majority of voters even when voters share the same payoff function, so that political disagreement derives solely from belief disagreement. Consider Example 1 in Section 5.2. Without the IP's experiment, both candidates support the same policy. Consequently, the candidate with the highest realized valence wins. From the point of view of majority voters A, the information provided by the IP's optimal experiment has a small positive effect on the policy of candidate A, but a large negative effect on the policy of candidate B. From the point of view of voters A, information moves the policy of candidate B by too much and in the wrong direction. Consequently, as in Proposition 5, this negative effect dominates if candidate B is sufficiently more likely to win (η^A is sufficiently small).

6.3 Political Disagreement vs. Polarization

It is important to highlight that our notion of "political disagreement" is different from the notion of "polarization" present in many papers in the literature.[16] That is, an increase in political disagreement does not imply an increase in polarization, and vice versa. To illustrate this point, change the spatial policy model from Section 2.4 as follows: all voters share the same Euclidean policy payoff $u(x, \theta) = -|x - \theta|$, but have different prior beliefs. In this case, the optimal policy equals the expected median of the state, $x^{i*}(q^i) = M[\theta|q^i]$. Political disagreement then becomes

$$D^A(q^A, q^B) = -\sum_{\theta' \in \Theta} q^A_{\theta'} \Big[\big| M[\theta|q^A] - \theta' \big| - \big| M[\theta|q^B] - \theta' \big| \Big].$$

[16] Although there are different definitions of polarization in the literature, here we define polarization as the Euclidean distance between the policies supported by the candidates as in Dixit and Weibull (2007).

To see that information might strictly increase polarization and strictly decrease political disagreement, let $\Theta = \{-2, -1, +1, +2\}$ and consider priors $p^A = (.06, .8, .1, .04)$ and $p^B = (.04, .1, .8, .06)$. At the prior belief, candidate A prefers policy -1 while candidate B prefers $+1$. The original degree of political disagreement is $D^A(p^A, p^B) = 36/25$. Initial political disagreement is high because voters in group A are very confident that the state is -1 and not $+1$. Consider a binary experiment that simply reveals if the state is in the partition $\{-1, +1\}$ or $\{-2, +2\}$. When the experiment reveals that the state is in partition $\{-2, +2\}$, updated beliefs become $q^A = (.6, 0, 0, .4)$ and $q^B = (.4, 0, 0, .6)$. Candidates' preferred policies change to -2 and $+2$. Hence, the information results in more polarized policies. However, there is now a lower degree of political disagreement, $D^A(q^A, q^B) = 4/5$. Although policies are more polarized (farther away from each other), voters in group A now believe that there is a much higher chance that the opposing policy championed by candidate B might be the correct policy. In a nutshell, optimal policies are further apart, but voters suffer a smaller loss from appointing the rival candidate.

6.4 Related Literature

Our paper relates to the recent papers on Bayesian persuasion that follow KG. As in Alonso and Câmara (2015b), the goal of the incumbent party ("sender") in our model is to sway elections in favor of its preferred alternative. However, in Alonso and Câmara (2015b) the sender simply wants to convince a majority of voters that the proposal is better than the status quo. An important feature of our model is that the IP would like to convince voters from the majority group not only that their candidate supports a good policy, but also that the minority candidate supports a bad policy. That is, the "relative" expected payoff from the policies (the degree of political disagreement) is crucial.

Kolotilin et al. (2015) study a Bayesian persuasion model with a single receiver that has private information about his type, and a sender with a payoff that is a linear increasing function of the expected state. Although their setup and focus are quite different from ours, they find (Theorem 2) that if the receiver's type has a log-concave (log-convex) p.d.f., then it is optimal to use an upper (lower) censorship signal.[17] Interestingly, we can extend

[17]We use the term upper-censoring, since it is more common on the statistics literature.

our Proposition 2 as follows: if f and h are log-concave, then there exists an optimal upper-censoring experiment; if f and h are log-convex, then there exists an optimal lower-censoring experiment.[18] Their proof relies more on a mechanism design approach, while our proof is closer to the concave-closure approach of KG.

Our paper also relates to the literature on how policy experimentation (learning how different policies map into payoffs) can influence future policies and electoral outcomes — e.g., Callander (2011) and Callander and Hummel (2014). Millner, Ollivier, and Simon (2014) focus on the role of heterogenous priors and show how a policy-motivated party may over experiment, in order to show to the opposite party that its belief is "wrong" and *reduce* belief disagreement. In our model the purely office-motivated IP strategically discloses information to *increase* belief disagreement and influence elections.

We also contribute to the literature on how access to information can increase polarization and disagreement (e.g., Dixit and Weibull 2007). In most papers, a higher disagreement is a somewhat unintended side-effect of the actions of individuals generating information, such as the media catering information to the demand of biased voters. In our model, the IP generates information with the sole purpose of increasing disagreement and benefiting its supported candidate.

Finally, note that our results do not follow from reputation concerns of candidates. That is, our candidates are not trying to signal competence to voters by their choice of experiment. Our result follows from the fact the marginal benefit from increasing political disagreement is lower for more competent politicians than for less competent politicians. See Bernecker, Boyer and Gathmann (2014) for a recent model in which politicians use their choice of policy experiment to signal competence.

7 Extensions

In this Section we extend our model to the cases where the IP supports the minority candidate and experiments are costly. In the on-line Appendix B we examine the role of post-election

[18]Note that if disagreement is a linear function of the expected state, then h is both log-concave and log-convex. In this case, our optimal experiment only depends on whether valence distribution f is log-concave or log-convex.

information and consider valence shocks that are independent across voters.

7.1 IP Supports the Minority

So far we have assumed that the IP supports the majority candidate, that is, candidate A is the incumbent. Now suppose that the minority party $\mathcal{B}$ is in power (hence controls the experiment) and supports the incumbent candidate B. Since the political advantage of the majority candidate is solely due to political disagreement, the IP now benefits from *decreasing* political disagreement. The results from Sections 3 now apply to the valence η^B of the incumbent: the IP uses less informative experiments when the minority incumbent is more competent (η^B is high) and more informative experiments when he is less competent. Interestingly, the optimal experiment in Proposition 2 becomes *lower-censoring*: the minority party pools low disagreement states and fully reveals high disagreement states.

Moreover, consider the models of Sections 5.2 and 5.3, where citizens share the same payoff function but hold different prior beliefs. In these cases, regardless of prior beliefs, full information disclosure is *always* optimal for the minority candidate. Complete transparency eliminates political disagreement and the policy advantage of the majority candidate, increasing the chances of the minority candidate.

7.2 Costly Policy Experiments

Our basic model assumes that experiments are costless. If every experiment is equally costly, then whenever the IP decides to implement an experiment she implements the optimal experiment we describe. The only change is that she only implements an experiment if the value of persuasion is higher than the fixed cost of implementation.

What if different experiments have different costs? Following Gentzkow and Kamenica (2014), suppose that the cost of an experiment is given by the expected relative entropy of the beliefs that it induces. Consequently, more informative experiments are more costly. In this case, one can show that our results from Section 3 continue to hold. However, perfectly revealing a state is infinitely costly. Therefore, full information disclosure and upper-censoring experiments are never optimal.

8 Conclusion

In this paper we study strategic policy experimentation by office-motivated parties. That is, we study strategic learning about how different policies map into the payoffs of different voters. In our environment, information changes the degree of political disagreement and sways future elections — experimental outcomes that increase disagreement increase the victory probability of the candidate that has similar preferences and beliefs as a majority of voters. Therefore, office-motivated party bureaucrats supporting this majority candidate benefit from policy experiments that create more dissent between the majority and the minority.

We derive several results regarding the informational content of optimal policy experiments. First, we show that the informativeness of optimal experiments has a monotonicity property with respect to the competence of the incumbent politician. Optimal experiments under a more competent politician are never more informative than optimal experiments under a less competent politician. Second, when disagreement is an increasing function of the expected state, there exists an optimal experiment that is upper censoring — it fully reveals low disagreement states and pools high disagreement states. Under these conditions, we establish the sharper result that the informativeness of the optimal experiment in fact decreases with the incumbent's competence. Finally, we consider cases in which all voters share the same payoff function, so that political disagreement is solely due to belief disagreement. We show that, even in these cases, policy experiments can be used to increase disagreement and benefit the majority candidate.

Our basic idea that experimentation can be strategically used to increase political disagreement and influence future elections can be exploited in different models. For example, one could consider a repeated-election model in which new information can be uncovered each period and influence all future elections and policies. One could also consider the case where the IP is both office and policy motivated. We find these alternative models promising and we leave them to future work.

A Appendix

Before we present the proof of Propositions 1, we provide the following Lemma.

Lemma A. 1 *Define*

$$G\left(a,b,\eta^{A}\right)\equiv\frac{F\left(b+\eta^{A}\right)-F\left(a+\eta^{A}\right)}{f\left(a+\eta^{A}\right)},\tag{15}$$

*where F and f satisfy **(A1)**. Fix any $a,b,c\in\mathbb{R}$. Then:*

(i) $G(a,b,\eta^{A})$ is non-increasing in η^{A}, and it is strictly decreasing if f is strictly log-concave and $a\neq b$;

(ii) There exists a $\eta^{A\prime}<+\infty$ such that for any $\eta^{A}\geq\eta^{A\prime}(a,b)$ we have $G(a,b,\eta^{A})\leq b-a$; and a $\eta^{A\prime\prime}<+\infty$ such that for any $\eta^{A}\leq\eta^{A\prime\prime}(a,b)$ we have $G(a,b,\eta^{A})\geq b-a$

(iii) If $b>a$ and $\lim_{\eta^{B}\to-\infty}\frac{f(\eta^{B})}{F(\eta^{B})}=\infty$, then there exists a $\eta^{A\prime}>-\infty$ such that for any $\eta^{A}\leq\eta^{A\prime}$ we have $\frac{F(b+\eta^{A})-F(a+\eta^{A})}{F(a+\eta^{A})}\geq c$;

Proof of Lemma 1:

Part (i): We first rewrite the function G as

$$G\left(a,b,\eta^{A}\right)=\int_{0}^{b-a}\frac{f\left(a+\eta^{A}+s\right)}{f\left(a+\eta^{A}\right)}ds.$$

Since f is log-concave, it exhibits decreasing ratios in the sense that for every $s>0$ and $\eta^{A}\geq\eta^{A\prime}$ we have

$$\frac{f\left(a+\eta^{A\prime}+s\right)}{f\left(a+\eta^{A\prime}\right)}\geq\frac{f\left(a+\eta^{A}+s\right)}{f\left(a+\eta^{A}\right)}.\tag{16}$$

Suppose first that $b>a$. Then integrating both sides of (16) between 0 and $b-a$ shows that $G\left(a,b,\eta^{A\prime}\right)\geq G\left(a,b,\eta^{A}\right)$. Now suppose that $a>b$. Then for any $s\in[0,a-b]$ we can rewrite (16) as

$$\frac{f\left(a+\eta^{A}-s\right)}{f\left(a+\eta^{A}\right)}\geq\frac{f\left(a+\eta^{A\prime}-s\right)}{f\left(a+\eta^{A\prime}\right)}.$$

Integrating between 0 and $b-a$ we conclude that

$$-G\left(a,b,\eta^{A}\right)=\int_{0}^{a-b}\frac{f\left(a+\eta^{A}-s\right)}{f\left(a+\eta^{A}\right)}ds\geq\int_{0}^{a-b}\frac{f\left(a+\eta^{A\prime}-s\right)}{f\left(a+\eta^{A\prime}\right)}ds=-G\left(a,b,\eta^{A\prime}\right),$$

or, in other words, $G\left(a,b,\eta^{A\prime}\right)\geq G\left(a,b,\eta^{A}\right)$.

To show that strict monotonicity follows from strict log-concavity, we first note that the pdf f must be continuous in its support (as the exponential of a concave, and hence continuous, function). Strict log-concavity implies that (16) holds with strict inequality whenever $\eta^A > \eta^{A\prime}$; the continuity of f implies that if $\eta^A > \eta^{A\prime}$, then the minimum separation between the left hand side and right hand side in (16) is bounded away from zero for any $s \in (0, a-b]$. Therefore, the integral is also bounded away from zero and $G\left(a, b, \eta^{A\prime}\right) > G\left(a, b, \eta^A\right)$.

Part (ii-iii): Fix any c and $b > a$. We now establish some general facts that will be used in the proofs of Parts ii-iv. Log-concavity of f and full support of F imply that F is also log-concave. Therefore, the ratio $\frac{f(\eta^B)}{F(\eta^B)}$ is everywhere decreasing. Log-concavity of f also implies that f is unimodal. Full support then ensures the existence of a finite η^{A*} such that f is weakly increasing for $\eta^A \leq \eta^{A*}$ and weakly decreasing for all $\eta^A \geq \eta^{A*}$.
Part (ii): Fix any $a, b \in \mathbb{R}$, and define $\eta^{A\prime} = \eta^{A*} - \min\{a, b\}$. This implies that for any $\eta^A \geq \eta^{A\prime}$, f is weakly decreasing in $[\min\left(a+\eta^A, b+\eta^A\right), \max\left(a+\eta^A, b+\eta^A\right)]$. We now show that for any $\eta^A \geq \eta^{A\prime}$, $G(a, b, \eta^A) \leq b-a$. The results hold trivially if $a=b$. Suppose $b > a$. Then for any $\eta^A \geq \eta^{A\prime}$ we have that monotonicity of f implies that

$$G(a,b,\eta^A) = \frac{\int_{a+\eta^A}^{b+\eta^A} f(y)\, dy}{f(a+\eta^A)} \leq \frac{\int_{a+\eta^A}^{b+\eta^A} f(a+\eta^A)\, dy}{f(a+\eta^A)} = b-a.$$

Now suppose $b < a$. By the same argument, for any $\eta^A \geq \eta^{A\prime}$ we have

$$G(a,b,\eta^A) = \frac{\int_{a+\eta^A}^{b+\eta^A} f(y)\, dy}{f(a+\eta^A)} = \frac{-\int_{b+\eta^A}^{a+\eta^A} f(y)\, dy}{f(a+\eta^A)} \leq \frac{-\int_{b+\eta^A}^{a+\eta^A} f(a+\eta^A)\, dy}{f(a+\eta^A)} = b-a,$$

which concludes the proof.
Part (iii): Define $\eta^{A\prime} = \eta^{A*} - b$ so that for any $\eta^A \leq \eta^{A\prime}$, f is weakly increasing in $[a+\eta^A, b+\eta^A]$. This implies that for any $\eta^A \leq \eta^{A\prime}$ we have

$$\frac{F(b+\eta^A) - F(a+\eta^A)}{F(a+\eta^A)} = \frac{\int_{a+\eta^A}^{b+\eta^A} f(y)\, dy}{F(a+\eta^A)} \geq \frac{\int_{a+\eta^A}^{b+\eta^A} f(a+\eta^A)\, dy}{F(a+\eta^A)} = \frac{f(a+\eta^A)(b-a)}{F(a+\eta^A)}.$$

Since $\lim_{\eta^B \to -\infty} \frac{f(\eta^B)}{F(\eta^B)} = \infty$, the right hand side is unbounded for fixed a and b. As $\frac{f(\eta^B)}{F(\eta^B)}$ is decreasing, it follows that there is a $\hat{\eta}^A > -\infty$ such that $\frac{f(a+\eta^A)(b-a)}{F(a+\eta^A)} \geq c$ for all $\eta^A \leq \hat{\eta}^A$, concluding the proof. ■

Proof of Lemma 1:

Consider an experiment π that generates a distribution $\sigma \in \Delta(\Theta)$ over posterior beliefs.

Note that this distribution is independent of valences. For any q in the support of σ the change in the victory probability of the majority candidate is

$$v(q;\eta^A) - v(p;\eta^A) = F(D(q_A) + \eta^A) - F(D(p) + \eta^A) = f(D(p) + \eta^A)G(D(p), D(q), \eta^A),$$

where G is defined by (15). Therefore, the expected change in victory probability from experiment π can be written as

$$E_\pi[v(q;\eta^A) - v(p;\eta^A)] \quad = \quad f(D(p) + \eta^A) \int_{q \in \text{ supp}(\sigma)} G(D(p), D(q), \eta^A) d\sigma.$$

Because $f > 0$ rewrite

$$\frac{E_\pi[v(q;\eta^A) - v(p;\eta^A)]}{f(D(p) + \eta^A)} \quad = \quad \int_{q \in \text{supp}(\sigma)} G(D(p), D(q), \eta^A) d\sigma. \tag{17}$$

From Lemma 1-(i) we know that G is non-increasing in η^A, hence the LHS of (17) is non-increasing in η^A. This implies that if $E_\pi[v(q;\eta^A) - v(p;\eta^A)] \leq 0$ then $E_\pi[v(q;\eta^{A\prime}) - v(p;\eta^{A\prime})] \leq 0$ for any $\eta^{A\prime} > \eta^A$, concluding part (i) of the Lemma. This also implies that if $E_\pi[v(q;\eta^A) - v(p;\eta^A)] \geq 0$ then $E_\pi[v(q;\eta^{A\prime}) - v(p;\eta^{A\prime})] \geq 0$ for any $\eta^{A\prime} < \eta^A$, concluding part (ii) of the Lemma. ■

Proof of Proposition 1:

Suppose π^* is an optimal experiment given valence η^A. Take any $\eta^{A\prime} > \eta^A$ and any π' that is Blackwell more informative than π^*. The proof has two steps. In the first step we construct a grand experiment $\{\pi^*, \{\pi_s\}_{s \in S}\}$ that is payoff equivalent to π'. In the second step we show that since π^* is weakly better than $\{\pi^*, \{\pi_s\}_{s \in S}\}$ when valence is η^A, then π^* is weakly better than $\{\pi^*, \{\pi_s\}_{s \in S}\}$ when the valence is higher. Consequently, π^* is weakly better than π' for any $\eta^{A\prime} > \eta^A$.

Step 1: Let Q^* be the set of posterior beliefs q in the support of π^*, and Q' be the set of posterior beliefs q in the support of π'. The fact that π' is Blackwell more informative than π^* implies that Q^* is contained in the convex hull of Q'. Consequently, we can decompose the more informative experiment into two parts. Players first observe the realization $s \in S$ of experiment π^*. Then they observe the realization of an additional experiment π_s. Note that experiment π_s can be different for different signal realizations. This grand experiment

$\{\pi^*, \{\pi_s\}_{s\in S}\}$ induces the same probability distribution over posterior beliefs as π', hence they are payoff equivalent.

Step 2: When the incumbent's valence is η^A, optimality of π^* implies that after each signal realization s of π^* the IP does not benefit from further disclosing information. That is, for every posterior belief q^* in the support of π^* and every experiment π_s, we have

$$E_{\pi_s}[v(q;\eta^A)|q^*] \quad \leq \quad v(q^*;\eta^A). \tag{18}$$

Apply Lemma 1-(i) to (18): for each posterior belief q^* in the support of π^*, for every $\eta^{A\prime} > \eta^A$, and every experiment π_s, we have $E_{\pi_s}[v(q;\eta^{A\prime})|q^*] \leq v(q^*;\eta^{A\prime})$. Consequently, after each signal realization of π^* the IP does not benefit from further disclosing information. Taking expectations over the realizations of π^* yields $E_{\{\pi^*,\{\pi_s\}_{s\in S}\}}[v(q;\eta^{A\prime})] \leq E_{\pi^*}[v(q;\eta^{A\prime})]$. ■

Proof of Corollary 1:

Suppose that for some η_2^A a completely uninformative experiment is optimal, and note that every possible experiment is Blackwell more informative than no information. Then Proposition 1 immediately implies that a completely uninformative experiment is weakly better than every other experiment for any $\eta^A > \eta_2^A$.

Suppose that for some η_1^A the fully informative experiment π^{FD} is optimal. Alonso and Câmara (2015a, Corollary 2) show that a fully informative experiment is optimal if and only if $E_{\pi^{FD}}[v(q';\eta_1^A)|q] \geq v(q;\eta_1^A)$ for all $q \in \Delta(\Theta)$. For every $\eta^A < \eta_1^A$, Lemma 1-(ii) implies that $E_{\pi^{FD}}[v(q';\eta^A)|q] \geq v(q;\eta^A)$ for all $q \in \Delta(\Theta)$. Hence π^{FD} is optimal for all $\eta^A < \eta_1^A$. ■

The following technical Lemma will be used in proof of Proposition 2.

Lemma A. 2 *Suppose F and H are twice differentiable, with derivatives f and h that are strictly positive and log-concave.*
(i) Define

$$K(e;\eta^A) \quad \equiv \quad F(H(e)+\eta^A).$$

Then, given η^A, there exists a unique cutoff $\bar{e}(\eta^A)$ in the extended real line such that K is concave if $e \geq \bar{e}(\eta^A)$ and strictly convex if $e < \bar{e}(\eta^A)$. Moreover, $\bar{e}(\eta^A)$ decreases in η^A.

(ii) Define

$$A(e;\eta^A,e') \equiv F(H(e')+\eta^A) - F(H(e)+\eta^A) + f(H(e')+\eta^A)h(H(e')+\eta^A)[e-e'].$$

Fix any $e' \geq \bar{e}(\eta^A)$, where $\bar{e}(\eta^A)$ was defined in Part (i). There exits a cutoff $\underline{e}(\eta^A) \leq \bar{e}(\eta^A)$ such that $A(e;\eta^A,e') < 0$ if $e < \underline{e}(\eta^A)$ and $A(e;\eta^A,e') > 0$ if $e \in (\underline{e}(\eta^A), \bar{e}(\eta^A))$.

Proof of Lemma 2:

Part (i): Take the first and second derivatives of K:

$$\begin{aligned} \frac{\partial K(e;\eta^A)}{\partial e} &= f(H(e)+\eta^A)h(e) \\ \frac{\partial^2 K(e;\eta^A)}{\partial e^2} &= f'(H(e)+\eta^A)[h(e)]^2 + f(H(e)+\eta^A)h'(e) \\ &= [h(e)]^2 f(H(e)+\eta^A)\left\{\frac{f'(H(e)+\eta^A)}{f(H(e)+\eta^A)} + \frac{h'(e)}{[h(e)]^2}\right\}. \end{aligned} \tag{19}$$

Since $[h(e)]^2 f(H(e)+\eta^A) > 0$, this second derivative has the same sign as the term in brackets. We next show that the term in brackets decreases in e. Log-concavity of f and monotonicity of H implies that f'/f is decreasing in e. The derivate of $\frac{h'}{[h]^2}$ is $\frac{h''}{[h]^2} - \frac{2[h']^2}{[h]^3} = \frac{1}{h}\left\{\frac{h''}{h} - \frac{2[h']^2}{h^2}\right\} \leq \frac{1}{h}\left\{\frac{h''}{h} - \frac{[h']^2}{h^2}\right\} \leq 0$, where we used the fact that $h > 0$ and the last inequality follows since h is log-concave.[19]

Therefore, define $\bar{e}(\eta^A)$ in the extended real line as the infimum e such that $\frac{\partial^2 K(e;\eta^A)}{\partial e^2} \leq 0$. It then follows that K is concave if $e \geq \bar{e}(\eta^A)$ and strictly convex if $e < \bar{e}(\eta^A)$. Moreover, since f'/f decreases in η^A, the cutoff $\bar{e}(\eta^A)$ also decreases in η^A.

Part (ii): Take the first and second derivatives of A:

$$\begin{aligned} \frac{\partial A(e;\eta^A,e')}{\partial e} &= -f(H(e)+\eta^A)h(e) + f(H(e')+\eta^A)h(e') \\ \frac{\partial^2 A(e;\eta^A,e')}{\partial e^2} &= - f'(H(e)+\eta^A)[h(e)]^2 - f(H(e)+\eta^A)h'(e) \\ &= -[h(e)]^2 f(H(e)+\eta^A)\left\{\frac{f'(H(e)+\eta^A)}{f(H(e)+\eta^A)} + \frac{h'(e)}{[h(e)]^2}\right\}. \end{aligned} \tag{20}$$

Note that (20) has the opposite sign as (19). Therefore, A is convex if $e \geq \bar{e}(\eta^A)$ and strictly concave if $e < \bar{e}(\eta^A)$. Fix any $e' \geq \bar{e}(\eta^A)$. We first show that for all $e \geq \bar{e}(\eta^A)$ we have $A(e;\eta^A,e') \geq A(e';\eta^A,e') = 0$. To see this, note that at $e = e'$ we have $A(e';\eta^A,e') = 0$, and

[19] Log-concavity implies h'/h decreases, that is, its derivative is negative, $h''/h - [h']^2/h^2 \leq 0$

at this point $\partial A(e;\eta^A,e')/\partial e = 0$. The result then follows since A is convex in the segment $e \geq \bar{e}(\eta^A)$.

In particular, $A(\bar{e}(\eta^A);\eta^A,e') \geq 0$. This together with strict concavity of A imply that, in the range $e < \bar{e}(\eta^A)$, function A crosses zero at most once and from below. Letting $\underline{e}(\eta^A)$ in the extended real line be this crossing point, we conclude part (ii) of the proof. ■

Proof of Proposition 2:

Suppose **(A.2′)** holds.

Step 1) Because $v(q;\eta^A) = F(H(\langle q,\theta\rangle)+\eta^A)$ and $\langle q,\theta\rangle$ is linear in q, Lemma 2 implies that there exists a $\bar{e}(\eta^A)$ in the extended real line such that: if $\langle q,\theta\rangle > \bar{e}(\eta^A)$, then v is a locally concave function of q; if $\langle q,\theta\rangle < \bar{e}(\eta^A)$, then v is a locally strictly convex function of q in any direction that changes expectation $\langle q,\theta\rangle$.

Step 2) Given η^A, suppose π^* is an optimal experiment. We first show that there exists an experiment π' that (i) is weakly better than π^*, hence π' is also optimal, and (ii) letting σ' be the distribution of posterior beliefs induced by π^*, there is at most one non degenerate belief in the support of σ'.

Let σ^* be the the distribution of posterior beliefs induced by the optimal experiment π^*. Using the cutoff $\bar{e}(\eta^A)$ from Step 1, all beliefs q in the support of σ^* such that $\langle q,\theta\rangle < \bar{e}(\eta^A)$ must be degenerate, since v is strictly convex in every direction of changes in q that change the expectation (changes in expectation are possible iff q is not degenerate). All remaining beliefs in the support of σ^* are then in the concave region $Q^+ \equiv \{q \in \Delta(\Theta) : \langle q,\theta\rangle \geq \bar{e}(\eta^A)\}$.

Concavity of v in the set Q^+ implies the following. We can construct an alternative experiment π' that induces (i) the same distribution of degenerate beliefs as π^* in the strictly convex region, and (ii) one non degenerate posterior belief q^+ in the concave region Q^+, such that π' is weakly better than π^*. Because it might be the case that π^* does not induce beliefs in the concave region Q^+, formally π' has at most on non degenerate belief in its support and π' is an optimal experiment.

Step 3) We now solve for the optimal experiment in the constrained set of experiments that induce at most one non degenerate belief. Using the result from Step 2, this implies that this constrained experiment is weakly better than any other unconstrained experiment.

Consider any experiment that induces at most one non degenerate belief. Without loss of generality, define the signal space as $S \equiv \{s_{\theta_1}, \ldots, s_{\theta_N}, s_{\text{pooling}}\}$. Each state $\theta \in \Theta$ induces the pooling signal s_{pooling} with probability $\alpha_\theta \in [0,1]$, and induces the fully revealing signal s_θ with probability $1-\alpha_\theta$. Given $\alpha = (\alpha_{\theta_1}, \ldots, \alpha_{\theta_N})$, let $q^+(\alpha) \equiv \left\{\frac{\alpha_\theta p_\theta}{\sum_{\theta' \in \Theta} \alpha_{\theta'} p_{\theta'}}\right\}_{\theta \in \Theta}$ be the updated posterior belief after observing s_{pooling}, and $E^+(\alpha) \equiv \langle q^+(\alpha), \theta\rangle = \frac{\sum_{\theta\in\Theta} \alpha_\theta p_\theta \theta}{\sum_{\theta\in\Theta}\alpha_\theta p_\theta}$ be the updated expectation of θ. The IP's problem then simplifies to choose the α the maximizes victory probability:

$$\max_{\alpha_\theta \in [0,1], \theta \in \Theta} \Pi(\alpha) \equiv \left(\sum_{\theta\in\Theta} \alpha_\theta p_\theta\right) F(H(E^+(\alpha)) + \eta^A) + \sum_{\theta\in\Theta} (1-\alpha_\theta) p_\theta F(H(\theta) + \eta^A).$$

We now solve for a optimal α^* and show that the optimal experiment is upper-censoring, that is, there exists a cutoff state θ_k such that $\alpha^*_\theta = 0$ if $\theta < \theta_k$ and $\alpha^*_\theta = 1$ if $\theta > \theta_k$.

If $\theta_N \le \bar{e}(\eta^A)$, then v is everywhere strictly convex and full information disclosure is optimal, $\alpha^*_\theta = 0$ for all θ. If $\bar{e}(\eta^A) \le \theta_1$ then v is everywhere concave and no information disclosure is optimal, $\alpha^*_\theta = 1$ for all θ. Note that full disclosure and no disclosure are the extreme cases of upper-censoring, where the cutoff state is θ_N and θ_1, respectively.

Consider the remaining case $\theta_1 < \bar{e}(\eta^A) < \theta_N$. From Steps 1 and 2, it must be the case that pooling posterior belief $q^+(\alpha^*)$ is in the concave region, $\langle q^+(\alpha^*), \theta\rangle \ge \bar{e}(\eta^A)$. Moreover, since posterior beliefs in the convex region must be degenerate, we have $\alpha^*_\theta = 1$ for all $\theta \ge \bar{e}(\eta^A)$. Now consider the strictly convex region $\theta < \bar{e}(\eta^A)$, and note that $E^+(\alpha^*) \ge \bar{e}(\eta^A)$ since it is in the concave region. Taking the derivative of the objective function with respect to $\alpha_{\theta'}$ for each state $\theta' < \bar{e}(\eta^A)$ yields

$$\begin{aligned} \frac{\partial \Pi(\alpha)}{\alpha_{\theta'}} &= p_{\theta'} F(H(E^+(\alpha)) + \eta^A) - p_{\theta'} F(H(\theta') + \eta^A) \\ &\quad + \left(\sum_{\theta\in\Theta} \alpha_\theta p_\theta\right) f(H(E^+(\alpha)) + \eta^A) h(E^+(\alpha)) \frac{\partial E^+(\alpha)}{\partial \alpha_{\theta'}}, \end{aligned}$$

where $\frac{\partial E^+(\alpha)}{\partial \alpha_{\theta'}} = \frac{p_{\theta'}\theta'}{\sum_{\theta\in\Theta}\alpha_\theta p_\theta} - \frac{p_{\theta'}[\sum_{\theta\in\Theta}\alpha_\theta p_\theta \theta]}{[\sum_{\theta\in\Theta}\alpha_\theta p_\theta]^2} = \frac{p_{\theta'}}{\sum_{\theta\in\Theta}\alpha_\theta p_\theta}\left[\theta' - E^+(\alpha)\right]$. Simplify

$$\frac{\partial \Pi(\alpha)}{\alpha_{\theta'}} = p_{\theta'} \left\{ F(H(E^+(\alpha)) + \eta^A) - F(H(\theta') + \eta^A) + f(H(E^+(\alpha)) + \eta^A) h(E^+(\alpha)) \left[\theta' - E^+(\alpha)\right] \right\}.$$

Because $p_{\theta'} > 0$, the derivative has the same sign as the term in brackets. Note that the term in brackets equals to $A(e; \eta^A, e')$ from Lemma 2(ii), evaluated at $e = \theta'$ and $e' = E^+(\alpha) \ge$

$\bar{e}(\eta^A)$. Therefore, there exists a cutoff $\underline{e}(\eta^A) \leq \bar{q}(\eta^A)$ such that the derivative $\frac{\partial \Pi(\alpha)}{\alpha_{\theta'}}$ evaluated at α^* is: strictly positive for any state $\theta' \in (\underline{e}(\eta^A), \bar{e}(\eta^A))$, hence the upper constraint binds and $\alpha_{\theta'} = 1$, and strictly negative for any state $\theta' < \underline{e}(\eta^A)$, hence the lower constraint binds and $\alpha_{\theta'} = 0$. Therefore, the optimal experiment is upper-censoring with cutoff $\underline{e}(\eta^A)$.

The fact that the optimal censoring cutoff weakly decreases with η^A follows immediately from Proposition 1: strictly increasing the censoring cutoff increases the informativeness of the experiment, but the IP does not benefit from a more informative experiment if η^A increases. ■

The proof of Proposition 3 will be based on the following Lemma:

Lemma A. 3 *Suppose political disagreement takes the form* $\mathcal{D}(q^A, q^B) = R(|E[\theta|q^A] - E[\theta|q^B]|)$*, where* $R \geq 0$ *and* $R' > 0$*. Define the vector* $v = r\left(\theta - E[\theta|q^B]\right)$*, the linear subspaces* $W_1 = \left\{x \in \mathbb{R}^{card(\Theta)} : \langle x, 1\rangle = 0\right\}$ *and* $W_{\theta - v} = \left\{x \in \mathbb{R}^{card(\Theta)} : \langle x, \theta - v\rangle = 0\right\}$*. If the projections of* θ *and* r *are not negatively collinear with respect to* $W_1 \cap W_{\theta-v}$*, then there exists an experiment* π *where all signal realizations increase political disagreement.*

Proof of Lemma 3:

Use (11) to rewrite

$$\mathcal{D}(q^A, q^B) \;=\; R(|E[\theta|q^A] - E[\theta|q^B]|) = R\left(\left|\left\langle q^A, \theta\right\rangle - \frac{\left\langle q^A r, \theta\right\rangle}{\left\langle q^A, r\right\rangle}\right|\right) \equiv D(q^A)$$

Define $q^A = \varepsilon\lambda + p^A$, with $\lambda \in W_1 = \{x : \langle x, 1\rangle = 0\}$ and $\varepsilon \in \mathbb{R}$, and let

$$L(\varepsilon; \lambda) = \left\langle q^A, \theta\right\rangle - \frac{\left\langle q^A r, \theta\right\rangle}{\left\langle q^A, r\right\rangle} = \varepsilon \left\langle \lambda, \theta\right\rangle + E[\theta|q^A] - \frac{\varepsilon \left\langle \lambda, r\theta\right\rangle + E[\theta|q^B]}{\varepsilon \left\langle \lambda, r\right\rangle + 1}$$

Disagreement is a strictly increasing function of the absolute value of $L(\varepsilon; \lambda)$. First suppose that $L(\varepsilon; \lambda) \geq 0$. We will show that under the conditions of the proposition one can always find a vector of "marginal beliefs" λ' such that L achieves a local minimum with respect to ε at $\varepsilon = 0$. This means that along the line λ' and in a neighborhood of 0, any belief $q^A = \varepsilon\lambda' + p^A$ with $\varepsilon > 0$ increases L, and thus $D(q^A) > D(p^A)$, while any belief $q^A = \varepsilon\lambda' + p^A$ with $\varepsilon < 0$ also increases L, yielding $D(q^A) > D(p^A)$. That is, we have found collinear beliefs that can average to the prior and that increase D.

First, we have

$$\begin{aligned}\frac{dL}{d\varepsilon} &= \langle\lambda,\theta\rangle - \frac{\langle\lambda,r\theta\rangle - \langle\lambda,r\rangle E\left[\theta|q^B\right]}{\left(\varepsilon\langle\lambda,r\rangle+1\right)^2}\\ \frac{d^2L}{d\varepsilon^2} &= \frac{2\langle\lambda,r\rangle\left[\langle\lambda,r\theta\rangle - \langle\lambda,r\rangle E\left[\theta|q^B\right]\right]}{\left(\varepsilon\langle\lambda,r\rangle+1\right)^3}\end{aligned}$$

For $L(\varepsilon;\lambda)$ to achieve a local minimum at $\varepsilon=0$ it is sufficient that there exists $\lambda\in W$ such that

$$\left.\frac{dL}{d\varepsilon}\right|_{\varepsilon=0} = 0 \Rightarrow \langle\lambda,\theta\rangle = \left\langle\lambda, r\left(\theta - E\left[\theta|q^B\right]\right)\right\rangle \tag{21}$$

$$\left.\frac{d^2L}{d\varepsilon^2}\right|_{\varepsilon=0} > 0 \Rightarrow \langle\lambda,r\rangle\left\langle\lambda, r\left(\theta - E\left[\theta|q^B\right]\right)\right\rangle > 0 \tag{22}$$

Since θ and r are not negatively collinear with respect to $W_1\cap W_{\theta-v}$ then there exists $\lambda'\in W_1\cap W_{\theta-v}$ with $\langle\lambda',\theta\rangle\langle\lambda',r\rangle>0$ (see Alonso and Camara 2015a). Since $\lambda'\in W_{\theta-v}$ then λ' satisfies (21). Then, given (21), the fact that $\langle\lambda',\theta\rangle\langle\lambda',r\rangle>0$ implies that λ' also satisfies (22). Therefore $L(\varepsilon;\lambda')$ achieves a local minimum at $\varepsilon=0$.

Now consider the remaining case, $L(\varepsilon;\lambda)<0$. Since disagreement strictly increases in the absolute value of L, we now can increase disagreement by decreasing L. The same steps of the proof above can be used to show that under the conditions of the proposition one can always find a vector of "marginal beliefs" λ'' such that L achieves a local maximum with respect to ε at $\varepsilon=0$. This follows as the fact that θ and r are not negatively collinear with respect to $W_1\cap W_{\theta-v}$ implies the existence of $\lambda''\in W_1\cap W_{\theta-v}$ with $\langle\lambda'',\theta\rangle\langle\lambda'',r\rangle<0$ (see Alonso and Camara 2015a), so that $L(\varepsilon;\lambda'')$ is locally concave at $\varepsilon=0$. ■

Proof of Proposition 3:

Lemma 3 above shows that if the projection of θ and r are not negatively collinear with respect to $W_1\cap W_{\theta-v}$ then persuasion is valuable. We now show that negative collinearity of θ and r with respect to $W_1\cap W_{\theta-v}$ is a non-generic property if $N\geq 4$. First note that $W_1\cap W_{\theta-v}$ has at least dimension $N-2$, and thus the projections of θ and r also have dimension $N-2\geq 2$. As collinearity is a non-generic property with vectors of dimension at least 2, this concludes the proof. ■

Proof of Proposition 4:

We will show that η_1^A and η_2^A from Corollary 1 are such that $\eta_1^A = -\infty < \eta_2^A < \infty$, which implies that partial information disclosure is optimal if and only if $\eta^A < \eta_2^A$. Condition **(C1)** implies that political disagreement is not maximized at the prior belief (Lemma B.1 in the on-line Appendix B). Proposition 1(b) in Appendix A shows that if political disagreement can be increased and $\lim_{\eta^B \to -\infty} \frac{f(\eta^B)}{F(\eta^B)} = \infty$, then $\eta_2^A > -\infty$. Proposition B.1(a) in the online Appendix B also shows that if D is concave (which is the case here) then $\eta_2^A < \infty$. Therefore η_2^A is finite. Full information disclosure is never optimal since it decreases disagreement to zero with certainty, thus $\eta_1^A = -\infty$. ■

Proof of Proposition 5:

Suppose $\beta^A\beta < 0$ or $|\beta^B| > 2|\beta^A|$. This implies that $\beta^A \neq \beta^B$ and $\beta^B(2\beta^A - \beta^B) < 0$.

We first show that, from the point of view of voter A, the expected policy payoff if candidate B is elected is strictly lower if the candidate observes a fully informative experiment, compared to no information. That is, full information makes candidate B choose a strictly worse policy on average.

Without further information, candidate B chooses policy $\beta^B E[\theta|p]$, which yields expected policy payoff $E[-(\beta^B E[\theta|p] - \beta^A\theta)^2|p]$ to voter A. With a fully informative signal, candidate B chooses policy $\beta^B\theta$ after learning that the state is θ. This yields an expected policy payoff $E[-(\beta^B\theta - \beta^A\theta)^2|p]$ to voter A. No information yields a strictly higher payoff than full information if and only if

$$\begin{aligned}
E[-(\beta^B E[\theta|p] - \beta^A\theta)^2|p] &> E[-(\beta^B\theta - \beta^A\theta)^2|p] \\
-(\beta^B)^2 E[\theta|p]^2 + 2\beta^A\beta^B E[\theta|p]^2 - (\beta^A)^2 E[\theta^2|p] &> -(\beta^B)^2 E[\theta^2|p] + 2\beta^A\beta^B E[\theta^2|p] - (\beta^A)^2 E[\theta^2|p] \\
(2\beta^A\beta^B - (\beta^B)^2) E[\theta|p]^2 &> (2\beta^A\beta^B - (\beta^B)^2) E[\theta^2|p] \\
0 &> \beta^B(2\beta^A - \beta^B)(E[\theta^2|p] - E[\theta|p]^2).
\end{aligned}$$

Since the variance $(E[\theta^2|p] - E[\theta|p]^2)$ is strictly positive given any interior prior belief, the inequality holds if and only if $0 > \beta^B(2\beta^A - \beta^B)$, concluding this step of the proof.

Disagreement is a convex function of the posterior belief, $D(q) = (\beta^B - \beta^A)^2(E[\theta|q])^2$. Consequently, if η^A is sufficiently low, then the IP's optimal experiment is fully informative.

From the point of view of voter A, compared to no information, full information leads candidate B to choose a worse policy on average, while it leads candidate A to choose a better policy when elected. Moreover, if η^A is sufficiently low, then candidate B is sufficiently likely to win the election, and the strictly negative effect of a worse policy from candidate B dominates the positive effect from the better policy from candidate A. ■

References

[1] Alonso, R. and O. Câmara (2015a): "Bayesian Persuasion with Heterogenous Priors," *mimeo.*

[2] Alonso, R. and O. Câmara (2015b): "Persuading Voters," *mimeo.*

[3] Bagnoli, Mark, and Ted Bergstrom(2005): "Log-concave probability and its applications," *Economic Theory*, 26(2), 445-469.

[4] Boleslavsky, R. and C. Cotton (2015): "Information and extremism in elections," *AEJ: Microeconomics*, 7(1), pp. 165-207.

[5] Bernecker, A., P. C. Boyer and C. Gathmann (2014): "Trial and Error? Reelection Concerns and Policy Experimentation during the US Welfare Reform," *mimeo.*

[6] Callander, S. (2011): "Searching for good policies," *American Political Science Review*, 105(04), pp. 643-662.

[7] Cover, T. M. and J. A. Thomas (2006): *Elements of Information Theory*, Second Edition, John Wiley & Sons.

[8] Dixit, A. K. and J. W. Weibull (2007): "Political polarization," *Proceedings of the National Academy of Sciences*, 104(18), pp. 7351-7356.

[9] Downs, A. (1957): *An Economic Theory of Democracy*, Harper and Row, New York.

[10] Gentzkow, M., and E. Kamenica (2014): "Costly Persuasion," *American Economic Review P&P*, 104(5), pp, 457-462.

[11] Groseclose, T. (2001): "A model of candidate location when one candidate has a valence advantage," *American Journal of Political Science*, 45(4), pp. 862-886.

[12] Iyengar, S. and A. F. Simon (2000): "New perspectives and evidence on political communication and campaign effects." *Annual review of psychology*, 51(1), pp. 149-169.

[13] Kamenica, E., and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, pp. 2590-2615.

[14] Kolotilin, A., M. Li, T. Mylovanov, and A. Zapechelnyuk (2015): "Persuasion of a Privately Informed Receiver," *mimeo.*

[15] Matsusaka, J. G. (1995): "Explaining voter turnout patterns: An information theory," *Public Choice*, 84(1-2), pp. 91-117.

[16] Millner, A., H. Ollivier, and L. Simon (2014): "Policy experimentation, political competition, and heterogeneous beliefs," *Journal of Public Economics*, 120, pp. 84-96.

[17] Prékopa, A. (1971): "Logarithmic concave measures with application to stochastic programming," *Acta Scientiarum Mathematicarum*, 32, pp. 301-316.

[18] Stokes, D. E. (1963) "Spatial models of party competition," *American Political Science Review*, 57(02), pp. 368-377.