Polarization through Persuasion

Jenny (Seoyeon) Kim*

Abstract

Policy experiments, which provide voters with information about a policy's effectiveness, are often implemented selectively. I propose a theoretical framework that explores the role of strategically designed policy experiments in shaping voter polarization. In the model, a politician seeking to maximize voter support tailors the informativeness of experiments based on her confidence in the policy's effectiveness and the predispositions of different voter groups. I show that this endogenous selection of experimental design intensifies polarization in districts with moderate initial polarization, but conversely reduces polarization in districts that were previously deeply divided. The model further examines the impact of asymmetric group sizes and incomplete voter attention.

*Department of Political Science, Emory University. Email: *seoyeon.kim@emory.edu*.

1 Introduction

Policy experiments allow voters to learn about a policy. Voters update their beliefs based on the result of a pilot project, which allows them to make more accurate judgments about the success probability of a policy. Importantly, however, a politician often has control over the design of a policy experiment, which determines *who* among voters will get *what* information (Alonso and Câmara, 2016; Arieli and Babichenko, 2019; Titova, 2022). For example, state legislatures often oversee pilot programs by selecting and granting awards to local municipalities or agencies that seek to administer policy experiments in their jurisdictions. Voters must then rely on the information strategically chosen by the politician who aims to maximize support for the policy.

Examples of such oversight range from simple construction projects—e.g., funding of on-demand microtransit service—to controversial issues such as police body cameras or universal basic income. For instance, as part of an effort to evaluate the impact of police body cameras on transparency, accountability, and public trust in law enforcement, the Massachusetts state government launched the Law Enforcement Body-Worn Camera (BWC) Program in 2021 and since offered competitive grants to equip police officers with body cameras. In 2023, Healey and Driscoll administration has awarded the New Bedford Police Department $250,000 of the $3.6 million in funding while a municipality similar in size, Fall River police, received $7,400. This highlights the endogeneity of pilot locations, i.e. information provision.

Why would politicians implement policy experiments in one district but not in another? In particular, what group of voters would they want to persuade the most? I propose a theoretical model where a politician who wants to maximize voters' support for a policy can utilize policy experiments to persuade them. The politician chooses how informative the policy experiment will be given his confidence in the policy's effectiveness and the voters' predispositions. I find that the politician's effort to maximize the share of support for a policy may lead to her conducting policy experiments only in selective districts. Under some conditions, such endogenous selection of information by the politician may exacerbate polarization among voters. In fact, districts with lower initial levels of polarization become more polarized as a result of the politician's endogenous choice of experiment. Polarization can arise from either a successful or a failed experiment. I extend the model further to account for variations in group sizes and the possibility of incomplete attention.

This project builds on the extensive literature on information control (Kamenica and Gentzkow, 2011; Schnakenberg, 2015; Alonso and Câmara, 2016; Arieli and Babichenko, 2019; Gitmez and Molavi, 2023) and targeted political advertising (Prummer, 2020; Titova, 2022). My theoretical framework contributes to this literature in two ways. First, I focus on variations in the *distribution* of receiver types. By doing so I attempt to answer the following questions: Is there a level of ideological polarization above which polarization feeds upon itself to become a runaway process? Under what conditions do politicians target districts with swing voters instead of focusing on their core supporters? In the model, voter distributions are exogenously determined; the politician chooses which districts to conduct pilot experiments in and, for each selected district, determines the accuracy of the experiments. This is different from previous works where candidates can target an arbitrary number of outlets (Prummer, 2020) or where the sender can segment the group of receivers and send different private messages to different coalitions (Titova, 2022).

2 Model Setup

A politician in a district wants to persuade voters to support a policy. The policy is either effective ($\omega = 1$) or not ($\omega = 0$). The voters' objective is to support policy if he believes it is effective and not otherwise, but they are uncertain about the true effectiveness of the policy. Specifically, the politician's prior belief about the efficacy of the policy is $b_P \in (0, 1)$. Additionally, let $b \in [0, 1]$ denote an individual voter's prior belief. There are two groups of voters in the district, Group 1 and Group 2. Group 1 generally believes that the policy is likely ineffective and is represented by distribution $b \sim \text{Uniform}[0, \bar{b}]$. Group 2—represented by distribution $b \sim \text{Uniform}[\underline{b}, 1]$—believes the policy will likely be effective. Let F_i denote Group i's cumulative distribution of belief and let $F = (F_1 + F_2)/2$ denote the voter's cumulative distribution of belief as a whole. This distributional assumption allows us to clearly examine between-group differences in a district.

Further define $s \in \{0,1\}$ as the voter's choice of whether or not to support the policy. Each individual voter wants to vote for the policy only when it is effective:

$$U_V = \mathbb{1}\{s = \omega\}.$$

The politician's objective is to maximize policy support:

$$U_P = \int_0^1 s(b)f(b)db.$$

Prior to taking a vote, the politician implements a policy experiment that signals the voters about the effectiveness of the policy. She chooses signal accuracy $k \in [0,1]$ which affects the experiment outcome $\sigma \in \{0,1\}$ that voters will observe. The signal is drawn as follows:

$$Pr(\sigma = \omega) = \frac{1+k}{2}.$$

The polarization between the two groups prior to the experiment can be expressed as a difference in prior medians.[1]

$$\rho = \frac{\underline{b}+1}{2} - \frac{\overline{b}}{2}.$$

In the following section, I solve for the politician's optimal choice of signal accuracy k given the prior distributions of two groups and examine its effect on the post-experiment level of polarization.

[1] I construct the actual polarization measure using the median rather than mean. The use of medians is the most conservative measure as it is the least influenced by party outliers, as suggested in the literature (e.g., Enders and Armaly, 2019; Shor, McCarty et al., 2022); no substantive results from any analysis differs employing the median measure.

3 Equilibrium Analysis

The solution concept is Perfect Bayesian Equilibrium. To solve this by backward induction, I first examine the voter's decision to support the policy and then determine what the politician's optimal choice of k is. An individual voter supports the policy if his or her belief about the effectiveness of the policy after the experiment is greater than 1/2 ($b > 1/2$). It follows that the voter supports the policy if

$$\mu(1) > 1/2 \Longleftrightarrow b > b^*(1) \equiv \frac{1-k}{2},$$
$$\mu(0) > 1/2 \Longleftrightarrow b > b^*(0) \equiv \frac{1+k}{2}.$$

The above inequalities state the conditions under which the voter supports the policy after a signal of 1 and 0, respectively. Note that $b^*(0)$ is always larger than $b^*(1)$, which implies that it is harder to persuade the voter after a signal of 0 than after a signal of 1.

The politician's utility as a function of choice k is

$$U_P(k) = Pr(\sigma = 1) \cdot [1 - F(b^*(1))] + Pr(\sigma = 0) \cdot [1 - F(b^*(0))],$$

which is equivalent to

$$\underbrace{\int_{b^*(0)}^{1} f(b)db}_{\text{guaranteed support}} + Pr(\sigma = 1) \underbrace{\int_{b^*(1)}^{b^*(0)} f(b)db}_{\text{conditional support}}.$$

The politician's utility is consisted of **guaranteed support** and **conditional support**. Guaranteed support represents the group of voters with $b > b^*(0)$ who always support the policy regardless of the outcome, and conditional support represents the group of voters with $b \in [b^*(1), b^*(0)]$ that supports the policy only if the experiment is successful.

Remark 1 An increase in k implies less guaranteed support but more conditional support.

Importantly, however, Remark 1 tells us that raising accuracy comes with a trade-off between guaranteed and conditional support. On the one hand, a less accurate experiment is a safe choice in terms of securing the support of voters who already believe that the policy is effective. On the other hand, a more accurate experiment is the relatively riskier choice that convinces voters who do not support the policy prior to the experiment.

Remark 2 An increase in b_p leads to an increase in the politician's choice of accuracy k^*.

Additionally, Remark 2 shows that an increase in the politician's prior belief about policy effectiveness incentivizes her to set a higher level of accuracy. The politician who is confident about the policy ($b_P > 1/2$) is willing to allow the signal to accurately convey the true policy effectiveness because she believes that it is highly likely to be effective; conversely, if the politician is not confident ($b_P < 1/2$), she believes the experiment is less likely to succeed and thus is less willing to set a higher accuracy.

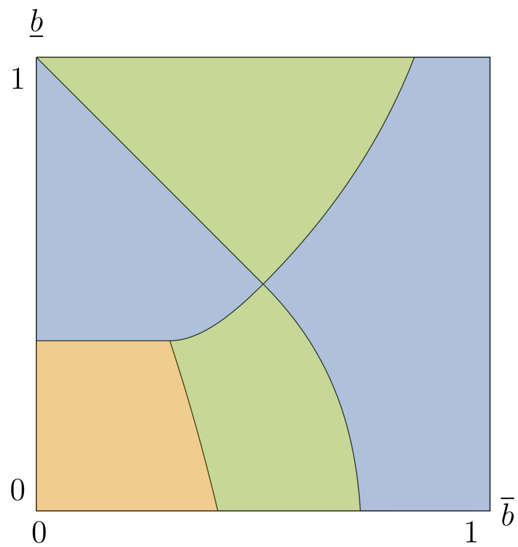


Figure 1: Equilibrium outcome depending on prior distribution
($b_P = 0.4$)

Figure 1 depicts how the politician's optimal choice of k depends on the bounds $\overline{b}$ and $\underline{b}$. Blue-, yellow-, and green-colored region refers to the politician's choice of no accuracy ($k^* = 0$), perfect accuracy ($k^* = 1$), and some accuracy ($k^* \in (0,1)$). The politician chooses no accuracy when $\underline{b}$ is moderate and $\overline{b}$ is either sufficiently small or large. Note that the support level of Group 2 is not sufficiently strong, so the politician is more cautious when choosing

k. For sufficiently small $\overline{b}$, Group 1 is too skeptical for the politician to try convincing. For sufficiently large $\overline{b}$, Group 1 is already somewhat in favor of the policy, so the politician does not want to risk losing support via a failed experiment. On the other hand, the politician chooses partial accuracy when $\overline{b}$ is moderate and $\underline{b}$ is either sufficiently small or large. The support level of Group 1 is moderate, so the politician is incentivized to persuade this unfavorable group via an informative experiment. For sufficiently small $\underline{b}$, the politician feels the need to persuade Group 2 as well. For sufficiently large $\underline{b}$—when Group 2 strongly believes that the policy is effective—a failed experiment will not significantly dissuade the voters that the politician is willing to undergo a more informative experiment. Lastly, the politician chooses perfect accuracy when both $\overline{b}$ and $\underline{b}$ are sufficiently small. This is when the joint prior distribution is skewed to the right, i.e., voters are not in favor of the policy in general. In this region, it is worth providing perfect information in order to significantly shift the voters' beliefs. The proposition below states the necessary conditions for the politician to choose certain k as the optimal accuracy in equilibrium.

Proposition 1 *Suppose that* $b_P < 1/2$. *Then,*

- *The necessary condition for the politician to choose* $k \in (0,1)$ *in equilibrium is* $\underline{b} < \min\{1-\overline{b}, 1/2\}$ *or* $\underline{b} > \max\{1-\overline{b}, \overline{b}\}$.
- *The necessary condition for the politician to choose* $k = 1$ *in equilibrium is* $\underline{b} < \min\{1-\overline{b}, 1/2\}$.

Suppose that $b_P > 1/2$. *Then,*

- *The politician never chooses* $k = 0$ *in equilibrium.*
- *The necessary condition for the politician to choose* $k \in (0,1)$ *in equilibrium is* $\underline{b} > 1-\overline{b}$.

Below I focus on one particular case where the politician is pessimistic about the effectiveness of the policy ($b_P = 0.1$) and Group 1 is very supportive of the policy prior to the experiment ($\underline{b} = 0.9$). I vary Group 2's prior level of support by shifting $\overline{b}$ and examine the politician's optimal choice of accuracy k^* and the resulting level of polarization between the two groups.

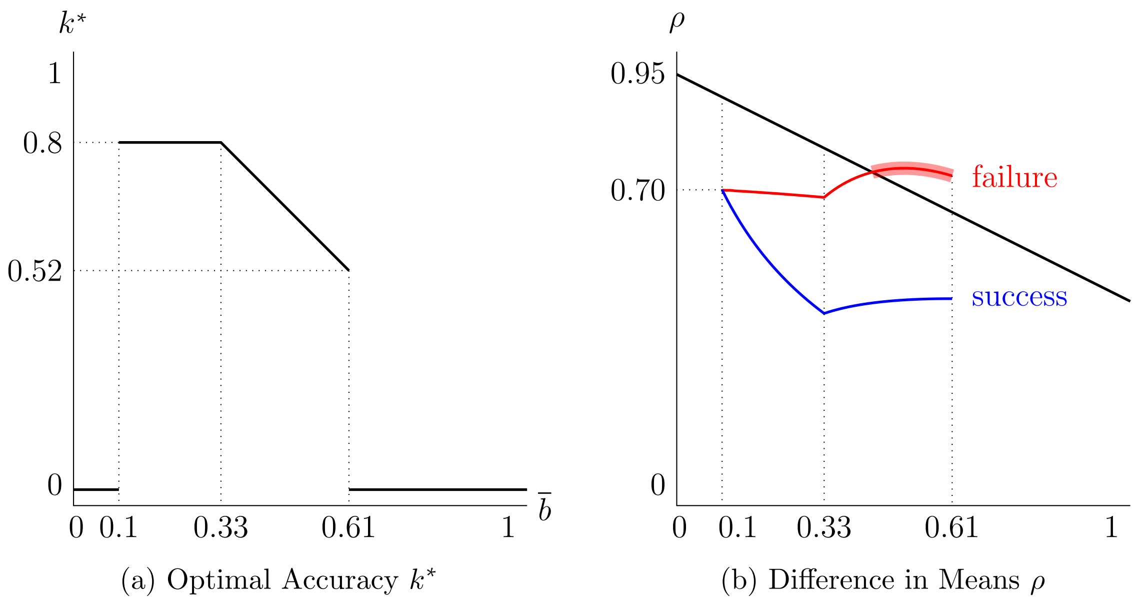


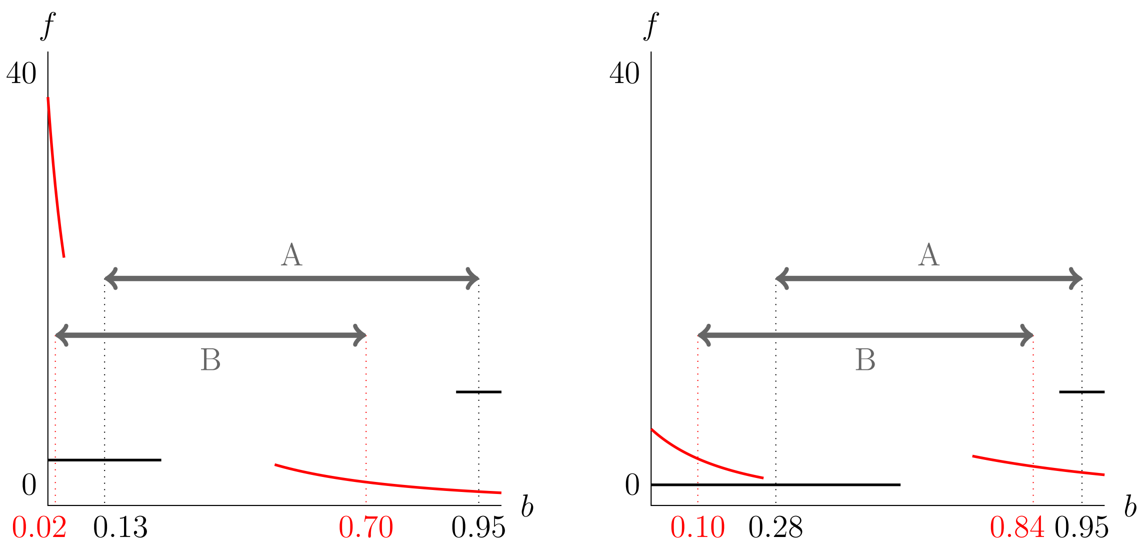
Figure 2: Optimal Choice of Accuracy and its Effect on Polarization
($b_P = 0.1$, $\underline{b} = 0.9$)

Figure 2a illustrates the politician's optimal choice of accuracy k with respect to Group 2's degree of support $\bar{b}$. Importantly, we observe that k^* is non-monotonic - for sufficiently low and high values of $\bar{b}$, the politician chooses zero accuracy $k^* = 0$, or equivalently, she conducts no experiment in equilibrium. For moderate values, the politician chooses some intermediate level of accuracy and this is the region where voters update their belief about the effectiveness of the policy after the experiment.

Figure 2b represents how equilibrium accuracy k^* affects the disparity in voters' posterior beliefs about the policy. The black line in the figure represents the prior difference-in-means between the two groups, while the blue and red lines represent the difference-in-means after the experiment's success and failure, respectively. If the blue or red line falls below the black line, this indicates that polarization has decreased after the experiment; if either line is above the black line, polarization has increased. Note that given these parameters, Group 1 is very supportive of the policy prior to the experiment ($\underline{b} = 0.9$). We see that the politician's choice of experiment *exacerbates* polarization among voters when the prior level of polarization between the two groups is moderate enough—the upper bound of Group 2's support $\bar{b}$ is moderate—but conversely *reduces* polarization when the groups were deeply divided prior to the experiment. The below proposition summarizes this result.

Proposition 2 *Let the distance between medians in Groups 1 and 2 measure the degree of polarization. It follows that*

- *The necessary condition for the failed experiment to cause more polarization is* $\underline{b} > \max\{\bar{b}, 1-\bar{b}\}$ *if* $b_P < 1/2$ *and* $\underline{b} > 1-\bar{b}$ *if* $b_P > 1/2$.
- *The necessary condition for the successful experiment to cause more polarization is* $b_P < 1/2$ *and* $\underline{b} < \min\{1/2, 1-\bar{b}\}$.



(a) High Disagreement ($\bar{b} = 0.25$) (b) Moderate Disagreement ($\bar{b} = 0.55$)

Figure 3: Examples of More/Less Polarization
($b_P = 0.1$, $\underline{b} = 0.9$)

Figures 3a and 3b illustrate this in more detail. Black lines represent the prior belief distribution of Groups 1 and 2; red lines represent the posterior distributions. 'A' and 'B' represent the prior and posterior levels of polarization between the two groups, respectively, measured by the difference-in-means. In Figure 3a where the prior level of disagreement between the two groups are high, $A > B$, meaning that the politician's choice of experiment reduces polarization. When the prior disagreement is lower, as can be seen in Figure 3b, the politician chooses a lower level of accuracy in equilibrium and this results in higher level of polarization between the groups ($B > A$).

Then what happens as a result of the experiment? To what extent does it lead to an increase in voter support? Below I compare the ex-ante support level for the policy before and after the experiment. Note that this is equivalent to the politician's utility, as the politician's objective is to maximize this value.

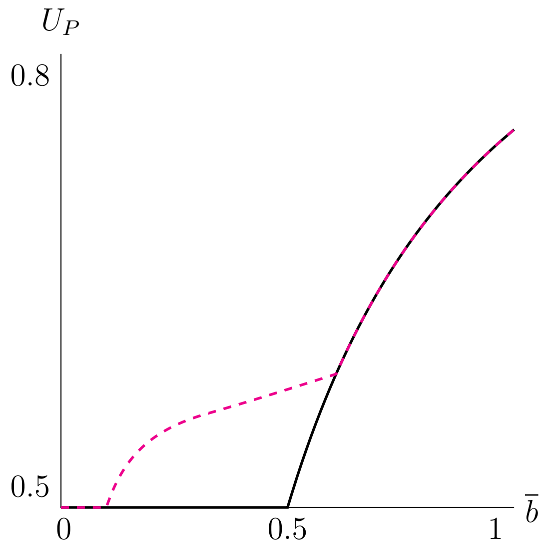


Figure 4: Ex-ante Support Level Before and After the Experiment
$(b_P = 0.1, \underline{b} = 0.9)$

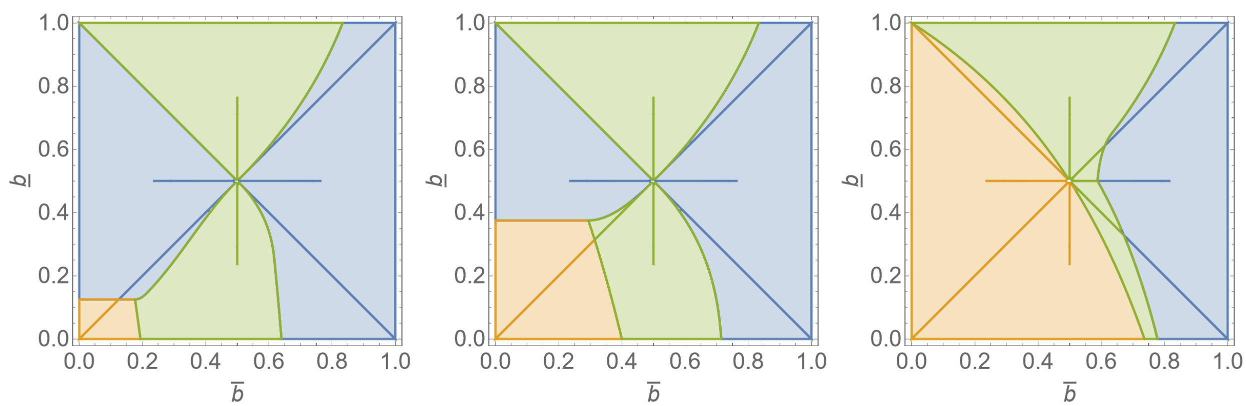
The black solid line represents the support level of voters prior to the experiment given $b_P = 0.1$ and $\underline{b} = 0.9$. For $\overline{b} \in [0, 0.5]$, the support level is constant at 0.5, as all voters in Group 1 does not support the policy; as $\overline{b}$ increases from 0.5, the overall support level monotonically increases. With the politician's endogenous choice of k^*, we see an increased ex-ante level of support for the policy. In Figure 2b, this refers to the region where either (1) both success and failure leads to less polarization or (2) the gains (of support) from success outweigh the loss from failure.

4 Extensions

In this section, I extend the model to incorporate several additional features of voters that affect the politician's optimal persuasion strategy.

4.1 Asymmetry in Population

The baseline model assumes that the proportion of Group 1 and Group 2 is equal. We could generalize this to account for scenarios where the relative proportion of one group is larger or smaller than the other. By allowing group sizes to differ, we are able to capture how the selective size of each group interacts with the characteristics of each group to influence the politician's choice to persuade. Let ω represent the relative group size of Group 1, which is the group predisposed against the politician.



(a) $\omega = 0.3$ (b) $\omega = 0.5$ (c) $\omega = 0.7$

Figure 5: Politician's Optimal Choice k^* for Different ω

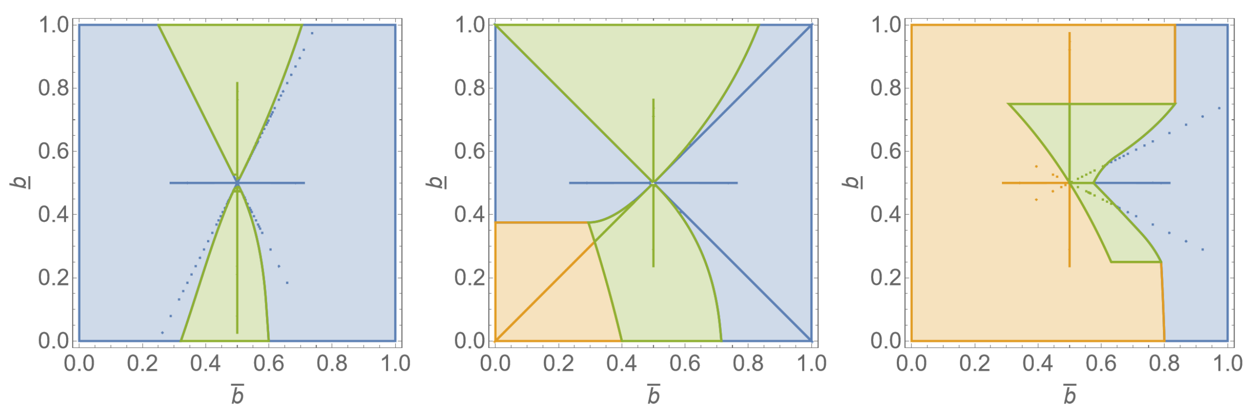
Figure 5 illustrates one example of how ω affects the politician's choice of accuracy. Note again that blue-, yellow-, and green-colored region refers to the politician's choice of no accuracy ($k^* = 0$), perfect accuracy ($k^* = 1$), and some accuracy ($k^* \in (0,1)$). From left to right, the values of ω are 0.3, 0.5, 0.7, respectively. We see that the region where the politician optimally provides some information (yellow- and green-colored regions) increase as ω increases. Based on this observation, we can make the following conjecture:

Conjecture 1 *Suppose $F = \omega F_1 + (1-\omega)F_2$. As ω increases, k^* increases.*

Simply put, this conjecture establishes that the politician aims to persuade the skeptics through more information but prefers to provide less information to the ones who already support the policy.

4.2 Scarce Attention

Additionally, I relax the presumption that the politician knows that voters will observe the outcome of the experiment with certainty. Instead, I suppose that each group of voters learns about the results of the policy experiment with some noise. This may be coming from voters' lack of attention or the politician's inability to persuade a certain group of voters. Importantly, even if the politician wants to fully reveal information, the voters receive less accurate information; formally, suppose the politician's optimal choice of k is discounted by $1 - \epsilon_i$ for Group i.



(a) $\epsilon_1 = 0.5, \epsilon_2 = 0$ (b) $\epsilon_1 = 0, \epsilon_2 = 0$ (c) $\epsilon_1 = 0, \epsilon_2 = 0.5$

Figure 6: Politician's Optimal Choice k^* for Different ϵ_i

Figure 6 now illustrates the politician's optimal choice of k for different values of ϵ_i. The left panel represents the case where $\epsilon_1 = 0.5$, i.e., Group 1 pays less attention to the politician's experiment; the middle panel is the case where $\epsilon_1 = \epsilon_2 = 0$; and the right panel is the case where $\epsilon_2 = 0.5$ and Group 2 conversely pays less attention. Note that as ϵ_1 decreases and ϵ_2 increases—equivalently, as the supporters pay less attention and skeptics pay more attention—the politician optimally generates more information.

Conjecture 2 *Suppose Group i receives $\sigma_i = \omega$.*

- *As ϵ_1 increases, k^* decreases.*

- *As ϵ_2 increases, k^* increases.*

The above conjecture highlights distinct incentives for the politician regarding Group 1 and Group 2. Specifically, the politician prefers to provide the skeptics with more information to persuade them to support the policy. In contrast, for those who already support the policy, the politician avoids "experimenting," as doing so could risk leading them to withdraw their support. This is similar to the trend in Section 4.1.

5 Conclusion

In this paper, I fully characterize the politician's optimal choice of experiment accuracy that reflect their confidence in the policy's effectiveness and voter predispositions and further examine its impact on polarization and the resulting vote share. I find that the prior level of division between the two groups of voters leads to more or less polarization through the politician's endogenous choice of *how much* information to generate to voters. In particular, an endogenous pilot experiment *exacerbates* polarization in districts with moderate prior polarization, but conversely *reduces* polarization when the groups were deeply divided prior to the experiment. A politician's objective to maximize the support for a policy may thus inadvertently lead to polarization among voters regarding their support for the policy.

My current measure of polarization only examines the distance between groups via difference-in-means. However, polarization manifests both *between* and *within* groups, where group members disagree with members of other groups or with their own, or both (Mehlhaff, 2024). Since the model assumes a uniform distribution for the two groups, the variance of the distribution for each group always weakly decreases after any experiment, which makes it difficult to observe the pilot experiment's effect on within-group polarization. Examining this problem with a different distribution, e.g., Beta(α, β), could capture both aspects of polarization. This allows us to examine whether a choice of accuracy that reduces between-group polarization may conversely increase within-group polarization, and what the politician's optimal choice is in equilibrium.

References

Alonso, Ricardo and Odilon Câmara. 2016. "Persuading voters." *American Economic Review* 106(11):3590–3605.

Arieli, Itai and Yakov Babichenko. 2019. "Private bayesian persuasion." *Journal of Economic Theory* 182:185–217.

Enders, Adam M and Miles T Armaly. 2019. "The differential effects of actual and perceived polarization." *Political Behavior* 41:815–839.

Gitmez, A Arda and Pooya Molavi. 2023. "Informational Autocrats, Diverse Societies.".

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101(6):2590–2615.

Mehlhaff, Isaac D. 2024. "A group-based approach to measuring polarization." *American Political Science Review* 118(3):1518–1526.

Prummer, Anja. 2020. "Micro-targeting and polarization." *Journal of Public Economics* 188:104210.

Schnakenberg, Keith E. 2015. "Expert advice to a voting body." *Journal of Economic Theory* 160:102–113.

Shor, Boris, Nolan McCarty et al. 2022. "Two decades of polarization in American state legislatures." *Journal of Political Institutions and Political Economy* 3(3–4):343–370.

Titova, Maria. 2022. Persuasion with verifiable information. Technical report working paper.

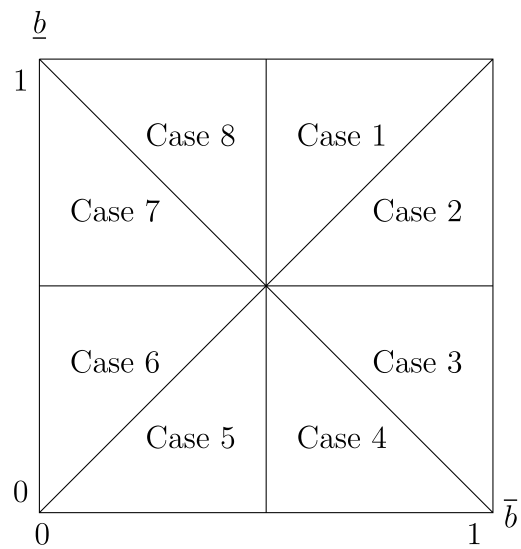


Figure A.1: Different cases depending on prior distribution shape

A Complete Characterization of Equilibrium

For a complete equilibrium analysis of the game, I consider sixteen distinct cases, defined by whether the politician is confident ($b_P > \frac{1}{2}$) and the shape of the prior distribution, as illustrated in Figure A.1.

A.1 Case 1: $1/2 < \overline{b} < \underline{b}$

Suppose $b_P < 1/2$.

Then, the politician's utility is decreasing in k for $k < 2\overline{b} - 1$; concave in k for $k \in [2\overline{b} - 1, 2\underline{b} - 1]$ with FOC satisfying k being

$$k_1^* \equiv \frac{-1 + \overline{b} + b_P - 2\overline{b}b_P}{-1 + 2b_P};$$

decreasing in k for $k > 2\underline{b} - 1$.

It follows that

1. If $k_1^* < 2\overline{b} - 1$, then the optimal choice is 0.

2. If $k_1^* \in [2\overline{b} - 1, 2\underline{b} - 1]$, then the optimal choice may be 0 or k_1^*.

3. If $k_1^* > 2\underline{b} - 1$, then the optimal choice may be 0 or $2\underline{b} - 1$.

Suppose $b_P > 1/2$.

Then, the politician's utility is increasing in k for $k < 2\underline{b} - 1$; convex in k for $k > 2\underline{b} - 1$ with FOC satisfying k being

$$k_2^* \equiv \frac{-1 + b_P + \underline{b} - b_P\underline{b} + \overline{b}(2 - 2\underline{b} + b_P(-3 + 4\underline{b}))}{(-1 + 2b_P)(1 + \overline{b} - \underline{b})}.$$

It follows that

1. If $k_2^* < 2\underline{b} - 1$, then the optimal choice is 1.

2. If $k_2^* \in [2\underline{b} - 1, 1]$, then the optimal choice may be 1 or $2\underline{b} - 1$.

3. If $k_2^* > 1$, then the optimal choice is $2\underline{b} - 1$.

A.2 Case 2: $1/2 < \underline{b} < \overline{b}$

Suppose $b_P < 1/2$.

Then, the politician's utility is decreasing in k for all k. It follows that the optimal choice is 0.

Suppose $b_P > 1/2$.

Then, the politician's utility is increasing in k for $k < 2\underline{b}-1$; convex in k for $k \in [2\underline{b}-1, 2\overline{b}-1]$ with FOC satisfying k being

$$k_3^* \equiv \frac{\overline{b}(1 - \underline{b} + b_P(-1 + 2\underline{b}))}{(-1 + 2b_P)(2 + \overline{b} - 2\underline{b})};$$

convex in k for $k > 2\overline{b} - 1$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_3^* < 2\underline{b} - 1$, then the optimal choice is 1.

2. If $k_3^* \in [2\underline{b} - 1, 2\overline{b} - 1]$, then the optimal choice may be 1 or $2\underline{b} - 1$.

3. If $k_3^* > 2\overline{b} - 1$ and $k_2^* < 2\overline{b} - 1$, then the optimal choice may be 1 or $2\underline{b} - 1$.

4. If $k_3^* > 2\overline{b} - 1$ and $k_2^* \in [2\overline{b} - 1, 1]$, then the optimal choice may be 1 or $2\underline{b} - 1$.

5. If $k_3^* > 2\overline{b} - 1$ and $k_2^* > 1$, then the optimal choice is $2\underline{b} - 1$.

A.3 Case 3: $\underline{b} < 1/2 < \overline{b}$ and $\underline{b} + \overline{b} > 1$

Suppose $b_P < 1/2$.

Then, the politician's utility is decreasing in k for all k. It follows that the optimal choice is 0.

Suppose $b_P > 1/2$.

Then, the politician's utility is increasing in k for $k < 1-2\underline{b}$; convex in k for $k \in [1-2\underline{b}, 2\overline{b}-1]$ with FOC satisfying k being k_3^*; convex in k for $k > 2\overline{b} - 1$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_3^* < 1 - 2\underline{b}$, then the optimal choice is 1.

2. If $k_3^* \in [1 - 2\underline{b}, 2\overline{b} - 1]$, then the optimal choice may be 1 or $1 - 2\underline{b}$.

3. If $k_3^* > 2\overline{b} - 1$ and $k_2^* < 2\overline{b} - 1$, then the optimal choice may be 1 or $1 - 2\underline{b}$.

4. If $k_3^* > 2\overline{b} - 1$ and $k_2^* \in [2\overline{b} - 1, 1]$, then the optimal choice may be 1 or $1 - 2\underline{b}$.

5. If $k_3^* > 2\overline{b} - 1$ and $k_2^* > 1$, then the optimal choice is $1 - 2\underline{b}$.

A.4 Case 4: $\underline{b} < 1/2 < \overline{b}$ and $\underline{b} + \overline{b} < 1$

Suppose $b_P < 1/2$.

Then, the politician's utility is decreasing in k for $k < 2\overline{b} - 1$; concave in k for $k \in [2\overline{b} - 1, 1 - 2\underline{b}]$ with FOC satisfying k being

$$k_4^* \equiv \frac{(1 - b_P + \overline{b}(-1 + 2b_P))(-1 + \underline{b})}{(-1 + 2b_P)(1 + 2\overline{b} - \underline{b})};$$

concave in k for $k > 1 - 2\underline{b}$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_4^* < 2\overline{b} - 1$, then the optimal choice is 0.

2. If $k_4^* \in [2\overline{b} - 1, 1 - 2\underline{b}]$, then the optimal choice may be 0 or k_4^*.

3. If $k_4^* > 1 - 2\underline{b}$ and $k_2^* < 1 - 2\underline{b}$, then the optimal choice may be 0 or $1 - 2\underline{b}$.

4. If $k_4^* > 1 - 2\underline{b}$ and $k_2^* \in [1 - 2\underline{b}, 1]$, then the optimal choice is k_2^*.

5. If $k_4^* > 1 - 2\underline{b}$ and $k_2^* > 1$, then the optimal choice is 1.

Suppose $b_P > 1/2$.

Then, the politician's utility is increasing in k for all k. It follows that the optimal choice is 1.

A.5 Case 5: $\underline{b} < \overline{b} < 1/2$

Suppose $b_P < 1/2$.

Then, the politician's utility is decreasing in k for $k < 1 - 2\overline{b}$; concave in k for $k \in [1 - 2\overline{b}, 1 - 2\underline{b}]$ with FOC satisfying k being k_4^*; concave in k for $k > 1 - 2\underline{b}$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_4^* < 1 - 2\overline{b}$, then the optimal choice is 0.

2. If $k_4^* \in [1 - 2\overline{b}, 1 - 2\underline{b}]$, then the optimal choice may be 0 or k_4^*.

3. If $k_4^* > 1 - 2\underline{b}$ and $k_2^* < 1 - 2\underline{b}$, then the optimal choice may be 0 or $1 - 2\underline{b}$.

4. If $k_4^* > 1 - 2\underline{b}$ and $k_2^* \in [1 - 2\underline{b}, 1]$, then the optimal choice may be 0 or k_2^*.

5. If $k_4^* > 1 - 2\underline{b}$ and $k_2^* > 1$, then the optimal choice may be 0 or 1.

Suppose $b_P > 1/2$.

Then, the politician's utility is increasing in k for all k. It follows that the optimal choice is 1.

A.6 Case 6: $\overline{b} < \underline{b} < 1/2$

Suppose $b_P < 1/2$.

Then, the politician's utility is decreasing in k for $k < 1 - 2\overline{b}$; concave in k for $k > 1 - 2\overline{b}$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_2^* < 1 - 2\overline{b}$, then the optimal choice is 0.
2. If $k_2^* \in [1 - 2\overline{b} - 1, 1]$, then the optimal choice may be 0 or k_2^*.
3. If $k_2^* > 1$, then the optimal choice may be 0 or 1.

Suppose $b_P > 1/2$.

Then, the politician's utility is increasing in k for $k < 1 - 2\underline{b}$; convex in k for $k \in [1 - 2\underline{b}, 1 - 2\overline{b}]$ with FOC satisfying k being

$$k_5^* \equiv \frac{1 - b_P}{-1 + 2b_P} + \underline{b};$$

increasing in k for $k > 1 - 2\overline{b}$.

It follows that

1. If $k_5^* < 1 - 2\underline{b}$, then the optimal choice is 1.
2. If $k_5^* \in [1 - 2\underline{b}, 1 - 2\overline{b}]$, then the optimal choice is 1.
3. If $k_5^* > 1 - 2\overline{b}$, then the optimal choice is 1.

A.7 Case 7: $\overline{b} < 1/2 < \underline{b}$ and $\underline{b} + \overline{b} < 1$

Suppose $b_P < 1/2$.

Then, the politician's utility is constant at 1/2 for $k < 2\underline{b} - 1$; decreasing in k for $k \in [2\underline{b} - 1, 1 - 2\overline{b}]$; concave in k for $k > 1 - 2\overline{b}$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_2^* < 1 - 2\overline{b}$, then the optimal choice is 0.
2. If $k_2^* \in [1 - 2\overline{b} - 1, 1]$, then the optimal choice is 0.
3. If $k_2^* > 1$, then the optimal choice is 0.

Suppose $b_P > 1/2$.

Then, the politician's utility is constant at 1/2 for $k < 2\underline{b} - 1$; convex in k for $k \in [2\underline{b} - 1, 1 - 2\overline{b}]$ with FOC satisfying k being k_5^*; increasing in k for $k > 1 - 2\overline{b}$.

It follows that

1. If $k_5^* < 2\underline{b} - 1$, then the optimal choice is 1.
2. If $k_5^* \in [2\underline{b} - 1, 1 - 2\overline{b}]$, then the optimal choice is 1.
3. If $k_5^* > 1 - 2\overline{b}$, then the optimal choice is 1.

A.8 Case 8: $\overline{b} < 1/2 < \underline{b}$ and $\underline{b} + \overline{b} > 1$

Suppose $b_P < 1/2$.

Then, the politician's utility is constant at 1/2 for $k < 1 - 2\overline{b}$; concave in k for $k \in [1 - 2\overline{b}, 2\underline{b} - 1]$ with FOC satisfying k being k_1^*; decreasing in k for $k > 2\underline{b} - 1$.

It follows that

1. If $k_1^* < 1 - 2\overline{b}$, then the optimal choice is 0.
2. If $k_1^* \in [1 - 2\overline{b} - 1, 2\underline{b} - 1]$, then the optimal choice is k_1^*.
3. If $k_1^* > 2\underline{b} - 1$, then the optimal choice is $2\underline{b} - 1$.

Suppose $b_P > 1/2$.

Then, the politician's utility is constant at 1/2 for $k < 1 - 2\overline{b}$; increasing in k for $k \in [1 - 2\overline{b}, 2\underline{b} - 1]$; convex in k for $k > 2\underline{b} - 1$ with FOC satisfying k being k_2^*.

It follows that

1. If $k_2^* < 2\underline{b} - 1$, then the optimal choice is 1.

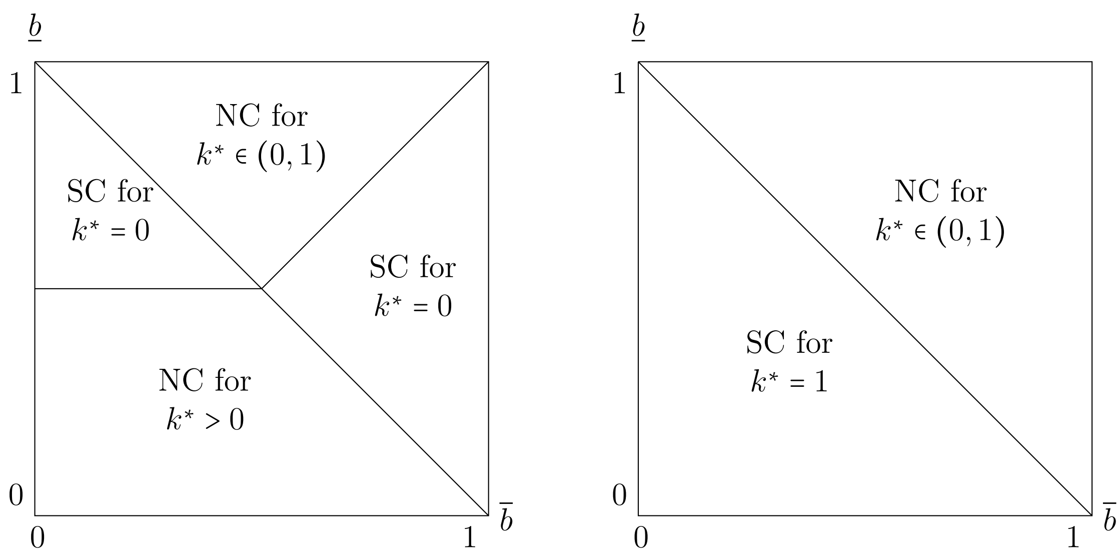
2. If $k_2^* \in [2\underline{b} - 1, 1]$, then the optimal choice may be 1 or $2\underline{b} - 1$.

3. If $k_2^* > 1$, then the optimal choice is $2\underline{b} - 1$.

To summarize,

	$b_P < 1/2$		$b_P > 1/2$	
	NC	SC	NC	SC
$k^* = 0$	all	case 2-3,7	none	none
$k^* \in (0,1)$	case 1,4-6,8	·	case 1-3,8	·
$k^* = 1$	case 4-6	·	all	case 4-7

Table A.1: Necessary/Sufficient Conditions for Optimal Choice of k



(a) $b_P < 1/2$ (b) $b_P > 1/2$

Figure A.2: Necessary/Sufficient Conditions for Optimal Choice of k

B Polarization Result

I use the distance between medians to measure the degree of polarization across two groups. The polarization prior to the experiment is

$$\rho \equiv \frac{\underline{b}+1}{2} - \frac{\overline{b}}{2}.$$

The polarization post the failure of the experiment is

$$\rho(0) \equiv \frac{(1+\underline{b})(1-k)}{(1+\underline{b})(1-k)+(1-\underline{b})(1+k)} - \frac{\overline{b}(1-k)}{\overline{b}(1-k)+(2-\overline{b})(1+k)}.$$

The necessary condition for $\rho(0) > \rho$ is that the politician in equilibrium chooses k_1^*; chooses $2\underline{b}-1$, or; if $b_P > 1/2$, chooses $1-2\underline{b}$.

The necessary condition for this is $b_P < 1/2$ and $\underline{b} > \max\{1-\overline{b}, \overline{b}\}$ or $b_P > 1/2$ and $\underline{b} > 1-\overline{b}$.

The polarization post the success of the experiment is

$$\rho(1) \equiv \frac{(1+\underline{b})(1+k)}{(1+\underline{b})(1+k)+(1-\underline{b})(1-k)} - \frac{\overline{b}(1+k)}{\overline{b}(1+k)+(2-\overline{b})(1-k)}.$$

The necessary condition for $\rho(1) > \rho$ is that the politician in equilibrium chooses k_2^*; chooses k_4^*, or; if $b_P < 1/2$, chooses $1-2\underline{b}$.

The necessary condition for this is $b_P < 1/2$ and $\underline{b} < \min\{1/2, 1-\overline{b}\}$.

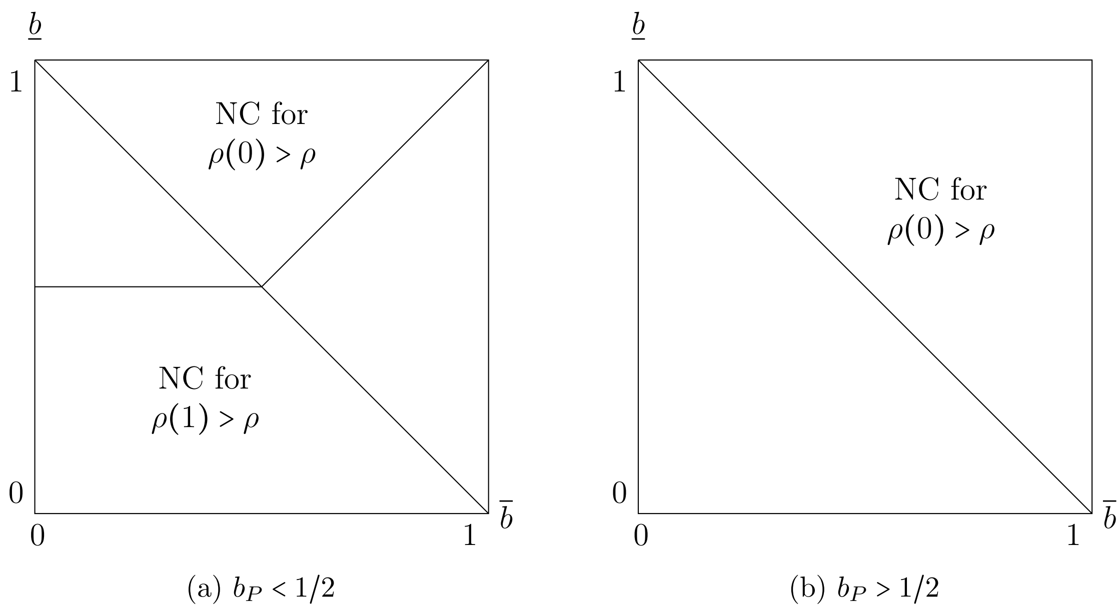


Figure B.3: Necessary/Sufficient Conditions for Optimal Choice of k