

Referenda and Partial Commitment in Policy Making ^{*}

Carl Heese [†]

We study *quasi-referenda*: two-stage decision processes in which a large number of privately informed agents each answer a binary question, and their collective response constrains the policy options from which a principal must then choose. Across all quasi-referenda, our analysis singles out procedures in which the collective acts as a gatekeeper that authorizes or blocks an extreme policy while otherwise delegating choice to the principal. These procedures maximize the principal's payoff guarantee (the worst-case payoff across all equilibria and information structures) within the class of all quasi-referenda, and when the policy space is fine, they robustly yield near-full-information outcomes. Our results provide a rationale for referenda used in practice and, more broadly, highlight commitment as an important dimension of political institutions.

Many collective decision processes, such as referenda or polls, serve not to specify a single policy to be implemented, but rather to reshape the space of policies available—opening the door to certain actions, blocking others. The final decision is left to a principal, who is free to select any policy within the constraints imposed by the agents.

For example, shareholders in a company may vote in a *cap referendum* to set a ceiling on executive pay, while leaving the design of the final pay package to the board. Similarly, in local public finance, citizens can vote on a cap override to determine whether officials may exceed a statutory spending limit, but leave the details of the budget to the officials. In other contexts, *gateway referenda* may trigger a departure from the status quo without committing to a specific alternative, as in Italy's abrogative referenda, which repeal an existing law or part of it without specifying the replacement, or as in constitutional referenda authorizing constitutional revision.¹

This paper develops a theory of such decision processes, which we call *quasi-referenda*. In a quasi-referendum, each of a group of agents takes a binary action (e.g., casts a yes/no

^{*}April 16, 2026. I thank Stephan Lauer mann, Wolfgang Pesendorfer, Wing Suen, Balázs Szentes, Marciano Siniscalchi, and audiences at the University of Vienna, the University of Hong Kong, NYU Shanghai, the Chinese University of Hong Kong, Hong Kong University of Science and Technology, the East Asia Theory Conference, and the 2025 CityU Microeconomics Conference for their valuable comments and suggestions. This work was supported by a grant from the University of Hong Kong (Projects 102010268 and 109001744) and from Hong Kong's Research Grant Council (ECS 27505725).

[†]The University of Hong Kong, Department of Economics, heese@hku.hk

vote); the principal observes the collective action (the average m of the individual actions); and m determines a set $Q(m)$ of feasible policies from which the principal then chooses. (For instance, in the shareholder example, $Q(m)$ excludes all pay packages above the cap unless m meets a certain cutoff.)

We allow the mapping Q to be arbitrary as long as it exhibits the defining feature of *partial commitment*: $Q(m)$ contains more than one policy, so that the collective action constrains but does not specify the outcome.² Our paper therefore studies collective decision processes under partial commitment. This shifts attention away from the familiar polar cases of full-commitment voting rules and non-binding cheap talk and toward the general question of how commitment affects policy outcomes.³

The main results provide a rationale for minimal commitment. In minimal-commitment processes, the collective acts as a gatekeeper for a single policy, being able to block or authorize only this policy. In the environments we study, we formalize a dual role of minimal commitment: it sustains robust aggregation of the agents' private information while it minimizes the principal's exposure to their coordination problems. As a result, the principal's worst-case outcomes are optimized. Gateway referenda and certain cap referenda emerge as the leading examples of minimal-commitment processes.

The baseline model considers a linear policy environment with common preferences under full information. A principal must choose a policy level x from a finite, ordered policy set. The policy has a common linear cost $c \cdot x$ and a common linear benefit $\omega \cdot x$ that depends on an unknown binary state ($\omega \in \{0, 1\}$).⁴ The principal would like to condition her policy choice on the state (choosing high x when the state is 1 and low x when the state is 0), but she does not observe the state. She therefore consults a population of N agents (where N is large) who hold private information in the form of noisy (conditionally independent) signals and heterogeneous prior beliefs. The agents participate in a quasi-referendum which results in a set of feasible policies for the principal to choose from.

The paper proceeds in four parts. The first two develop the logic behind the main results by identifying the coordination failures arising from partial commitment and characterizing when quasi-referenda robustly aggregate information. The third presents the main results on minimal commitment. The fourth discusses robustness and extensions.

¹The same logic characterizes city-level votes to initiate a planning process and neighborhood referenda to authorize redevelopment, as in the UK.

²We also study *generalized quasi-referenda*, which allow for commitment to either the status quo or a non-singleton set of policies. These decision processes formalize gateway referenda.

³Classic work on the Condorcet jury theorem studies voting rules that commit to a fully specified outcome as a function of votes, whereas cheap talk models assume no commitment. See, e.g., Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997), Bhattacharya (2018), and Krishna and Morgan (2012) for seminal contributions to the Condorcet jury theorem literature, and Wolinsky (2002), Morgan and Stocken (2008), Battaglini (2017), and Levit and Malenko (2011) for the cheap talk literature.

⁴Linearity is a standard assumption; compare for example to the classic public-goods formulation with linear cost and benefit in Ledyard and Palfrey (1999).

Partial commitment creates a strategic environment in which each agent has two incentives. His action transmits information and thereby affects the principal’s posterior beliefs; but it can also influence the set of feasible policies. In the first part of the paper, we show that the interaction of these two incentives gives rise to two *coordination failures*, that is, to two types of inefficient equilibrium sequences in which the agents miscoordinate on how to use their collective influence.

The first of these is as follows. One might expect that when the principal and the agents are largely aligned in their preferences, informative behavior leads to efficient outcomes. We show that this need not be the case: Near-truthful behavior often arises in equilibrium, with almost all agents matching actions to signals; however, such equilibria can be inefficient, because the agents’ collective action may rule out the ex-post optimal policy (Proposition 1).

The logic is easiest to see in a simple example. Suppose that in state 1 the highest policy is ex-post optimal, in state 0 the lowest policy instead, and that the quasi-referendum restricts the principal to low policies unless the collective action m exceeds a high cutoff—say, 85%—and to high policies otherwise. If the agents’ private information is informative but not too precise, then even in state 1 only a moderate majority of agents (say, around 60%) will receive a signal favoring high policies, so m misses the threshold with probability approaching one, as $N \rightarrow \infty$. Thus, even though the collective action is informative enough for the principal to learn the state (as $N \rightarrow \infty$), the induced constraint prevents her from acting on what she learns.

A second coordination failure arises if the principal is perceived as biased. In many applications, the agents may believe that the principal’s prior is systematically too conservative (or too aggressive) relative to a large part of the population. In that case, the agents may use the quasi-referendum not primarily to inform the principal, but to discipline her anticipated use of discretion by shifting the constraint she will face ex post. Our next result formalizes this logic. We consider a broad class of quasi-referenda in which the feasible policies $Q(m)$ increase monotonically at cutoff points. We show that when a sufficiently large share of agents is more optimistic about the value of the policy than the principal, *any* such quasi-referendum admits equilibrium sequences with *disciplining* behavior: Many agents choose the pro-policy action even when their private information points against it, in order to counteract the principal’s bias. With a large population ($N \rightarrow \infty$), m concentrates above the highest cutoff, so the set with the highest feasible policies is selected with probability approaching one (Proposition 2). In the example of the previous paragraph, with a cutoff of 85%, this prevents the principal from choosing low policies, so outcomes are inefficient in state 0.

Ex-post efficient equilibrium sequences often also exist (e.g., in the 85% example), but they require more delicate coordination than near-truthful or disciplining behavior. Both inefficiencies above thus demonstrate genuine failures to coordinate efficiently.

In the second part of the paper we characterize the informativeness of quasi-referenda. Despite the incentives for strategic exaggeration, quasi-referenda can still be robustly informative: In many quasi-referenda, the principal learns the state from the collective action along *every* equilibrium sequence, uniformly across a broad class of admissible information structures. We provide a sharp characterization of when this robust information aggregation obtains for an important class—monotone quasi-referenda with a single cutoff (Proposition 3). This result will imply that very limited commitment can be sufficient for learning, making minimal commitment procedures viable.

In the third part of the paper, we present the main results. We compare quasi-referenda in terms of their robustness. Our yardstick of comparison is the principal’s *payoff guarantee*: her worst-case payoff across all equilibria and all admissible distributions of priors and signals. This is a natural benchmark given the equilibrium multiplicity induced by partial commitment and the possible coordination failures, which are a common concern in the information aggregation literature (see, e.g., Ekmekci and Lauermann, 2020, and Ali, Mihm and Siga, 2025).

Across our broad class of quasi-referenda, we obtain a sharp performance ranking, led by two simple procedures: *collective vetoes of the maximum* (specific cap referenda) and *gateway referenda*. Their common feature is minimal commitment, where the collective acts as a gatekeeper for one polar option—either the maximal policy or the status quo—while deliberately leaving the choice between the remaining alternatives to the principal.

Under a collective veto of the maximum, the maximal policy is excluded if m falls short of a given cutoff; otherwise, the principal has full discretion. We show that collective vetoes of the maximum maximize the principal’s payoff guarantee. When the policy space is fine, they even yield near-full-information outcomes in every equilibrium (Theorem 1).

The key to this result is how commitment affects both information transmission and exposure to the agents’ coordination issues. With no commitment, actions are pure cheap talk, and babbling equilibria without any information transmission exist. With strong commitment—as in the 85% example above—miscoordination can leave the cutoff unmet, forcing the principal to make a quite inefficient choice. A collective veto of the maximum threads this needle via minimal commitment. The existence of *some* commitment makes the collective action consequential and thus robustly sustains learning (as Proposition 3 implies). On the other hand, the process leaves the principal almost full discretion, thereby minimizing exposure to coordination failures. At worst, coordination failures may rule out the maximum; when the policy grid is fine, this yields only a negligible loss.

A gateway referendum enforces the status quo if the collective action falls short of a certain cutoff; otherwise, it excludes the status quo but leaves the principal full discretion in choosing the alternative. To encompass these institutions, we generalize our definition of quasi-referenda to allow singleton commitment to the status quo ($Q(m) = \{x_1\}$) while maintaining non-singleton commitment sets $Q(m)$ otherwise. Gateway referenda satisfy

a theorem analogous to Theorem 1: The gateway referendum with the simple-majority cutoff $m_1 = \frac{1}{2}$ maximizes the principal’s payoff guarantee among all *generalized* quasi-referenda, and it robustly yields near-full-information outcomes when the policy space is fine (Theorem 2).

In the fourth part of the paper, we discuss robustness and several extensions. In Section 5.1, we provide conditions under which our optimality results persist under a weaker robustness criterion that fixes the information structure and takes the worst case only across equilibria. In Sections 5.2 and 5.3, we relax the assumptions on preferences. In Section 5.2, we expand the model to include a preference aggregation problem, allowing groups of agents to have opposing preferences depending on the state.⁵ This allows the principal’s preferences to conflict with those of some agents even under full information, thereby aggravating the incentives for strategic behavior. In Section 5.3, we relax the assumption of linear cost and benefit, instead considering players with ex-ante preferences that are single-plateau over the policy space. In Section 5.4, we show that the linear baseline model also has an interpretation in terms of stochastic commitment.

This paper makes two main contributions. First, it provides an analysis of commitment in referenda and related political institutions. A main insight from this analysis is that minimal commitment may be optimal under realistic informational frictions and imperfect commitment. Second, the logic of minimal commitment provides a rationale for two widely used procedures—cap referenda (including collective vetoes) and gateway referenda. In environments where information is dispersed across a large group but efficient collective action is difficult to coordinate, these referenda optimally navigate a central tension between information aggregation and exposure to coordination failures.⁶

Our model connects the two polar cases of full commitment and no commitment that have been used in the literature to study formal voting rules (e.g. Feddersen and Pesendorfer, 1997) and informal political processes such as protests (Battaglini, 2017). We discuss the related literature and benchmark results in detail in Section 6, and further related literature in Section 7. In particular, our analysis suggests that coordination problems may be more central to policy making than the full-commitment benchmark implies.

Although we study commitment in the context of referenda, the concept admits several broader interpretations. First, as in Battaglini (2017), commitment can be interpreted as a measure of how formal a political process is. From this perspective, our results imply that formality matters less for information aggregation than existing impossibility results

⁵Models with such preferences have been used to study distributive politics, in which the state of the world determines which group will benefit from a policy. See Fernandez and Rodrik (1991), Ali *et al.* (2025), and Bhattacharya (2018).

⁶These results are also relevant to current policy debates on democratic innovation (see, e.g., OECD, 2020) and to reforms aimed at expanding citizen participation through alternative institutions, such as the UK’s Localism Act 2011, which includes neighborhood planning referenda resembling gateway referenda.

may suggest (see Section 6). Second, commitment can be viewed as capturing how direct or indirect a democratic institution is. More commitment corresponds to less residual discretion for the principal and hence to a more direct institution; less commitment corresponds to more discretion and hence to a more indirect one. Third, partial-commitment processes may be understood as modeling mandates: Collective support confers a degree of political authorization for subsequent action. Stronger support translates into a broader mandate—that is, a larger set of policies from which the principal may choose—while weaker support yields a narrower mandate.⁷

1 Model and Incentives

A policy x needs to be chosen from a finite set of options $\mathcal{P} = \{x_1, \dots, x_l\}$ with $x_1 < x_2 < \dots < x_l$. To simplify the algebra, we let $x_1 = 0$, $x_l = 1$, and $x_2 = 1 - x_{l-1} = \varepsilon > 0$. The policy has a common and constant marginal cost of $c = \frac{1}{2}$, and a common and constant marginal benefit given by an uncertain state $\omega \in \{0, 1\}$. Thus, all players' payoff from x in ω is

$$x(\omega - \frac{1}{2}).$$

There is a principal who has a commonly known prior⁸

$$\frac{1}{2} < \Pr(\omega = 1) \leq 1 - \varepsilon.$$

In addition, there is a set of agents $\{1, \dots, N\}$ who hold private information.

- First, each agent is a non-strategic *partisan* with a positive probability and chooses a prescribed action. This happens with probabilities $0 < \rho_a < \frac{1}{2}$, for the prescribed actions $a = 0$ and $a = 1$, independent of the state. The partisans imply minimal noise and trembling-hand perfection (cf. the remark at the end of this section).⁹ With the remaining probability $(1 - \rho_1 - \rho_0)$, each agent is strategic.
- Second, each strategic agent i receives a binary private signal $s_i \in \{0, 1\}$ drawn independently from a common distribution conditional on the state, and satisfying

⁷The idea that officials have mandates to govern, and that these mandates are stronger for officials with greater support, has been explored before but without modeling a strategic principal; see, e.g., Herrera, Llorente-Saguer and McMurray (2019) and Damiano, Li and Suen (2025).

⁸ The right constraint ensures that the payoffs from an equilibrium without any information transmission cannot be arbitrarily close to the full-information payoffs. The left constraint is without loss. If instead $\varepsilon < \Pr(\omega = 1) < \frac{1}{2}$, the analogous main results hold. For example, under that condition, a collective veto of the minimum maximizes the principal's payoff guarantee within the class of all quasi-referenda.

⁹This approach is standard in the existing work on information aggregation, which similarly considers agents with partisan preferences and only equilibria in which they vote for their preferred policy (see, e.g., Bhattacharya, 2013; Feddersen and Pesendorfer, 1997). In particular, it allows us to discuss existing benchmark results within our framework.

$0 < \Pr(s_i = 1 \mid \omega = 0) < \Pr(s_i = 1 \mid \omega = 1) < 1$; so, signal 1 is an indication for state 1 and signal 0 an indication for state 0.

- Third, each strategic agent i holds a private prior belief $p_i \in [0, 1]$ about the likelihood of state 1; we refer to p_i simply as the agent i 's *type*. Types are drawn independently from a differentiable distribution F with support $[0, 1]$.

The distributions of the private signals and types (together with ρ_0 and ρ_1) constitute an *agents' information structure*. Given a *quasi-referendum* Q (defined below), the timing is as follows.

1. Each agent i observes his private information and takes a binary action $a_i \in \{0, 1\}$.
2. The principal observes the quantity $m = \frac{\sum_{i=1}^N a_i}{N}$, which we call the *collective action*, and chooses a policy $x \in \mathcal{P}$ subject to a constraint $Q(m)$ determined by the quasi-referendum Q .

A *quasi-referendum* is a left-continuous mapping from $[0, 1]$ to the set of all non-singleton subsets of \mathcal{P} , with at most finitely many discontinuities, and none at 0. A quasi-referendum Q thus maps a collective action m to a policy set $Q(m)$.¹⁰ Every quasi-referendum Q takes the form of a step function; that is,

- there exist a finite number of cutoffs $0 < m_1 < \dots < m_R < m_{R+1} = 1$
- and $Q(m)$ is constant on $I_1 := [0, m_1]$ and on $I_j := (m_{j-1}, m_j]$ for $j = 2, \dots, R + 1$.
- We denote the constant value on the interval I_j by Q_j .

Given a quasi-referendum Q , a principal's pure strategy is a sequence $\left(x(k)\right)_{k=0, \dots, N}$ mapping each possible realization $m = \frac{k}{N} \in [0, 1]$ to a policy $x(k) \in \left(Q(m)\right)$. We allow for randomization. A symmetric agents' strategy is a mapping $\sigma : [0, 1] \times \{0, 1\} \rightarrow [0, 1]$, where $\sigma(p, s)$ represents the likelihood that a non-partisan agent with prior p and signal s chooses action 1. Strategy profiles are denoted by η . The analysis that follows focuses on weak perfect Bayesian equilibria in symmetric agents' strategies. The presence of partisans implies trembling-hand perfection (Selten, 1988)—see the online appendix—and that the mean action for each state is interior, i.e.,

$$q(\omega'; \sigma) := \left((1 - \rho_1 - \rho_0) \mathbb{E} \left(\sigma(p, s) \mid \omega = \omega' \right) + \rho_1 \right) \in (0, 1) \text{ for } \omega' \in \{0, 1\}.$$

¹⁰The assumption that Q has at most finitely many discontinuities is without loss of generality for monotone quasi-referenda, i.e., those where $\min Q(m)$ and $\max Q(m)$ are weakly increasing, since the policy space \mathcal{P} is finite.

1.1 Principal's Best Responses

The principal's best response can be described via a single cutoff since it is driven by monotone Bayesian updating. Upon observing that k of the N agents have chosen action 1, the principal's posterior is¹¹

$$\frac{\Pr(\omega = 1 \mid k; \sigma, N)}{\Pr(\omega = 0 \mid k; \sigma, N)} = \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \binom{N}{k} \left(\frac{q(1; \sigma)}{q(0; \sigma)} \right)^k \left(\frac{1 - q(1; \sigma)}{1 - q(0; \sigma)} \right)^{N-k}.$$

If $q(1; \sigma) \geq q(0; \sigma)$, the posterior $\Pr(\omega = 1 \mid k; \sigma, N)$ is increasing in k . Since the principal's expected payoff from x is

$$x \left(\Pr(\omega = 1 \mid k; \sigma, N) - \frac{1}{2} \right),$$

and since her prior exceeds $\frac{1}{2}$, she can be indifferent after at most one k .

This allows us to characterize the best response in terms of a single cutoff \bar{k} : Either $\frac{1}{2} < \Pr(\omega = 1 \mid k; \sigma, N)$ for all $0 \leq k \leq N$ —in which case we set $\bar{k} = N$ —or there is a minimal \bar{k} with $-1 \leq \bar{k} < N$ such that¹²

$$\Pr(\omega = 1 \mid \bar{k}; \sigma, N) < \frac{1}{2} \leq \Pr(\omega = 1 \mid \bar{k} + 1; \sigma, N). \quad (1)$$

If $q(1; \sigma) < q(0; \sigma)$, the posterior is decreasing in k and \bar{k} is defined analogously.

The principal can only be indifferent at $k = \bar{k} + 1$. Any mixed best response thus can only involve mixing at $k = \bar{k} + 1$. In the remainder, we only consider principal's strategies that are best responses.

1.2 Agents' Best Responses

Fix a quasi-referendum Q with cutoffs (m_1, \dots, m_{R+1}) and a strategy profile η . An agent i 's best response is driven by two incentives. His action transmits information and thereby affects the principal's posterior beliefs and her preference (for 0 versus 1) but it can also influence which choices are feasible ex-post. Importantly, agent i 's choice affects the policy outcome x only if a *pivotal event* occurs, and what pins down the agents' incentives is the average effect on policy across multiple pivotal events.

Pivotality depends on the realized number k_{-i} of other agents choosing action 1. Let piv^k denote the event that $k_{-i} = k$.

- Since the principal's preference for 0 versus 1 switches at the cutoff \bar{k} , the events $\text{piv}^{\bar{k}}$ and $\text{piv}^{\bar{k}+1}$ are the only ones in which agent i 's choice possibly affects the principal's preference, and thereby the policy outcome.

¹¹We typically indicate posteriors of an agent i with the subscript i , e.g., $\Pr_i(\omega = 1 \mid p_i = p, s_i = s)$, but do not use a subscript for the principal's beliefs.

¹²We abuse notation here and set $\Pr(\omega = 1 \mid k = -1; \sigma, N) = 0$.

- In any event piv^k with $k = \lfloor m_j \cdot N \rfloor$ for $j = 1, \dots, R$,¹³ agent i 's choice changes the policy set available to the principal, and may thereby affect the policy outcome.
- In any other event, agent i 's choice does not affect the policy outcome.

An agent i with signal $s_i = s$ and prior $p_i = p$ compares the average policy effect in the two states, weighted by his posterior, and best-responds with action 1 if

$$\Pr_i(\omega = 1 \mid p_i = p, s_i = s) U(1; \eta) - \Pr_i(\omega = 0 \mid p_i = p, s_i = s) U(0; \eta) > 0, \quad (2)$$

For each state, the average effect $U(\omega'; \eta)$ summarizes the effect of his choice across pivotal events and realized strategies of the principal:

$$U(\omega'; \eta) := \sum_{k=0, \dots, N-1} \Pr(\text{piv}^k \mid \omega = \omega'; \eta, N) \cdot r(k) \text{ for } r(k) = \mathbb{E}(x(k+1) - x(k) \mid \eta). \quad (3)$$

We denote the pivotal events related to a shift of the principal's preference by $\text{piv}_0 = \text{piv}^{\bar{k}} \cup \text{piv}^{\bar{k}+1}$, and those related to a change in the feasible policy set by $\text{piv}_j = \text{piv}^{\lfloor m_j \cdot N \rfloor}$ for $j = 1, \dots, R$.

2 Coordination Failures

We show that in a large referendum, the interaction of the agents' two incentives gives rise to two *coordination failures*, that is, to two types of inefficient equilibrium sequences in which the agents miscoordinate on how to use their collective influence.

2.1 Approximately Truthful Equilibria

Given any quasi-referendum Q and any of its policy sets Q_j , we construct an equilibrium sequence in which the chosen policy set is Q_j with probability converging to 1 as $N \rightarrow \infty$. Thus, whenever Q_j excludes an ex-post optimal policy, i.e., $x = 0$ or $x = 1$, the principal is constrained to choose a suboptimal policy in at least one state.¹⁴

Proposition 1. *Consider any quasi-referendum with cutoffs $0 < m_1 < \dots < m_R < m_{R+1} = 1$. For any $j^* = 1, \dots, R+1$, there exist an agents' information structure and a sequence of equilibrium strategies $(\sigma_N)_{N \in \mathbb{N}}$ for which*

$$\lim_{N \rightarrow \infty} \Pr(m \in I_{j^*} \mid \sigma_N, N) = 1,$$

¹³Here, for any $z > 0$, $\lfloor z \rfloor$ denotes the largest non-negative integer that lies weakly below z .

¹⁴On the other hand, if Q never excludes $x = 1$ and also never excludes $x = 0$, then there is always an (inefficient) equilibrium without any information transmission: in this equilibrium, all non-partisan agents choose the same action and the principal chooses the prior-optimal policy $x = 1$, independent of the observed collective action.

hence the realized policy set is Q_{j^*} with probability approaching one.

As we will show in the online appendix, many quasi-referenda have also efficient equilibrium sequences for *all* agents' information structures. So Proposition 1 shows that the agents face a coordination problem: They may miscoordinate to achieve an inefficient equilibrium instead of an efficient one. For comparison, the Condorcet jury theorem (as in Bhattacharya (2013)) implies that when the principal has unlimited commitment power, simple majority voting between $x = 0$ and $x = 1$ implies efficient outcomes in *all* equilibrium sequences as $N \rightarrow \infty$ in our setting. That is, the possibility of miscoordination is a consequence of the partial commitment of quasi-referenda.

The formal proof of Proposition 1 is in the appendix. Here, we provide a sketch. The idea of the proof is to construct a sequence of equilibria in “approximately truthful” strategies. Formally, for $\delta > 0$, an agents' strategy σ is (δ -)approximately truthful if, for any given realized signal, a share of at least $1 - \delta$ non-partisan types match their action to their signal. Equilibria in approximately truthful strategies (which we call approximately truthful equilibria) exist for small enough δ and some agents' information structure with the following two properties. First, the agents' signals are relatively uninformative:

$$\Pr(s_i = 1|\omega = 1) - \Pr(s_i = 1|\omega = 0) = \gamma \quad (4)$$

for small enough $\gamma > 0$. Second, the agents' priors are relatively close to the principal's prior:

$$\Pr_F\left(p_i \in [p_1, \Pr(\omega = 1)]\right) > 1 - \frac{\delta}{4}, \quad (5)$$

for a certain bound p_1 .

The relevance of the two properties is best illustrated by connecting them to the scenario where the agents' actions are cheap talk and the prior is common. This is a pure common-value game, and, as such, it has an equilibrium in which *all* agents truthfully match their actions to their signals.

In the following, we sketch how, for small enough parameters $\gamma > 0$ and $\delta > 0$ and some information structures with the properties (4) and (5), the agents' incentives given any approximately truthful strategy approximate those in the common-value cheap-talk game. As the proof shows, this will imply the existence of an approximately truthful equilibrium.

The agents' incentives are driven by the pivotal events piv_j corresponding to the principal's and quasi-referendum's cutoffs. In the cheap-talk game, only the principal's cutoff \bar{k} and its limit $m_0 := \lim_{N \rightarrow \infty} \frac{\bar{k}}{N}$ matter for incentives, as $N \rightarrow \infty$. In a quasi-referendum, instead multiple cutoffs may be relevant, but the following explains, with an example, when only m_0 matters.

Intuitively, a cutoff m_j is more relevant in shaping the agents' incentives when it is

closer to the mean action, because in that case, the pivotal event in which the realized collective action equals the cutoff is more likely. It is therefore key to compare the distance of m_0 to the mean action in each state with that of the other cutoffs. We will do this via a bound $M > 0$.

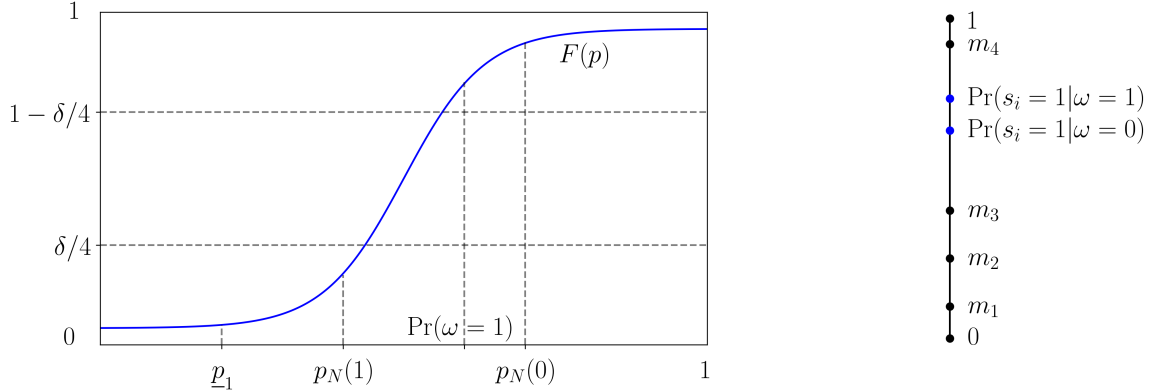


Figure 1: The agents' information structure, given by the prior distribution F (left) and the signal probabilities $\Pr(s_i = 1|\omega = \omega')$ (right). The approximately truthful equilibrium is given by the types $p_N(0)$ and $p_N(1)$ that are indifferent after signals 0 and 1, respectively: An agent i with signal $s \in \{0, 1\}$ chooses $a_i = 1$ if and only if $p_i \geq p_N(s)$.

Figure 1 depicts an example quasi-referendum and an agents' information structure satisfying (4) and (5) for small parameters $\gamma > 0$ and $\delta > 0$. The left panel shows the distribution of priors, nearly all of whose mass lies between the principal's prior and a bound $p_1 < \Pr(\omega = 1)$ close to it. The right panel shows the signal probabilities $\Pr(s_i = 1|\omega = 1)$ and $\Pr(s_i = 1|\omega = 0)$.

The signal probabilities are chosen together with the parameter $\gamma > 0$ so that they lie in between the quasi-referendum's cutoffs m_3 and m_4 , and with relatively large distance to these cutoffs, compared to γ .

The parameter $\delta > 0$ is chosen small enough so that any approximately truthful strategy σ_N has mean actions $q(\omega'; \sigma_N)$ so close to the signal probabilities that

$$q(1; \sigma_N) - q(0; \sigma_N) > \frac{\gamma}{2} \quad \text{and} \quad \left| m_j - q(\omega'; \sigma_N) \right| > M\gamma \quad \text{for } j \in \{3, 4\}, \omega \in \{0, 1\} \quad (6)$$

with $M > 0$ bounding the distance to the cutoffs m_3 and m_4 . The left ordering implies that the principal's posterior is strictly increasing in the number of observed actions 1. Her posterior becomes extreme (close to 0 or 1 as $N \rightarrow \infty$) for observations close to the mean $q(0; \sigma_N)$ or the mean $q(1; \sigma_N)$, respectively, by an application of the law of large numbers. Hence, the principal's best response cutoff (at which her posterior is not far from $\frac{1}{2}$) must lie in between the mean actions, ensuring

$$\left| m_0 - \lim_{N \rightarrow \infty} q(\omega'; \sigma_N) \right| \leq \gamma \quad \text{for } \omega' \in \{0, 1\}.$$

Comparing both displayed equations, one sees that M bounds the relative distance of

the principal's cutoff to the mean actions compared to the referendum's cutoffs. As the proof shows, when M is relatively large (as illustrated by the configuration in the figure), the likelihood ratio of the pivotal events becomes extreme when the number of agents grows large. The agents become almost certain to influence the principal's preference for low versus high policies (and not the policy set she chooses from), conditional on being pivotal, i.e.

$$\lim_{N \rightarrow \infty} \Pr_i(\text{piv}_0 | \text{piv}; \sigma_N, N) = 1. \quad (7)$$

In other words, the agents' incentives are almost the same as if their actions were cheap talk. For small $\delta > 0$, the players' prior distribution is close to a common prior, so that altogether incentives are close to those in the common value cheap-talk game.

The appendix then uses a formal fixed-point argument to construct an approximate truthful equilibrium, mirroring the truthful equilibrium of the common value cheap-talk game.

The left panel of Figure 1 shows the equilibrium strategy in terms of its cutoff types, namely, the agent types $p_N(1)$ and $p_N(0)$ that are indifferent after a signal 1 or a signal 0, with types above the cutoff choosing action 1 and types below action 0. Note that $p_N(1)$ lies below the δ -quantile and $p_N(0)$ above the $(1 - \delta)$ -quantile. Thus, the equilibrium is δ -approximately truthful.

Finally, the mean actions of the equilibrium lie strictly between m_3 and m_4 ; see (6). So, the realized collective action is in $I_4 = (m_3, m_4]$ and the realized policy set Q_4 with probability approaching one as $N \rightarrow \infty$, by an application of the law of large numbers. The proof of Proposition 1 extends the ideas of this example to show that *any* policy set Q_j of *any* quasi-referendum can become binding with probability approaching one.

2.2 Disciplining a Biased Principal

Differences between the principal's and the agents' priors can as well give rise to inefficiencies. If the principal is more pessimistic about the possible benefits from the policy than a sufficient mass of agent types, there are equilibria in which the agents miscoordinate on "disciplining" the principal's perceived bias: They constrain the principal to the highest policy range available, given the announced quasi-referendum, with probability close to 1. To state the result formally, we define $\bar{p} > \Pr(\omega = 1)$ as the type for which the type's posterior likelihood ratio conditional on signal 0 equals the principal's prior likelihood ratio multiplied with $\frac{\Pr(s_i=1|\omega=1) \Pr(s_i=0|\omega=0)}{\Pr(s_i=1|\omega=0) \Pr(s_i=0|\omega=1)} > 1$, and call a quasi-referendum *monotone* if both $\min Q(m)$ and $\max Q(m)$ are weakly increasing.

Proposition 2. *Consider any non-constant, monotone quasi-referendum with cutoffs $0 < m_1 < \dots < m_R < m_{R+1} = 1$. There is some $\bar{q} \in (m_R, 1 - \rho_0)$ such that, when*

$\rho_1 + (1 - \rho_0 - \rho_1)(1 - F(\bar{p})) > \bar{q}$, there is a sequence of equilibrium strategies $(\sigma_N)_{N \in \mathbb{N}}$ for which $\lim_{N \rightarrow \infty} \Pr(m_R < m \mid \sigma_N, N) = 1$.

The proof of Proposition 2 is in the appendix and constructs equilibrium sequences in which the mean action in both states exceeds the highest cutoff,

$$m_R < \bar{q} \leq q(0; \sigma_N) < q(1; \sigma_N) - \delta \text{ for some } \delta > 0. \quad (8)$$

This way, the highest cutoff becomes almost certainly binding by the law of large numbers, as the theorem claims.

The proof provides a formal fixed-point argument, showing that the best response is a self-map on the set of strategies satisfying (8). Here, we sketch it. Take any strategy satisfying (8) and a best response of the principal to it. Analogous to the argument in the previous Section 2.1, if \bar{q} is sufficiently high, the principal's cutoff is much closer to the mean actions $q(\omega'; \sigma_N)$ than the quasi-referendum's cutoffs. As before, the proof shows that then, the agents become almost certain to influence the principal's preference for low versus high policies (and not the policy set she chooses from), conditional on being pivotal; i.e. (7) holds.

In contrast to the previous Section 2.1, the principal and the agents' priors are not close. The condition $\rho_1 + (1 - \rho_0 - \rho_1)(1 - F(\bar{p})) > \bar{q}$ means that the principal is more pessimistic about the possible benefits from the policy than a large mass of non-partisan types. Consequently, it is not optimal for many agents to truthfully act in line with their signal. Instead, an agent's best response arises from the following strategic reasoning: conditional on the agent's action changing the principal's preference for low versus high policies, the principal must be close to indifferent. Now, if the agent's prior is sufficiently higher than the principal's, regardless of his signal, the agent is not indifferent but strictly prefers to influence the principal toward higher policies. The proof shows that indeed all non-partisan types with prior above \bar{p} best-respond by choosing action 1; the best response then again satisfies (8), given $\bar{q} < \rho_1 + (1 - \rho_0 - \rho_1)(1 - F(\bar{p}))$.

3 Information Aggregation

Despite the coordination problems unveiled in the previous two sections, many quasi-referenda possess robust information aggregation properties. We say that an equilibrium sequence *aggregates information* if the principal learns the state from observing the agents' collective action (with probability approaching one, as $N \rightarrow \infty$).

In this section, we characterize information aggregation for *monotone* quasi-referenda with a single cutoff. We identify two elementary properties of such quasi-referenda that jointly guarantee information aggregation uniformly across equilibria and (admissible) information structures (Proposition 3 below).

The first property requires the maximal feasible choice to increase in the collective action,

$$\max Q(0) < \max Q(1).$$

This rules out pure cheap talk. If the maximal feasible choice were constant, then there is always a trivial equilibrium without information transmission: all non-partisans choose the same action, so the collective action conveys no information; the principal then optimally chooses the same maximal feasible policy after every observation (given the prior preference for high policies); and this makes every agent type indifferent, sustaining the agents' uninformative behavior.

The second property rules out a “balanced” split of decision authority between the principal and the agent body. A single-cutoff quasi-referendum has *no balance* if

$$\max Q(0) \neq \min Q(1),$$

and it has *balance* otherwise. Balance admits equilibrium sequences that do not aggregate information. For example, consider

$$Q(m) = \begin{cases} \{0, \dots, \frac{1}{2}\} & \text{for } m \leq \frac{1}{2}, \\ \{\frac{1}{2}, \dots, 1\} & \text{for } m > \frac{1}{2} \end{cases}$$

Because $\frac{1}{2}$ is feasible on both sides of the cutoff, there exist *deadlock* equilibria in which the principal chooses $x = \frac{1}{2}$ after every observation, even though she updates her beliefs from the collective action: Once the principal plays such a constant strategy, all agents are indifferent between all strategies (their action has no effect on the policy outcome). What matters is that for large N , there is an agents' strategy that induces the principal's constant play as a best response. This way, the strategy together with the constant choice $x = \frac{1}{2}$ constitute an equilibrium.

The agents' equilibrium strategy has mean actions with $0 < q(1; \sigma_N) < q(0; \sigma_N) < \frac{1}{2}$ calibrated so that the principal is indifferent when the majority cutoff is just met, i.e.¹⁵

$$\Pr\left(\omega = 1 | k = \lfloor \frac{N}{2} \rfloor + 1; \sigma_N, N\right) = \frac{1}{2}.$$

Since $q(1; \sigma_N) < q(0; \sigma_N)$ the principal's posterior is decreasing in k . Thus, if she observes $k < \lfloor \frac{N}{2} \rfloor + 1$, she prefers high policies but can choose at most $x = \frac{1}{2}$. If she observes $k \geq \lfloor \frac{N}{2} \rfloor + 1$, she prefers low policies but must choose at least $x = \frac{1}{2}$. Thus, it is optimal for her to choose $x = \frac{1}{2}$ constantly.

Information does not aggregate, as $q(0; \sigma_N) < \frac{1}{2}$ for all N . This way, in state 0, the realized number k is smaller than $\lfloor \frac{N}{2} \rfloor + 1$ and the principal's posterior is greater than $\frac{1}{2}$,

¹⁵The online appendix formally shows existence of such a strategy.

with a non-vanishing probability. So the principal does not learn the state when it is 0.

To conclude, both properties—no balance and $\max Q(0) < \max Q(1)$ —are *necessary* for robust information aggregation. The following result shows that together they are also *sufficient*.

Proposition 3. *Consider any monotone quasi-referendum Q with a single cutoff $0 < m_1 < 1$ and any agents' information structure. Information aggregates in all equilibrium sequences if the quasi-referendum has no balance and $\max Q(0) < \max Q(1)$.*

The characterization is technically very demanding and in the appendix. Below, we present a sketch outlining the main steps.

First, the condition $\max Q(0) < \max Q(1)$ rules out the possibility of “uninformative” equilibria σ_N , which we define via the property $q(0; \sigma_N) = q(1; \sigma_N)$.

Suppose σ_N were uninformative. Then the principal learns nothing from the realized collective action; so her equilibrium choice depends only on feasibility: above the cutoff she selects the largest element of $Q(1)$, and below the cutoff she selects the largest element of $Q(0)$ (given her prior preference for high policies). Because $\max Q(0) < \max Q(1)$, an agent is then pivotal only at the referendum cutoff m_1 . The unformativeness of the equilibrium implies that this pivotal event has the same probability in both states, so any agent with a uniform prior is indifferent before observing his signal. His best response (and that of nearby types) is to match action to signal, implying on average higher actions in state 1, $q(0; \sigma_N) < q(1; \sigma_N)$, thereby contradicting the initial assumption of unformativeness.

Second, non-balance causes the logic of the deadlock equilibrium to fail. The deadlock equilibrium is supported by the fact that agents cannot affect the policy outcome. Non-balance eliminates this possibility: It implies that an agent's average effect $U(\omega'; \eta_N)$ on the policy outcome, given by (3), is non-zero and has the same sign in both states.

Lemma 1. *Consider any monotone quasi-referendum Q with a single cutoff $0 < m_1 < 1$, no balance and $\max Q(0) < \max Q(1)$. Any equilibrium η_N satisfies*

$$\begin{aligned} & \text{either } U(\omega'; \eta_N) > 0 \text{ for all } \omega' \in \{0, 1\}, \\ & \text{or } U(\omega'; \eta_N) < 0 \text{ for all } \omega' \in \{0, 1\}. \end{aligned}$$

To get an intuition, let us revisit the above referendum with balance, where $\frac{1}{2} = \max Q(0) = \min Q(1)$, and decrease $\max Q(0)$. The principal's best response to the agents' deadlock equilibrium strategy is no longer constant: she chooses $\max Q(0)$ for $m \leq m_1$ and $\min Q(1)$ for $m > m_1$, but these policies do not coincide any longer. The agents' actions now affect the policy outcome; an action 1 moves the policy upwards at the cutoff, so the average effect is positive, $U(\omega'; \eta_N) > 0$ for all $\omega' \in \{0, 1\}$.

Lemma 1 implies a simple cutoff form of the agents' equilibrium strategies. Any agent trades off the average effect in the two states, weighted by her belief about the states.

The agents' best response characterization (2) entails for each signal s a *unique* type $0 < p_N(s) < 1$ that is indifferent after observing s :

$$\frac{U(0; \eta_N)}{U(1; \eta_N)} = \frac{\Pr_i(\omega = 1 \mid p_i = p_N(s), s_i = s)}{\Pr_i(\omega = 0 \mid p_i = p_N(s), s_i = s)}. \quad (9)$$

Information aggregation follows once we show that the cutoff types do not drift to the extremes. In particular, it is enough to establish¹⁶

$$0 < \lim_{N \rightarrow \infty} p_N(1) < \lim_{N \rightarrow \infty} p_N(0) < 1. \quad (10)$$

Then, as $N \rightarrow \infty$, the mean action differs across signals and thus across the two states. Since the realized collective action concentrates around the mean action in each state, the principal learns the state from observing it.

Finally, the key observation underlying (10) is that, under partial commitment, a hypothetical failure of (10) would generically imply that, conditional on being pivotal, an agent becomes certain (as $N \rightarrow \infty$) that he is pivotal for the principal's preference for low versus high policies, that is,

$$\lim_{N \rightarrow \infty} \Pr_i(\text{piv}_0 \mid \text{piv}; \eta_N, N) = 1. \quad (11)$$

The intuition is as follows. If (10) failed, then all non-partisans would choose the same action, implying extreme limit mean actions

$$\lim_{N \rightarrow \infty} q(0; \sigma_n) = \lim_{N \rightarrow \infty} q(1; \sigma_N) \in \{\rho_0, 1 - \rho_1\}$$

The proof shows that then the principal's limit cutoff $m_0 = \lim_{N \rightarrow \infty} \frac{\bar{k}}{N}$ must coincide with the limit mean actions. Whenever $m_1 \notin \{\rho_0, 1 - \rho_1\}$, it is therefore much closer to them than to the referendum's cutoff m_1 . As already noted in Proposition 1's proof sketch, this relative closeness implies (11).

Given (11), the crucial point is that the principal's updating from piv_0 is bounded: It shifts her prior to a belief close to $\frac{1}{2}$, the indifference point. The agents' updating from piv_0 is therefore bounded as well. This, in turn, yields bounds on the limit indifferent types and hence implies (10).¹⁷ In conclusion, (10) holds generically.

The appendix provides the full proof, including that of Lemma 1, the precise comparison of pivotal probabilities and the remaining *non-generic* knife-edge cases in which cutoff limits may fail to be strictly interior.

¹⁶Throughout, whenever limits of sequences such as $p_N(s)$ or $q(\omega'; \sigma_N)$ are considered, we work along convergent subsequences; since these sequences lie in compact sets, this is without loss.

¹⁷Precisely, $\lim_{N \rightarrow \infty} \Pr_i(\text{piv}_0 \mid \text{piv}; \eta_N, N) = 1$ implies $\frac{U(0; \eta_N)}{U(1; \eta_N)} \approx \frac{\Pr_i(\text{piv}_0 \mid \omega=0; \eta_N, N)}{\Pr_i(\text{piv}_0 \mid \omega=1; \eta_N, N)}$, and the bounded updating from piv_0 , i.e., $\lim_{N \rightarrow \infty} \frac{\Pr_i(\text{piv}_0 \mid \omega=0; \eta_N, N)}{\Pr_i(\text{piv}_0 \mid \omega=1; \eta_N, N)} \in (0, \infty)$, then implies (10).

4 Main Results: Robustness of Minimal Commitment

We compare quasi-referenda in terms of their robustness. Our yardstick of comparison is the principal's *payoff guarantee*, the proportion of the full-information payoff that the principal obtains in the worst-case scenario, as $N \rightarrow \infty$. Formally, for a quasi-referendum Q , the principal's payoff guarantee is defined as

$$G(Q) := \inf_{(\eta_N)_{N \in \mathbb{N}}, \pi} \left(\liminf_{N \rightarrow \infty} \mathbb{E}(x \mid \omega = 1; \eta_N, N) - \frac{\Pr(\omega = 0)}{\Pr(\omega = 1)} \mathbb{E}(x \mid \omega = 0; \eta_N, N) \right),$$

where we take the infimum over all equilibrium sequences $(\eta_N)_{N \in \mathbb{N}}$ and all agents' information structures π .¹⁸

Our main results characterize quasi-referenda that maximize the principal's payoff guarantee; we call them *robust-optimal*. The characterizations are driven by a common logic, grounded in the observations from the preceding sections.

On the one hand, some commitment is necessary to make the collective action consequential and thereby sustain information transmission (compare to Proposition 3). On the other hand, stronger commitment than necessary makes outcomes more vulnerable to the agents' coordination failures (compare to Propositions 1 and 2). The robust-optimal quasi-referenda therefore turn out to use minimal commitment: the collective acts as a gatekeeper for one polar option, while the principal otherwise retains discretion.

4.1 Collective Vetoes of the Maximum

We first consider the baseline class of all quasi-referenda. In this class, the preceding logic singles out quasi-referenda in which the collective can exclude the maximal policy while the principal otherwise retains discretion. We call these *collective vetoes of the maximum*. They are examples of practically important cap-referendum procedures, such as shareholder votes on remuneration caps and budget- or tax-cap referenda in local public finance.

Formally, a *collective veto of the maximum* is a quasi-referendum Q of the form

$$Q(m) = \begin{cases} \mathcal{P} \setminus \{1\} & \text{if } m \leq m_1, \\ \mathcal{P} & \text{if } m > m_1. \end{cases} \quad (12)$$

¹⁸Here, \liminf denotes the smallest accumulation point of a sequence. For any equilibrium sequence $(\eta_N)_{N \in \mathbb{N}}$, the smallest accumulation point of the principal's payoff is $\liminf_{N \rightarrow \infty} \frac{1}{2} \left(\Pr(\omega = 1) \mathbb{E}(x \mid \omega = 1; \eta_N, N) - \Pr(\omega = 0) \mathbb{E}(x \mid \omega = 0; \eta_N, N) \right)$. When the principal knows the state, she can achieve the full-information payoff $\frac{1}{2} \Pr(\omega = 1)$. Dividing the former quantity by the latter and taking the infimum over all π and $(\eta_N)_{N \in \mathbb{N}}$ yields $G(Q)$. The appendix provides a general equilibrium existence result, which clarifies that the payoff guarantee of a quasi-referendum is well-defined: For *any* N , *any* quasi-referendum, and *any* agents' information structure, an equilibrium exists.

for some $m_1 \in (0, 1)$.¹⁹ Theorem 1 shows that collective vetoes are robust-optimal.

Theorem 1. *Any collective veto of the maximum has a payoff guarantee of $1 - \varepsilon$ and is robust-optimal among all quasi-referenda.*

Robust optimality requires balancing two forces. On the one hand, some commitment is necessary to make the collective action consequential: if a quasi-referendum Q *never* excludes the maximal policy $x = 1$, then there are uninformative equilibria in which the principal chooses $x = 1$ in both states, as argued in Section 3. This would imply an upper bound for the payoff guarantee of

$$G(Q) \leq 1 - \frac{\Pr(\omega = 0)}{\Pr(\omega = 1)} < 1 - \varepsilon.$$

(Here the last inequality holds because we assumed the principal's prior is not extreme, $\Pr(\omega = 1) \leq 1 - \varepsilon$; see Section 1.)²⁰

On the other hand, stronger commitment than necessary increases the principal's exposure to coordination failures. If a quasi-referendum excludes $x = 1$ after *some* collective action, then Proposition 1 implies equilibrium sequences in which the corresponding constrained policy set binds with probability approaching one. Along such sequences, the principal cannot choose $x = 1$ in state 1, and therefore chooses at most $1 - \varepsilon$. This yields an upper bound for the payoff guarantee of

$$G(Q) \leq 1 - \varepsilon$$

If the quasi-referendum were to exclude $x = 1$ and *additional* policies after some collective action, the same argument implies an even lower upper bound for the payoff guarantee.

Collective vetoes Q^* balance the two forces by using minimal commitment: the collective can only exclude the maximal policy, while the principal otherwise retains discretion. This ensures robust information aggregation—an immediate implication of the previous analysis (Proposition 3)—while minimizing the principal's exposure to miscoordination of the agents. Since the principal learns the state under Q^* and can choose any policy except $x = 1$ regardless of the agents' equilibrium play, in the worst-case, she chooses $x = 0$ in state 0 and $x = 1 - \varepsilon$ in state 1. Consequently, a lower bound for the payoff guarantee is

$$G(Q^*) \geq 1 - \varepsilon.$$

¹⁹More generally, in *cap referenda* the collective action decides which policy cap the principal has to obey. For low collective actions, a more stringent policy cap $x_c < x_C$ applies than for high collective actions. Depending on whether $m \leq m_1$ or $m > m_1$ for a cutoff $m_1 \in (0, 1)$, the set of feasible policies is either $Q(m) = [0, \dots, x_c]$ or $Q(m) = [0, \dots, x_C]$. Thus, collective vetoes are cap-referenda with $x_c = 1 - \varepsilon$ and $x_C = 1$.

²⁰The calculation is as follows: $1 - \frac{\Pr(\omega=0)}{\Pr(\omega=1)} \leq 1 - \frac{\varepsilon}{1-\varepsilon} < 1 - \varepsilon$.

Since the two inequalities above show that $1 - \varepsilon$ is an upper bound for the payoff guarantee of *every* quasi-referendum, collective vetoes are robust-optimal, with a payoff guarantee of $1 - \varepsilon$.

Two remarks are in order. First, the theorem also shows that collective vetoes achieve near full-information outcomes if the policy space is fine, i.e. if ε is small. Second, collective vetoes also maximize the agents' welfare guarantee given the common ex-post preferences: The agents' welfare guarantee is the ratio of the mean equilibrium payoff of an agent in the worst-case scenario to his mean full-information payoff. For a quasi-referendum Q , it is²¹

$$\inf_{(\eta_N)_{N \in \mathbb{N}}, \pi} \left(\liminf_{N \rightarrow \infty} \mathbb{E}(x \mid \omega = 1; \eta_N, N) - \frac{\mathbb{E}(1 - p_i)}{\mathbb{E}(p_i)} \mathbb{E}(x \mid \omega = 0; \eta_N, N) \right)$$

The agents' welfare guarantee arises by replacing the principal's prior with the agents' mean prior in the definition of $G(Q)$. The analogues of the inequalities above hold provided the agents' mean prior satisfies the same extremeness bound as the principal's prior, $\mathbb{E}(p_i) \leq 1 - \varepsilon$. Thus, the quasi-referenda defined by (12) are also agent-optimal except in these extreme scenarios.

4.2 Gateway Referenda

We now enlarge the class of decision processes by allowing commitment to the status quo (interpreted as the policy $x = 0$). Such commitment is feasible in many applications. In fact, many real-world referenda decide whether to retain the status quo or to mandate some change, while delegating the details of that change to a subsequent institutional procedure. A constitutional referendum, for example, may decide whether to keep the current constitution or to allow its revision without specifying the outcome.

We call a mapping Q a *generalized quasi-referendum* if for every $m \in [0, 1]$, either $Q(m) = \{0\}$ or $Q(m)$ is a non-singleton policy set. This relaxes the quasi-referendum requirement that $Q(m)$ be non-singleton for all m .

Our analysis singles out *gateway referenda*. These are generalized quasi-referenda of the form

$$Q(m) = \begin{cases} \{0\} & \text{if } m \leq m_1, \\ \mathcal{P} \setminus \{0\} & \text{if } m > m_1, \end{cases} \quad (13)$$

for some cutoff $m_1 \in (0, 1)$. Thus, if the cutoff is not met, the status quo is kept; if it is met, the status quo is excluded and the principal can choose any positive policy. Conditional on change, the referendum is therefore only minimally binding.

²¹To evaluate payoffs of partisan agents, we assume here and in the following that partisans hold extreme priors that conform with their prescribed choice $a \in \{0, 1\}$, i.e. $p_i = a$.

The next theorem shows that the simple majority gateway referendum is robust-optimal in this larger class of decision processes.

Theorem 2. *The gateway referendum with the simple-majority cutoff $m_1 = \frac{1}{2}$ has a payoff guarantee larger than $1 - \varepsilon$ and is robust-optimal among all generalized quasi-referenda.*

The proof is in the appendix. Here we highlight the key novel step relative to the argument for Theorem 1. What remains unchanged is the argument for information aggregation. The simple-majority gateway referendum allows the collective to retain or block the status quo and thereby gives sufficient incentives for information to aggregate in all equilibrium sequences. Formally, it is a single-cutoff process with no balance and with a strict increase in the maximal feasible policy across regimes, $\max Q(0) < \max Q(1)$, so the same argument as in the proof sketch of Proposition 3 implies information aggregation.²²

However, unlike in the case of collective vetoes, information aggregation does not by itself imply outcomes close to the ex-post optimum. The new difficulty is that, while the gateway referendum is only minimally binding conditional on change, it uses full commitment to the status quo $x = 0$ otherwise. If this commitment were to bind in the state where $x = 1$ is ex-post optimal, the outcome $x = 0$ would be far from optimal.

The reason why commitment to $x = 0$ does not obstruct robust optimality is that, under the simple-majority gateway referendum, it can bind asymptotically only in the state where $x = 0$ is ex-post optimal. To establish this, we show that in any equilibrium sequence the mean action in state 1 exceeds the majority threshold as $N \rightarrow \infty$. Hence the threshold is met with probability approaching one in state 1, and the status-quo commitment cannot bind there.

Lemma 2. *For any agents' information structure and any sequence of equilibrium strategies $(\sigma_N)_{N \in \mathbb{N}}$ of the gateway referendum with $m_1 = \frac{1}{2}$, it holds*

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) > \frac{1}{2}.$$

The logic behind Lemma 2 is closely related to that of the classical Condorcet jury theorem (Bhattacharya, 2013). The Condorcet jury theorem establishes that simple majority voting between $x = 0$ and $x = 1$ leads to ex-post optimal policies, as $N \rightarrow \infty$. In the standard majority-voting environment, one proves $\lim_{N \rightarrow \infty} q(1; \sigma_N) > \frac{1}{2}$ by contradiction. One assumes $\lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}$ and combines this with *two ingredients*: first, that the equilibrium mean actions satisfy $\lim_{N \rightarrow \infty} q(0; \sigma_N) < \lim_{N \rightarrow \infty} q(1; \sigma_N)$; and second, that agents' incentives are only driven by the pivotal event piv_1 at the majority

²²Proposition 3 was stated for quasi-referenda Q , but the relevant part of the proof uses only that Q has a single cutoff, no balance, and satisfies $\max Q(0) < \max Q(1)$. It does not use that every feasible set is non-singleton.

threshold $\frac{1}{2}$. Under these two conditions, the mean action is closer to the majority threshold in state 1, $\lim_{N \rightarrow \infty} |q(0; \sigma_N) - \frac{1}{2}| < \lim_{N \rightarrow \infty} |q(1; \sigma_N) - \frac{1}{2}|$. The standard argument implies then that pivotality at the majority threshold becomes an overwhelmingly strong indication for state 1, so that almost all non-partisans strictly prefer action 1, implying $\lim_{N \rightarrow \infty} q(1; \sigma_N) = 1 - \rho_0 > \frac{1}{2}$, a contradiction.

In our setting, neither ingredient is immediate. The ordering of the mean actions may be reversed in equilibrium (such as in the deadlock equilibria of Section 3), and the agents' incentives may also depend on the additional pivotal event piv_0 . The appendix shows, however, that under the simple-majority gateway referendum both ingredients are recovered asymptotically: the reverse ordering cannot arise in equilibrium, and under the contradiction hypothesis the additional pivotal event piv_0 becomes asymptotically irrelevant for the agents' incentives. The standard Condorcet-jury-theorem contradiction therefore applies and yields Lemma 2.

As a consequence of Lemma 2, the simple majority gateway referendum minimizes the principal's exposure to miscoordination and is robust-optimal: Since information aggregates and since the status-quo commitment does not bind in state 1, she chooses the ex-post optimal in state 1. The only potential loss arises in state 0 if the principal is forced to choose some $x > 0$; in this worst case, she then chooses the smallest feasible policy, $x = \varepsilon$. Consequently,

$$G(Q^*) \geq 1 - \frac{\Pr(\omega = 0)}{\Pr(\omega = 1)} \varepsilon > 1 - \varepsilon.$$

Similar constructions of approximately truthful equilibria, uninformative equilibria, and deadlock equilibria as in the proof of Theorem 1 and Proposition 3 rule out strictly higher payoff guarantees among any other generalized quasi-referenda, and establish the gateway referendum's robust-optimality. For example, one might consider $Q(m) = \{0\}$ for $m \leq \frac{1}{2}$ and $Q(m) = \mathcal{P}$ for $m > \frac{1}{2}$. However, this design has balance $\left(\max Q(0) = \min Q(1) \right)$ and therefore admits a deadlock equilibrium with constant policy outcome $x = 0$, implying a lower payoff guarantee.

5 Robustness and Extensions

5.1 A Weaker Robustness Criterion

The logic underlying the optimality results carries over when we weaken the robustness criterion and fix the simple majority cutoff. Under this weaker criterion, we fix the agents' information structure and compare quasi-referenda in terms of the principal's payoff guarantee across all equilibrium sequences, defined as in $G(Q)$ except that the infimum is taken only over equilibrium sequences under the fixed information structure.

For the comparison, we consider the simple class of monotone quasi-referenda with a single cutoff $m_1 = \frac{1}{2}$.

Take any such referendum and suppose $\Pr(\omega = 1) > \frac{1}{2}$, so that the principal's ex-ante preferred policy is the maximal one. As before, the key tension concerns how commitment affects both information transmission and exposure to agents' coordination issues.

First, some commitment is necessary to sustain information transmission. If $\max Q(0) = \max Q(1)$, there exists an equilibrium in which agents' behavior is uninformative and the principal implements the same maximal policy regardless of the collective action. Thus, avoiding such equilibria requires $\max Q(0) < \max Q(1)$.

Second, stronger commitment increases exposure to coordination failures. When the principal's prior is sufficiently high, an argument analogous to that underlying Proposition 2 implies equilibrium sequences in which $Q(0)$ binds with probability approaching one, as $N \rightarrow \infty$. Intuitively, as $\Pr(\omega = 1) \rightarrow 1$, Proposition 2's cutoff \bar{p} also approaches 1, implying $\rho_1 + (1 - \rho_0 - \rho_1) \left(1 - F(\bar{p})\right) < \bar{q}$ for any $\rho_1 < \bar{q} < \frac{1}{2}$; thus, an analogue of the proposition's sufficient condition is fulfilled (details are in the online appendix.)

Along sequences that make $Q(0)$ bind, the principal's maximal feasible policy in state 1 is $\max Q(0)$, implying that the payoff guarantee (across all equilibrium sequences) is bounded above by

$$\max Q(0) \leq 1 - \varepsilon.$$

Hence maximizing the payoff guarantee requires $\max Q(0)$ to be maximal subject to $\max Q(0) < 1$. The collective veto of the maximum fulfills this requirement and achieves a payoff guarantee of $1 - \varepsilon$ (Theorem 1). Hence, it is optimal among monotone quasi-referenda with cutoff $m_1 = \frac{1}{2}$, as the next proposition concludes.

Proposition 4. *Fix an agents' information structure. If the principal's prior is sufficiently high, the collective veto of the maximum maximizes the principal's payoff guarantee across all equilibrium sequences among all monotone quasi-referenda with a single cutoff $m_1 = \frac{1}{2}$.*

Now consider $\Pr(\omega = 1) < \frac{1}{2}$. Again, to rule out uninformative equilibria, the quasi-referendum must vary the minimum feasible policy, i.e. $\min Q(0) < \min Q(1)$. At the same time, when the principal's prior is sufficiently small, an argument analogous to that underlying Proposition 2 implies equilibrium sequences in which $Q(1)$ binds with probability approaching one. Along such sequences, the principal's minimal feasible policy in state 0 is $\min Q(1)$, implying that the payoff guarantee (across all equilibrium sequences) is bounded above by

$$1 - \frac{\Pr(\omega = 0)}{\Pr(\omega = 1)} \min Q(1) \leq 1 - \frac{\Pr(\omega = 0)}{\Pr(\omega = 1)} \varepsilon.$$

Hence maximizing the payoff guarantee requires $\min Q(1)$ to be minimal subject to $\min Q(1) > 0$. The simple majority gateway referendum fulfills this requirement and attains the upper bound $1 - \frac{\Pr(\omega=0)}{\Pr(\omega=1)}\varepsilon$.²³ Hence, it is optimal among monotone generalized quasi-referenda with cutoff $m_1 = \frac{1}{2}$.

Proposition 5. *Fix an agents' information structure. If the principal's prior is sufficiently low, the gateway referendum maximizes the principal's payoff guarantee across all equilibrium sequences among all monotone generalized quasi-referenda with a single cutoff $m_1 = \frac{1}{2}$.*

5.2 Heterogeneous Ex-Post Preferences

In the baseline model, all players agree on the best policy once the state is known. Such unanimity is often unrealistic. For instance, a policy with distributive consequences may benefit different groups depending on the realized state; see, e.g., Fernandez and Rodrik (1991) or Ali *et al.* (2025). Similarly, the principal's ex-post ranking of policies may conflict with that of a nontrivial share of agents.

We extend the model by allowing for state-dependent preference heterogeneity. In addition to partisans and "aligned" types who prefer higher policies in state 1 and lower ones in state 0, we allow for "opposed" types whose policy ranking reverses the principal's state-contingent ranking; they prefer lower policies in state 1 and higher ones in state 0. Their preferences are strictly opposite to those of the principal in each state, suggesting strong incentives to misrepresent or withhold information. Formally, a non-partisan's type is (p_i, \mathbf{t}_i) , where $p_i \in [0, 1]$ is her prior and $\mathbf{t}_i = (t_i(0), t_i(1)) \in [0, 1]^2$ are her state-dependent marginal benefits from the policy. The marginal cost $0 < c < 1$ is common, as in the baseline. The opposed types are those with $t_i(0) > c > t_i(1)$.

Despite the extended scope for misrepresentation of private information, *all* main results, in particular, the results about information aggregation and robust-optimality (Proposition 3, Theorems 1 and 2) extend to the heterogeneous-preference setting under a monotonicity condition for the agents' aggregate preferences (proofs are in the online appendix). A central observation is that, even with heterogeneous ex-post preferences, the aggregate best response can be summarized by a map Φ , which varies only in a one-dimensional incentive index. This map Φ is exogeneously given via the type distribution.

Let us explain this observation. Fix a strategy profile η and let $U(0; \eta)$ and $U(1; \eta)$ denote the average effect of an additional action 1 on the final policy in states 0 and 1, respectively. An individual agent type (p_i, \mathbf{t}_i) with signal s best-responds with action 1

²³Theorem 2 has established this payoff guarantee under the otherwise standing assumption that $\Pr(\omega = 1) > \frac{1}{2}$. The same proof, with minor and straightforward modifications, extends to the case $\Pr(\omega = 1) < \frac{1}{2}$.

if

$$U(0; \eta)(1 - p_i) \Pr(s_i = s | \omega = 0) (t_i(0) - c) + U(1; \eta)p_i \Pr(s_i = s | \omega = 1) (t_i(1) - c) \geq 0.$$

The aggregate best response can then be summarized via a map Φ , as follows. Fix s (equivalently, fix the likelihood ratio $l := \frac{\Pr(s_i=s|\omega=0)}{\Pr(s_i=s|\omega=1)} \in (0, \infty)$) and η . Let $\Phi(U(0; \eta), U(1; \eta), l)$ denote the probability that a randomly drawn type (p_i, \mathbf{t}_i) with signal s satisfies (14).

The aggregate best response (the value of Φ at $(U(0; \eta), U(1; \eta), l)$) turns out to depend only on the single incentive index

$$z_1 := \frac{U(0; \eta)}{U(1; \eta)} \cdot l.$$

whenever $U(1; \eta) \neq 0$. This is evident from rewriting the above best-response condition for individual types: If $U(1; \eta) > 0$, a type (p_i, \mathbf{t}_i) with signal s weakly prefers the action 1 if

$$z_1 \cdot (1 - p_i) (c - t_i(0)) \leq p_i (t_i(1) - c), \quad (14)$$

with the inequality reversed when $U(1; \eta) < 0$.

Based on this one-dimensional incentive representation, we sketch now how preference monotonicity matters. We say that preferences are *monotone* if, for each fixed sign of $U(1; \eta)$, the function Φ is continuously differentiable in z_1 and $\partial\Phi/\partial z_1$ has the same non-zero sign for all $z_1 \in (0, \infty)$.²⁴ Intuitively, monotonicity rules out non-monotone response patterns in which strengthening the incentive index z_1 could first increase and then decrease the share of types who favor the action 1; instead, aggregate best responses move in a single direction as incentives change.

Monotonicity of preferences allows us to modify the argument in Proposition 3's proof sketch, showing that monotone quasi-referenda with a single cutoff and $\max Q(0) < \max Q(1)$ only have informative equilibria.

Suppose an equilibrium were uninformative. Then the principal learns nothing from the realized collective action; so her equilibrium choice depends only on feasibility: above the cutoff she selects the largest element of $Q(1)$, and below the cutoff she selects the largest element of $Q(0)$ (given her prior preference for high policies). Because $\max Q(0) < \max Q(1)$, an agent is then pivotal only at the referendum cutoff m_1 . The unformativeness of the equilibrium implies that this pivotal event has the same probability in both states; thus the average effect of an additional action 1 is the same in both states, $U(1; \eta_N) = U(0; \eta_N) > 0$. However, given $U(0; \eta_N)/U(1; \eta_N) \in (0, \infty)$, the mean actions under the agents' best response satisfy $q(0; \sigma_N) \neq q(1; \sigma_N)$. This is a direct implication of (14) and the monotone likelihood ratio property of the signals, given preference

²⁴This monotonicity condition adapts Bhattacharya (2013)'s notion of "strong preference monotonicity."

monotonicity. It contradicts the initial assumption of uninformativeness.

This step establishing informativeness is key. Based on it, the same proofs as in the baseline analysis, with minor modifications, establish that all main results continue to hold in the heterogeneous-preference environment.

We conclude with three remarks. First, the monotonicity condition is satisfied in the baseline model, since all non-partisans share the same preference type. Second, non-monotonicity can overturn information aggregation of quasi-referenda for the same reason it can overturn the Condorcet jury theorem’s aggregation result (see Bhattacharya, 2013). Third, the robust-optimal quasi-referenda of the form (12) and (13) also maximize the agents’ ex-ante welfare guarantee, provided two additional conditions hold. The first is that the mean marginal benefit crosses the marginal cost across states,

$$0 \leq E(t_i(0)) < c < E(t_i(1)) \leq 1,$$

so that the principal and the mean agent share the same full-information ranking.²⁵ This is satisfied, for example, when the principal is a utilitarian social planner. Such incentives may arise from political economy forces under a broad set of conditions, cf. the literature on political agency and electoral accountability.²⁶ The second condition is

$$\frac{1 - E(p_i)}{E(p_i)} \cdot \frac{c - E(t_i(0))}{E(t_i(1)) - c} \geq \varepsilon,$$

which ensures that the agents’ mean payoff from the constant policy outcome $x = 1$ remains bounded away from the mean full-information payoff.

5.3 Non-Linear and Single-Plateau Preferences

We relax the baseline model’s assumption of constant marginal costs and benefits and consider non-linear payoffs $u(x, \omega)$. This extension broadens the scope of applications by allowing intermediate policies to be optimal. The analysis is substantially more involved and is therefore deferred to the online appendix. Here we introduce the setup, state the result established there, and briefly explain the additional difficulties relative to the baseline model.

A natural example captured by the extension is public-good provision with diminishing returns to scale, where providing the good is desirable but full capacity is inefficient. Similar examples arise in applications where compromises between opposing extremes are desirable.

²⁵To evaluate payoffs of partisan agents, we assume here that partisans hold extreme priors that conform with their prescribed choice $a_i \in \{0, 1\}$, i.e. $p_i = a_i$.

²⁶For the political agency literature, see Barro (1973); Ferejohn (1986) among others; for the literature on electoral accountability, see, e.g., the survey in Ashworth (2012). Relatedly, Battaglini (2017) provides an excellent discussion, with several explicit examples.

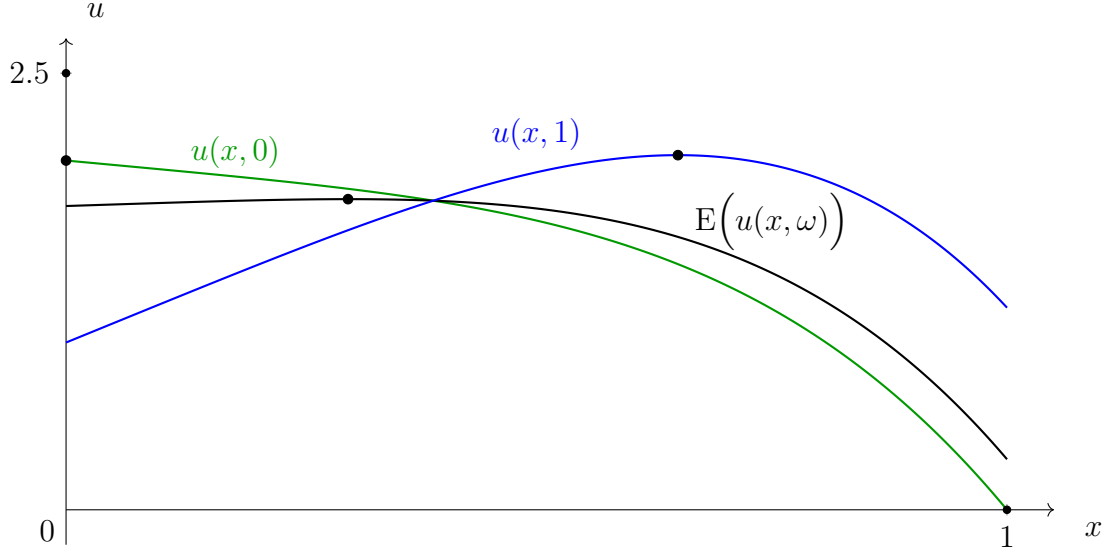


Figure 2: An example of non-linear and single-plateau preferences: $u(x, \omega) = -c(x) + b(x)\omega$ with $c(x) = -2 + \frac{1}{2}x + \frac{3}{2}x^4$ and $b(x) = -\frac{667}{640} + \frac{27}{10}x - \frac{1}{2}x^4$, so that $\frac{c'(x)}{b'(x)} = \frac{5+60x^3}{27-20x^3}$. The black dots show the peaks of the expected utility in state 0, in state 1, and given a prior of $\Pr(\omega = 1) = \frac{1}{4}$.

We consider common payoffs $u(x, \omega)$, continuously differentiable in x , such that

- in state 0, the status quo $x = 0$ is optimal, with $u'(x, 0) < 0$ for all $x \in [0, 1]$;
- in state 1, a unique non-zero policy is optimal.

Letting $c(x) = -u(x, 0)$ and $b(x) = u(x, 1) - u(x, 0)$, we can express the payoffs as before in terms of cost and benefits,

$$u(x, \omega) = -c(x) + b(x)\omega \quad \text{for } \omega \in \{0, 1\}.$$

The expected utility given a fixed prior p has a positive derivative if $\frac{c'(x)}{b'(x)} \leq p$. We focus on the case of “diminishing returns to scale,” where²⁷

$$\frac{c'(x)}{b'(x)} \text{ is strictly increasing.}$$

Under this condition, each player’s ex-ante expected utility is single-plateau, with the peaks increasing in p . As in the baseline model, we assume that, given the principal’s prior, there is a unique, non-zero optimal policy. Figure 2 illustrates the preferences in an example.

The online appendix shows that the simple-majority gateway referendum still aggregates information and achieves a payoff guarantee above $1 - \varepsilon$; see Theorem 3 there. We conjecture that the coordination-failure results of Proposition 1 and Proposition 2, as

²⁷The baseline model assumed that $\frac{c'(x)}{b'(x)}$ is constant.

well as the robust optimality of the simple-majority gateway referendum, also extend to this environment. To keep the exposition focused, we confine attention here to the result just stated.

The extension is not straightforward. In the baseline model, the principal's best-response correspondence can be described via a single cutoff \bar{k} . Under single-plateau preferences, by contrast, there may be multiple cutoffs $\bar{k}_{j+1,j}$ at which the principal's preference between policies x_j and x_{j+1} switches. This complicates the principal's side of the analysis.

It also complicates the agents' best response problem, which compares how, in each state, choosing action 1 (instead of 0) affects the principal's policy choice and thereby average payoffs

$$U(1; \eta) := \mathbb{E}\left(u(x, 1) | a_i = 1; \eta, N\right) - \mathbb{E}\left(u(x, 1) | a_i = 0, \eta, N\right), \quad (15)$$

$$U(0; \eta) := -\left(\mathbb{E}\left(u(x, 0) | a_i = 1; \eta, N\right) - \mathbb{E}\left(u(x, 0) | a_i = 0, \eta, N\right)\right), \quad (16)$$

cf. (2). In addition, because payoffs are now non-linear in x , a policy switch from x_j to x_{j+1} need not have proportional implications across states. In the linear baseline model, by contrast, the corresponding effects were proportional. Nevertheless, as the online appendix shows, the strict monotonicity of $\frac{c'(x)}{b'(x)}$ structures equilibrium incentives for the agents and the principal in a way that ensures information aggregation, and, with it, a payoff guarantee above $1 - \varepsilon$ for the simple-majority gateway referendum.

5.4 Stochastic Commitment: An Interpretation of the Linear Preference Model

The linear baseline model can also be interpreted as a model of supermajority voting between the policies $x' = 0$ or $x' = 1$ under stochastic commitment. The stochastic model retains the common payoff specification

$$x'(\omega - \frac{1}{2}) \text{ for } x' \in \{0, 1\} \text{ and } \omega \in \{0, 1\},$$

as well as the specification of the agents' side from the baseline model. The only change is in how outcomes are determined. With probability $1 - \varepsilon'$ the principal is bound by a supermajority rule and must choose $x' = 0$ if $m \leq m_1$ and $x' = 1$ if $m > m_1$. With the remaining probability $\varepsilon' > 0$, however, she is unconstrained and can choose either $x' = 0$ or $x' = 1$ after observing the vote share $m = \frac{k}{N}$.

This implies that, after any vote share m , a pure uncommitted choice induces an

outcome probability

$$x(k) = \Pr(x' = 1|k) \in \begin{cases} \{0, \varepsilon\} & \text{if } m \leq m_1, \\ \{1 - \varepsilon, 1\} & \text{if } m > m_1. \end{cases}$$

Hence, the principal's set of pure strategies is the same as in the baseline model with quasi-referendum

$$Q(m) = \begin{cases} \{0, \varepsilon'\} & \text{if } m \leq m_1, \\ \{1 - \varepsilon', 1\} & \text{if } m > m_1. \end{cases}$$

Because the common payoffs are linear in the probability of implementing $x' = 1$, corresponding principal strategies have the same payoff consequences in the two formulations. It follows that the two models are strategically equivalent and therefore have the same set of equilibria.

6 Discussion: Commitment in Policy Making

We relate our analysis of partial commitment to two strands of the information-aggregation literature: (i) models with full commitment to voting rules that map vote profiles into single policies, and (ii) models in which collective input is cheap talk (no commitment), which the literature has used to study informal political processes such as protests. These two strands provide natural benchmarks for our setting. Relative to the first, our results show that coordination failures are more central to policy making under dispersed information than the full-commitment benchmark suggests. Relative to the second, they show that formality (commitment) matters less for information aggregation than existing results suggest.

Under full commitment, classic results such as the Condorcet jury theorem (CJT) and its extensions show that, as $N \rightarrow \infty$, majority voting leads to the same outcome as under full information with probability approaching one under broad conditions (see, e.g., Bhattacharya, 2013; Feddersen and Pesendorfer, 1997). These results provide a benchmark in which the strategic coordination problem among agents disappears in large groups. Our analysis implies that the full-information-equivalence benchmark result is fragile to arbitrarily small departures from full commitment. Here and throughout, we distinguish information aggregation (learning the state) from full-information equivalence of outcomes.

This fragility can be seen in the most stark way when recalling the interpretation of our baseline model as one of majority voting between $x' = 0$ or $x' = 1$, but with stochastic commitment (see Section 5.4): With probability $1 - \varepsilon'$ the principal is bound by a majority rule and has to choose $x' = 0$ if $m \leq m_1$ and $x' = 1$ if $m > m_1$. With

probability $\varepsilon' > 0$, however, she best-responds by choosing either $x' = 0$ or $x' = 1$. As discussed, this stochastic model is equivalent to our linear baseline model with the quasi-referendum given by $Q(m) = \{0, \varepsilon'\}$ for $m \leq m_1$ and $Q(m) = \{1 - \varepsilon', 1\}$ for $m > m_1$, with outcomes $x \in Q(m)$ interpreted as choice probabilities $\Pr(x' = 1)$.

Our analysis thus implies that, for any $\varepsilon' > 0$, any supermajority rule admits equilibrium sequences (for some information structures) in which policy $x' = 1$ is chosen in all states with probability of at least $1 - \varepsilon'$ as $N \rightarrow \infty$ (Propositions 1 and 2). Hence, even a minimal likelihood of discretion can enlarge the equilibrium set in a way that upsets full-information equivalence. This observation complements work that identifies coordination failures and inefficiencies in related information-aggregation environments (see, e.g., Ali *et al.*, 2025; Barelli, Bhattacharya and Siga, 2022; Ekmekci and Lauermann, 2020; Mandler, 2012).

The literature has used models without commitment (pure cheap talk) to study informal political processes such as protests and contrasted their properties with formal voting rules. Impossibility results therein show that only babbling equilibria exist when a large group of agents strategically communicates with a principal whose prior is sufficiently separated from the agents' priors (see, e.g., Battaglini, 2017; Chen, 2025; Levit and Malenko, 2011; Morgan and Stocken, 2008). In contrast, in our setting even procedures with almost no commitment robustly aggregate information in all equilibrium sequences as $N \rightarrow \infty$ (Proposition 3). Thus, if commitment is interpreted as a measure of how formal a political process is, our results imply that formality matters less for information aggregation than the impossibility results suggest.

To understand the logic behind the cheap-talk impossibility results, consider the following example in our notation. Take a cheap-talk mechanism (for example, $Q(m) = \mathcal{P}$ for all m , so collective input never constrains the principal). Suppose the principal's prior odds are $\frac{\Pr(\omega=1)}{\Pr(\omega=0)} = 9$, each agent has a lower prior with odds $\frac{1}{2}$, and a favorable signal $s_i = 1$ has likelihood ratio 3, while an unfavorable signal $s_i = 0$ has likelihood ratio $\frac{1}{3}$. Suppose moreover that each player prefers high policies whenever her posterior likelihood ratio exceeds 1.

Write m_{-i} for the other agents' reports, and let \bar{k} denote the principal's cutoff, so that she chooses a high policy if and only if the total number of high reports is at least $\bar{k} + 1$. At any profile m_{-i} at which agent i is pivotal, reporting 1 must push the principal weakly above the cutoff while reporting 0 must keep her below it. Hence the principal's posterior odds conditional on m_{-i} and a report 1 must be at least 1, whereas her posterior odds conditional on m_{-i} and a report 0 must be below 1. This means

$$9 \cdot \frac{\Pr(m_{-i} \mid \omega = 1)}{\Pr(m_{-i} \mid \omega = 0)} \cdot 3 \geq 1 > 9 \cdot \frac{\Pr(m_{-i} \mid \omega = 1)}{\Pr(m_{-i} \mid \omega = 0)} \cdot \frac{1}{3}.$$

Equivalently,

$$\frac{1}{27} \leq \frac{\Pr(m_{-i} \mid \omega = 1)}{\Pr(m_{-i} \mid \omega = 0)} < \frac{1}{3}.$$

Now consider an agent who received a signal $s_i = 1$. Since her prior odds are $1/2$, her posterior odds conditional on the pivotal profile and her signal satisfy

$$\frac{\Pr_i(\omega = 1 \mid m_{-i}, s_i = 1)}{\Pr_i(\omega = 0 \mid m_{-i}, s_i = 1)} = \frac{\Pr_i(\omega = 1)}{\Pr_i(\omega = 0)} \cdot \frac{\Pr(m_{-i} \mid \omega = 1)}{\Pr(m_{-i} \mid \omega = 0)} \cdot \frac{\Pr(s_i = 1 \mid \omega = 1)}{\Pr(s_i = 1 \mid \omega = 0)} < \frac{1}{2} \times \frac{1}{3} \times 3 = \frac{1}{2}.$$

So even an agent with a signal 1 strictly prefers the report 0 at any pivotal profile. All non-partisan agents therefore babble, contradicting informativeness. This is the core pivotal-inference logic underlying the impossibility results.

Under partial commitment, by contrast, agents are pivotal in multiple pivotal events, not just at the event described above where the principal's preference changes. This upsets the existing pivotal-inference logic and allows informative equilibria to survive even when the usual impossibility conditions on the players' priors hold and even when commitment is just minimal.²⁸

7 Further Related Literature

We relate to a broader literature on voting models with multiple pivotal events. This paper is not the first to study environments with multiple pivotal events. In particular, Razin (2003) studies an environment in which multiple pivotal events arise since voters have signaling incentives, i.e., their actions may signal information to second-stage decision-makers—as in our setting. In his setting, information aggregation fails when decision makers and agents have disjoint sets of ex-post optimal policies. Our results differ in that we do not focus on such extreme ex-post preference conflicts; instead, our inefficiencies stem from coordination failures that can prevent full-information equivalence *despite* information aggregation, and our analysis characterizes robust approaches to these.

In other voting contexts, the literature has analyzed different types of signaling incentives²⁹ and multiple pivotal events that arise for different reasons than signaling, e.g. due to simultaneous voting on multiple issues (Ahn and Oliveros, 2012) or multiple voting

²⁸Strictly speaking, our full-support assumption on the agents' type distribution violates the exact conditions used in the mentioned impossibility results. However, even in a common-type version of our baseline model, the presence of multiple pivotal events implies the existence of informative equilibria for a broad class of quasi-referenda.

²⁹These contributions analyze settings in which early-stage collective behavior serves a signaling role: In multiple-round voting, a first-round vote can convey information to second-round voters (Piketty, 2000) or shape candidates' second-round proposals (Castanheira, 2003; Meirowitz and Shotts, 2009). Likewise, pre-election polls (Coughlan, 2000; Fey, 1997) or costly protest activity (Lohmann, 1994) may transmit information that strategically affects subsequent political choices.

thresholds (Damiano *et al.*, 2025).

Technically, we develop new methods to deal with the complexity created by multiple pivotal events. Most existing work relies primarily on statistical tools such as large-deviation theory. We combine large-deviation arguments with novel polynomial-algebraic methods, in particular leveraging Descartes’ rule of signs from polynomial algebra. Specifically, we relate the question of information aggregation to bounding the number of roots of an associated polynomial; cf. the proof of Lemma 1, Proposition 3, and Theorem 3 in the appendix.

We conclude the literature discussion with two remarks. The first remark concerns sincere voting. A longstanding question in political science is the extent to which citizens vote “sincerely” (see, e.g., Farquharson, 1969 and Palfrey, 2009). Previous theoretical work has shown that in elections with full commitment power and large electorates, sincere voting typically fails to be an equilibrium when voters are privately informed (see, e.g., Austen-Smith and Banks, 1996). In contrast, Proposition 1 and its proof show that, under partial commitment, strategy profiles in which approximately all agents act sincerely can constitute an equilibrium for a nontrivial range of information structures. Roughly speaking, the presence of multiple pivotal events weakens the incentives to deviate from sincere behavior that arise under full commitment. Related observations have been made in settings with participation costs (Krishna and Morgan, 2012) or aggregate uncertainty about the fraction of uninformed voters (Acharya and Meiwitz, 2017).

Finally, we note a conceptual connection to the delegation literature (Alonso and Matouschek, 2008; Holmström, 1984). A quasi-referendum can be viewed as a form of *collective delegation* to a principal: The collective action determines the feasible set the principal faces *ex post*. This perspective connects delegation-style reasoning to settings in which information is dispersed among many agents and to collective-choice applications such as shareholder voting and public finance.

8 Concluding Discussion

This paper studies commitment in referenda and related political institutions. When policy-relevant information is dispersed across many agents, the relevant institutional question is not simply whether collective input should matter, but how arrangements for incorporating that input shape policy outcomes in a strategic environment. Our analysis shows that, with a large group of agents and under realistic frictions, there is a case for minimal commitment. Minimal commitment gives strong enough incentives to sustain robust information aggregation; at the same time, it limits the extent to which coordination failures within the group distort outcomes. Stronger commitments make policy more exposed to such failures and therefore magnify the distortions they generate.

The paper identifies practically important referenda that implement this logic. Col-

lective vetoes and gateway referenda allow the collective to block or authorize one polar option while leaving the detailed policy specification to a principal. These referenda are simple and, when the policy space is fine, deliver outcomes arbitrarily close to the full-information benchmark. They also maximize the principal’s worst-case payoff guarantee across information structures and equilibria, within a broad class of decision processes.

The paper isolates one rationale for limited commitment, but it is unlikely to be the only one relevant in practice. Partial commitment may also serve purposes of legitimacy, accountability, or political cover. It may also preserve discretion for a principal with expertise or access to information that becomes relevant only after collective input has been observed (cf. Kartik, Van Weelden and Wolton (2017)). Our analysis abstracts from these considerations in order to isolate a single rationale: the trade-off between information aggregation and coordination failures. The observation is that this alone can rationalize minimal commitment.

Our model of (quasi-)referenda is a first step toward a broader theory of commitment in policy making under dispersed information. Future work could incorporate additional structure present in particular applications. For example, in shareholder voting, what factors determine how policy is decided after collective input is observed? Answering this may require introducing legislative bargaining, competing principals, or other strategic specification problems. Such extensions would further clarify how policy outcomes depend on the institutional environment in which final policy is chosen—that is, on the micro-foundations of partial commitment.

More broadly, the concept of partial commitment developed here may prove useful well beyond the setting studied in this paper, and may open up a broader research agenda on commitment in collective choice institutions.

Appendix

Throughout the appendices (except Appendix J), the following disclaimers apply.

1. For the principal, we only consider mixed strategies that are best responses. Given the best response characterization (1), we can thus identify a principal’s mixed strategy with a pair (\bar{k}, \tilde{x}) where $\bar{k} + 1$ is the principal’s cutoff at which she mixes and chooses a policy x randomly according to the distribution \tilde{x} .
2. We identify converging subsequences with the original sequence and omit the subsequence notation. Converging subsequences exist in each instance since the sequences lie in compact sets.

A Mathematical Preliminaries

We collect the tools that will be used repeatedly in the proofs below. In particular, we will approximate probabilities of pivotal events as $N \rightarrow \infty$.

A.1 Basics of Large Deviation Theory

Take a binomial distribution X_n with success probability $q \in (0, 1)$ and sample size n . Given any $m \in (0, 1)$ with $mn \in \mathbb{N}$, the probability of exactly mn successes out of n trials is well known to be³⁰

$$\Pr(X_n = mn) = \exp\left(-n\text{KL}(m, q) + o(n)\right), \quad (17)$$

where KL denotes the Kullback–Leibler divergence,

$$\text{KL}(m, q) = m \log\left(\frac{m}{q}\right) + (1 - m) \log\left(\frac{1 - m}{1 - q}\right).$$

A convenient derivation, due to Cramér (1938), uses a change of measure (Esscher transform, Escher (1932)). Consider the binomial distribution under which the event is not rare but rather typical, $Z_n \sim \mathcal{B}(n, m)$. Then (17) follows from observing that

$$\frac{\Pr(X_n = mn)}{\Pr(Z_n = mn)} = \exp\left(-n\text{KL}(m, q)\right) \text{ and } \Pr(Z_n = mn) = \exp\left(o(n)\right). \quad (18)$$

For the equation on the left, note that

$$\begin{aligned} \frac{\Pr(Z_n = mn)}{\Pr(X_n = mn)} &= \left(\frac{m}{q}\right)^{nm} \left(\frac{1 - m}{1 - q}\right)^{n(1 - m)} \\ &= \exp\left(\log\left(\left(\frac{m}{q}\right)^{nm} \left(\frac{1 - m}{1 - q}\right)^{n(1 - m)}\right)\right) \\ &= \exp\left(n\left(m \log\left(\frac{m}{q}\right) + (1 - m) \log\left(\frac{1 - m}{1 - q}\right)\right)\right). \end{aligned}$$

The equation on the right holds because the probability mass function (PMF) of the binomial distribution peaks at its mean, implying $\Pr(Z_n = mn) \in [\frac{1}{n}, 1]$. But for any sequence $(x_n)_{n \in \mathbb{N}}$ with $x_n \in [\frac{1}{n}, 1]$, it holds that $x_n = \exp\left(\log(x_n)\right) = \exp\left(o(n)\right)$.

A.2 Taylor Approximations of the Kullback–Leibler Divergence

Below we give two approximations of the Kullback–Leibler divergence

$$\text{KL}(m, q) = m \log\left(\frac{m}{q}\right) + (1 - m) \log\left(\frac{1 - m}{1 - q}\right).$$

³⁰Recall that a function f is $o(n)$ if $\frac{|f(n)|}{n}$ converges to 0 as $n \rightarrow \infty$.

The first is for $m \approx q$.³¹ For $m = q + \varepsilon'$ with small ε' , we expand the log terms using the Taylor expansion $\log(1+x) \approx x - \frac{x^2}{2}$ around $x = 0$ to obtain

$$\begin{aligned}\log \frac{m}{q} &= \log \left(1 + \frac{\varepsilon'}{q} \right) \approx \frac{\varepsilon'}{q} - \frac{(\varepsilon')^2}{2q^2}, \\ \log \frac{1-m}{1-q} &= \log \left(1 - \frac{\varepsilon'}{1-q} \right) \approx -\frac{\varepsilon'}{1-q} - \frac{(\varepsilon')^2}{2(1-q)^2},\end{aligned}$$

and substitute:

$$\begin{aligned}\text{KL}(m, q) &\approx (q + \varepsilon') \left(\frac{\varepsilon'}{q} - \frac{(\varepsilon')^2}{2q^2} \right) \\ &\quad + (1 - q - \varepsilon') \left(-\frac{\varepsilon'}{1-q} - \frac{(\varepsilon')^2}{2(1-q)^2} \right).\end{aligned}$$

Expanding the products and discarding all third-order terms, we have

$$\begin{aligned}(q + \varepsilon') \left(\frac{\varepsilon'}{q} - \frac{(\varepsilon')^2}{2q^2} \right) &\approx \varepsilon' - \frac{(\varepsilon')^2}{2q} + \frac{(\varepsilon')^2}{q}, \text{ and} \\ (1 - q - \varepsilon') \left(-\frac{\varepsilon'}{1-q} - \frac{(\varepsilon')^2}{2(1-q)^2} \right) &\approx -\varepsilon' - \frac{(\varepsilon')^2}{2(1-q)} + \frac{(\varepsilon')^2}{1-q}.\end{aligned}$$

Noticing that ε' cancels out, and simplifying the coefficients of $(\varepsilon')^2$, we have

$$\begin{aligned}\text{KL}(m, q) &\approx \left(-\frac{(\varepsilon')^2}{2q} + \frac{(\varepsilon')^2}{q} \right) + \left(-\frac{(\varepsilon')^2}{2(1-q)} + \frac{(\varepsilon')^2}{1-q} \right) \\ &= \frac{(\varepsilon')^2}{2q} + \frac{(\varepsilon')^2}{2(1-q)} \\ &= \frac{(\varepsilon')^2}{2} \left(\frac{1}{q} + \frac{1}{1-q} \right).\end{aligned}$$

We thus obtain the quadratic approximation

$$\text{KL}(m, q) \approx \frac{(m - q)^2}{2q(1 - q)} \text{ for } m \approx q. \quad (19)$$

For the second approximation, consider $q_1 \approx q_2$. Note that $\frac{\partial}{\partial q} \text{KL}(m, q) = -\frac{m}{q} + \frac{1-m}{1-q}$. We use linear Taylor approximations of $\text{KL}(m, q_1)$ and $\text{KL}(m, q_2)$ around the midpoint $\bar{q} = \frac{q_1 + q_2}{2}$,

$$\begin{aligned}\text{KL}(m, q_1) &\approx \text{KL}(m, \bar{q}) + \frac{\partial}{\partial q} \text{KL}(m, q)|_{q=\bar{q}}(q_1 - \bar{q}), \\ \text{KL}(m, q_2) &\approx \text{KL}(m, \bar{q}) + \frac{\partial}{\partial q} \text{KL}(m, q)|_{q=\bar{q}}(q_2 - \bar{q}),\end{aligned}$$

³¹For two sequences $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$, we write $a_n \approx b_n$ if $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$. Note that we do not retain the subscript.

to approximate the difference of these two quantities:

$$\begin{aligned} \text{KL}(m, q_1) - \text{KL}(m, q_2) &\approx \frac{\partial}{\partial q} \text{KL}(m, q)|_{q=\bar{q}}(q_1 - q_2) \\ &= \left(\frac{1-m}{1-\bar{q}} - \frac{m}{\bar{q}} \right) (q_1 - q_2). \end{aligned} \quad (20)$$

A.3 Properties of Binomial Distributions

In this section we record a property of the probability mass functions of binomial distributions. The binomial distribution $X_n \sim B(n, q)$ has an inverse-U-shaped PMF, with a unique mode at $\lfloor (n+1)q \rfloor$ if $(n+1)q$ is not an integer, and otherwise with two adjacent modes given by $(n+1)q$ and $(n+1)q - 1$; see, e.g., page 112 in Chapter 3.4 of Johnson, Kemp and Kotz (2005).

Claim 1. *Take a binomial distribution $X_n \sim \mathcal{B}(n, q)$. If $(n+1)q \in \mathbb{N}$, then*

$$\begin{aligned} \Pr(X_n = k) &< \Pr(X_n = k') \text{ for all } k < k' \leq (n+1)q - 1, \\ \Pr(X_n = k) &= \Pr(X_n = k') \text{ for } k = (n+1)q - 1 \text{ and } k' = (n+1)q, \\ \Pr(X_n = k) &> \Pr(X_n = k') \text{ for all } k' > k \geq (n+1)q. \end{aligned}$$

If $(n+1)q \notin \mathbb{N}$, then

$$\begin{aligned} \Pr(X_n = k) &< \Pr(X_n = k') \text{ for all } k < k' \leq \lfloor (n+1)q \rfloor, \\ \Pr(X_n = k) &> \Pr(X_n = k') \text{ for all } k' > k \geq \lfloor (n+1)q \rfloor. \end{aligned}$$

A.4 Pivotal Probabilities

The relevance of the mathematical preliminaries in Sections A.1-A.3 for our collective choice model derives from the fact that, for any symmetric strategy σ of the agents, the number of actions 1 taken by $N - 1$ agents follows a binomial distribution with success probability $q(\omega'; \sigma) = (1 - \rho_1 - \rho_0)E(\sigma(s)|\omega = \omega') + \rho_1$. The pivotal events in our model thus correspond to point events of a binomial distribution, and (17) provides a suitable approximation of their likelihood. To be precise, if we let $q = q(\omega', \sigma)$ and $m = \frac{\lfloor m_j N \rfloor}{N}$ for $j > 0$, then (17) becomes

$$\Pr(\text{piv}_j | \omega = \omega'; \sigma, N) = \exp \left(- (N-1) \text{KL} \left(\frac{\lfloor m_j N \rfloor}{N}, q(\omega'; \sigma) \right) + o(N) \right). \quad (21)$$

Similarly, if we let $q = q(\omega', \sigma)$ and $m = \frac{k}{N}$ for $k \in \{\bar{k}, \bar{k} + 1\}$, then (17) becomes

$$\Pr(\text{piv}^k | \omega = \omega'; \sigma, N) = \exp \left(- (N-1) \text{KL} \left(\frac{k}{N}, q(\omega'; \sigma) \right) + o(N) \right). \quad (22)$$

B Proof of Proposition 1

Our proof strategy is as follows. Take any quasi-referendum with cutoffs m_1, \dots, m_{R+1} . Fix a target $j^* \in \{1, \dots, R+1\}$.

1. First, we present candidate equilibrium strategies for the principal and the agents (Section B.1).
2. Second, we show an auxiliary result about the ordering of the principal's best response cutoff $\frac{\tilde{k}}{N}$ and the agents' mean actions in the two states (Section B.2).
3. Third, we choose an information structure such that, for every candidate strategy, (a) for any principal's best response to it, the agents' incentives approximate those in the cheap-talk game as $N \rightarrow \infty$ and (b) the realized feasible policy set is Q_{j^*} with probability approaching one (Section B.3).
4. Fourth, we construct the equilibrium sequences with a fixed-point argument (Section B.4).

B.1 Candidate Equilibrium Strategies

For any $x \in (0, 1)$, denote by q_x the x -quantile of the prior distribution, $q_x = F^{-1}(x)$. We consider the following candidate strategies. The principal follows a mixed strategy randomizing between the pure strategy where she chooses

$$\begin{aligned} & \min Q\left(\frac{k}{N}\right) \text{ if } k \leq k^*, \\ & \max Q\left(\frac{k}{N}\right) \text{ if } k > k^*, \end{aligned}$$

and the pure strategy where she chooses

$$\begin{aligned} & \min Q\left(\frac{k}{N}\right) \text{ if } k \leq k^* + 1, \\ & \max Q\left(\frac{k}{N}\right) \text{ if } k > k^* + 1, \end{aligned}$$

for some $k^* \in \{1, \dots, N-1\}$. We identify these two pure strategies with the cutoffs k^* and $k^* + 1$, and a mixture of the two with a random cutoff \tilde{k} or its associated probability $z = \Pr(\tilde{k} = k^*)$.

The non-partisan agents follow strategies $\sigma_{\mathbf{p}}$ under which, after observing signal $s_i \in \{0, 1\}$, agent i chooses $a_i = 1$ if and only if $p_i \geq p(s_i)$, for some $p(s_i) \in (0, 1)$. We slightly abuse notation and write $p_{s_i} = p(s_i)$ for readability. We identify $\sigma_{\mathbf{p}}$ with $\mathbf{p} := (p_0, p_1)$ and consider the set

$$D(\delta) = \{\mathbf{p} : q_{\frac{\delta}{4}} \leq p_1 \leq q_{\frac{\delta}{2}}, \Pr(\omega = 1) \leq p_0 < 1\}.$$

For any $\mathbf{p} \in D(\delta)$, the strategy $\sigma_{\mathbf{p}}$ is δ -approximately truthful under any information structure satisfying (5) and $\rho_a < \frac{\delta}{4}$ for $a \in \{0, 1\}$ (constructed in Section B.3 below). This is because after signal 0, the likelihood of action 1 is below

$$\rho_1 + (1 - \rho_1 - \rho_0) \left(1 - F(\Pr(\omega = 1))\right) \leq \rho_1 + (1 - \rho_1 - \rho_0) \frac{\delta}{4} < \delta;$$

where we used that $\Pr(\omega = 1) > q_{1-\frac{\delta}{4}}$, by (5), and for the second inequality we used the

Binomial formula. After signal 1, the likelihood of action 1 exceeds

$$\rho_1 + (1 - \rho_1 - \rho_0) \left(1 - F(q_{\frac{\delta}{2}})\right) \geq \left(1 - \frac{\delta}{4}\right)^2 \geq 1 - \delta,$$

where for the first inequality we used that $\rho_a < \frac{\delta}{4}$ for $a \in \{0, 1\}$.

We denote by

$$\hat{\mathbf{p}}(\mathbf{p}, z) = \left(\hat{p}(0; \mathbf{p}, z), \hat{p}(1; \mathbf{p}, z)\right) \in [0, 1]^2$$

the cutoffs of the best response given $\mathbf{p} \in D(\delta)$ and z . The best-response cutoffs are well-defined since the properties of $\mathbf{p} \in D(\delta)$ and z imply $U(\omega') > 0$ for $\omega' \in \{0, 1\}$. They are given by (9). Often we drop the arguments (\mathbf{p}, z) from the notation.

B.2 Ordering of the Principal's Cutoff and the Mean Actions

The principal's best response cutoff $\frac{\bar{k}}{N}$ to any strategy $\sigma_{\mathbf{p}}$ with $\mathbf{p} \in D(\delta)$ lies in between the strategy's mean actions when N is large and δ is small.

Lemma 3 (Interior Principal's Cutoff). *There is some $N_1 \in \mathbb{N}$ and some $\gamma_1 > 0$ such that*

$$\frac{\bar{k}}{N} \in \left(q(0; \mathbf{p}) + \gamma_1, q(1; \mathbf{p}) - \gamma_1\right) \quad (23)$$

for all $N \geq N_1$, $\mathbf{p} \in D(\delta)$, and $\delta \leq \frac{1}{4}$.

Proof. The restriction $\delta \leq \frac{1}{4}$ implies $q_{\frac{\delta}{2}} < q_{1-\frac{\delta}{2}}$; thus, there is a uniform lower bound $\eta > 0$ such that

$$q(1; \mathbf{p}) - q(0; \mathbf{p}) \geq \eta \text{ for all } \mathbf{p} \in D(\delta) \text{ with } \delta \leq \frac{1}{4}, \quad (24)$$

given the full support and absolute continuity of the distribution of the agents' priors and the different likelihood ratios of the two signals in the two states.

We start by applying the first equation of (18) to obtain

$$\frac{\Pr(k|\omega = 1; \mathbf{p})}{\Pr(k|\omega = 0; \mathbf{p})} = \exp\left(-N\left(\text{KL}\left(\frac{k}{N}, q(1; \mathbf{p})\right) - \text{KL}\left(\frac{k}{N}, q(0; \mathbf{p})\right)\right)\right) \quad (25)$$

for any $k \in \{0, \dots, N\}$. Consider

$$A_1 := \inf_{\mathbf{p} \in D(\delta), \delta \leq \frac{1}{4}} \left(\text{KL}\left(q(0; \mathbf{p}) + \gamma_1, q(1; \mathbf{p})\right) - \text{KL}\left(q(0; \mathbf{p}) + \gamma_1, q(0; \mathbf{p})\right)\right),$$

$$A_2 := \inf_{\mathbf{p} \in D(\delta), \delta \leq \frac{1}{4}} -\left(\text{KL}\left(q(1; \mathbf{p}) - \gamma_1, q(1; \mathbf{p})\right) - \text{KL}\left(q(1; \mathbf{p}) - \gamma_1, q(0; \mathbf{p})\right)\right),$$

which are strictly positive for $\gamma_1 > 0$ sufficiently small by (24) and since the Kullback–Leibler divergence $\text{KL}(m, q)$ has uniformly bounded partial derivatives on the parameter set $\{(m, q) : m \in [q(\omega'; \mathbf{p}) - \gamma_1, q(\omega'; \mathbf{p}) + \gamma_1] \text{ for some } \omega' \in \{0, 1\}, q \in \{q(0; \mathbf{p}), q(1; \mathbf{p})\}\}$.

Since $q(1; \mathbf{p}) > q(0; \mathbf{p})$, the function $\exp\left(-N\left(\text{KL}\left(\frac{k}{N}, q(1; \mathbf{p})\right) - \text{KL}\left(\frac{k}{N}, q(0; \mathbf{p})\right)\right)\right)$ is strictly increasing in k . This monotonicity implies

$$\frac{\Pr(k|\omega = 1; \mathbf{p}, N)}{\Pr(k|\omega = 0; \mathbf{p}, N)} < \exp(-NA_1) \text{ for any } k \text{ with } \frac{k}{N} \leq q(0; \mathbf{p}) + \gamma_1, \text{ and}$$

$$\frac{\Pr(k|\omega = 1; \mathbf{p}, N)}{\Pr(k|\omega = 0; \mathbf{p}, N)} > \exp(NA_2) \text{ for any } k \text{ with } \frac{k}{N} \geq q(1; \mathbf{p}) - \gamma_1.$$

Since A_1 and A_2 are strictly positive, for any $\kappa > 0$, there is some $N(\kappa) \in \mathbb{N}$ such that for all $N \geq N(\kappa)$,

$$\frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \exp(-NA_1) < \kappa \text{ and}$$

$$\frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \exp(NA_2) > \frac{1}{\kappa};$$

and the same bounds apply to the posterior likelihood ratio,

$$\frac{\Pr(\omega = 1|k; \mathbf{p}, N)}{\Pr(\omega = 0|k; \mathbf{p}, N)} < \kappa, \text{ for any } \frac{k}{N} \leq q(0; \mathbf{p}) + \gamma_1, \quad (26)$$

$$\frac{\Pr(\omega = 1|k; \mathbf{p}, N)}{\Pr(\omega = 0|k; \mathbf{p}, N)} > \frac{1}{\kappa} \text{ for any } \frac{k}{N} \geq q(1; \mathbf{p}) - \gamma_1. \quad (27)$$

Finally, we argue that we can choose $\kappa > 0$ small enough so that (23) holds uniformly, that is, for all $N \geq N_1 := N(\kappa)$, $\mathbf{p} \in D(\delta)$, and $\delta \leq \frac{1}{4}$. Suppose, on the contrary, that either $\frac{\bar{k}}{N} \geq q(1; \mathbf{p}) - \gamma_1$ or $\frac{\bar{k}}{N} \leq q(0; \mathbf{p}) + \gamma_1$. First, (26) and (27) imply that the posterior likelihood ratio crosses 1, so \bar{k} is not degenerate, $\bar{k} \neq N$. Second, since each private signal realization is boundedly informative, we can choose $\kappa > 0$ small enough so that for $N \geq N(\kappa)$, (26) and (27) imply

$$\frac{\Pr(\omega = 1|\bar{k} + 1; \mathbf{p}, N)}{\Pr(\omega = 0|\bar{k} + 1; \mathbf{p}, N)} < 1 \text{ if } \frac{\bar{k}}{N} \leq q(0; \mathbf{p}) + \gamma_1,$$

$$\frac{\Pr(\omega = 1|\bar{k}; \mathbf{p}, N)}{\Pr(\omega = 0|\bar{k}; \mathbf{p}, N)} > 1 \text{ if } \frac{\bar{k}}{N} \geq q(1; \mathbf{p}) - \gamma_1.$$

However, this would contradict the minimality of \bar{k} ; compare to its definition (1). \square

B.3 The Information Structures

Lemma 4 (Approximate Cheap-Talk). *There exists $\delta_1 > 0$ small enough and an agent's information structure so that for any sequence of strategies $(\sigma_{\mathbf{p}_N})_{N \in \mathbb{N}}$ with $\mathbf{p}_N \in D(\delta_1)$ and any sequence of principal's best responses to $(\sigma_{\mathbf{p}_N})_{N \in \mathbb{N}}$,*

$$\lim_{N \rightarrow \infty} \Pr(\text{piv}_0 | \text{piv}; \eta_N, N) = 1. \quad (28)$$

Proof. First, we define the agents' information structure. Fix the previously chosen $j^* \in \{1, \dots, R + 1\}$ and consider an agents' information structure satisfying (4), (5) for

the bound \underline{p}_1 given by

$$\frac{\underline{p}_1}{1 - \underline{p}_1} = \frac{\Pr(\omega = 1) \Pr(s_i = 0|\omega = 1) \Pr(s_i = 1|\omega = 0)}{\Pr(\omega = 0) \Pr(s_i = 0|\omega = 0) \Pr(s_i = 1|\omega = 1)}.$$

Further, the partisan probabilities $\rho_0 > 0$ and $\rho_1 > 0$ are taken to satisfy

$$\rho_a < \frac{\delta}{4}.$$

These choices imply that any strategy $\sigma_{\mathbf{p}_N}$ with $\mathbf{p}_N \in D(\delta)$ is δ -approximately truthful, as discussed in the preceding section.

The parameter $\delta > 0$ is taken to be small enough (smaller than some $\delta_1 > 0$) and the signal probabilities to satisfy the ordering

$$m_{j^*-1} < \Pr(s_i = 1|\omega = 0) < \Pr(s_i = 1|\omega = 1) < m_{j^*}$$

and to be much closer to each other than to the cutoffs m_{j^*-1} and m_{j^*} , so that there is $\nu > 0$ and given any strategy $\sigma_{\mathbf{p}_N}$ with $\mathbf{p}_N \in D(\delta)$,³²

$$q(\omega'; \sigma_{\mathbf{p}_N}) \in I_{j^*} \text{ for all } \omega' \in \{0, 1\}, \quad (29)$$

and

$$\begin{aligned} & \nu + \text{KL}\left(q(\omega'; \sigma_{\mathbf{p}_N}), q(\omega''; \sigma_{\mathbf{p}_N})\right) \\ & < \min_{\omega \in \{\omega', \omega''\}} \left(\text{KL}(m_j, q(\omega; \sigma_{\mathbf{p}_N}))\right) \text{ for all } j > 0 \text{ and } \omega', \omega'' \in \{0, 1\}, \end{aligned} \quad (30)$$

We now establish the claim of the lemma. Combining (29)-(30) and (23), we see that for $\delta \leq \delta_1$ and $N \geq N_1$ as in Lemma 3, and any strategy $\sigma_{\mathbf{p}_N}$ with $\mathbf{p}_N \in D(\delta)$, it holds

$$\begin{aligned} & \nu + \text{KL}\left(\frac{\bar{k}}{N}, q(\omega''; \sigma_{\mathbf{p}_N})\right) \\ & < \min_{\omega \in \{\omega', \omega''\}} \left(\text{KL}(m_j, q(\omega; \sigma_{\mathbf{p}_N}))\right) \text{ for all } j > 0 \text{ and } \omega', \omega'' \in \{0, 1\}, \end{aligned} \quad (31)$$

holds. Given (31), (21) and (22) imply (28). \square

We close this section, observing that, given (29)-(30), an application of the law of large numbers implies

$$\lim_{N \rightarrow \infty} \Pr(m \in I_{j^*} | \eta_N, N) = 1. \quad (32)$$

Remark 1. *Unlike in the main text, in the appendix we work with the Kullback-Leibler distances instead of Euclidean distances, which is more handy for a precise analysis. In the main text, we mentioned that it is key that the Euclidean distance of $q(\omega'; \sigma_N)$ to m_j and m_{j+1} is large enough, as measured by the bound $M > 0$ as defined in (6). The relevance of a large enough $M > 0$ is that it implies (30), given (6) and $\delta > 0$ small*

³²In the case where $j^* = 1$, we use the convention $m_{j^*-1} = 0$, slightly abusing the notation since m_0 is also the principal's limit cutoff.

enough.³³

B.4 The Equilibrium Construction

We construct equilibria as follows. We show there are $\mathbf{p}^* = (p_0^*, p_1^*) \in D(\delta)$ and a principal's mixing probability $z \in [0, 1]$ so that

$$\hat{\mathbf{p}}(\mathbf{p}^*, z) = \mathbf{p}^*, \quad (33)$$

and

$$\frac{\Pr(\omega = 1) \Pr(k = \bar{k} + 1 | \omega = 1; \mathbf{p}^*, N)}{\Pr(\omega = 0) \Pr(k = \bar{k} + 1 | \omega = 0; \mathbf{p}^*, N)} = 1; \quad (34)$$

This means the agents' cutoff strategy is a best response to itself, given z , and the principal is indifferent after observing $\bar{k} + 1$ actions 1, where \bar{k} is the principal's cutoff defined in (1). The principal's mixed strategy z is thus also a best response, and together the players' strategies constitute an equilibrium.

The construction reduces equilibrium existence to a fixed point of a one-dimensional map T_N in the cutoff p_0 . It works in three steps.

1. The first step shows that there is a continuous map from $p_0 \in [\Pr(\omega = 1), 1)$ to $p_1^*(p_0)$ so that the principal is indifferent at $\bar{k} + 1$ given $\mathbf{p} = (p_0, p_1^*(p_0))$ (Lemma 5 in Section B.4.2).
2. The second step shows that there is a continuous function that maps each $p_0 \in [\Pr(\omega = 1), 1)$ to a number $z^*(p_0) \in [0, 1]$ such that the first part of the fixed-point equation (33) holds for $\mathbf{p} = (p_0, p_1^*(p_0))$ (Lemma 6 in Section B.4.3) i.e.

$$\hat{p}(1; \mathbf{p}, z^*) = p_1^*(p_0). \quad (35)$$

3. The third step considers the mapping from p_0 to

$$T_N(p_0) := \max \left(\Pr(\omega = 1), \hat{p}(0; \mathbf{p}, z^*(p_0)) \right),$$

and shows that, for any N large enough, there is $\gamma_5 > 0$ so that (a) it is a continuous self-map on the compact interval $[\Pr(\omega = 1), 1 - \gamma_5]$ and (b) it (only) has interior fixed points $p_{0,N}^* > \Pr(\omega = 1)$ (Lemma 7 in Section B.4.4).

In all steps, we will identify a uniform upper bound on δ and a lower bound on N that are required for the arguments to hold.

³³To see this, take $\delta \leq \gamma$; then $|q(1; \sigma_N) - q(0; \sigma_N)| \leq 3\gamma$ given any δ -approximately truthful strategy σ_N . Together with (6) this means that $|q(1; \sigma_N) - q(0; \sigma_N)| \leq \frac{3}{M} |q(\omega'; \sigma_N) - m_j|$ for $j > 0$ and $\omega' \in \{0, 1\}$. So, arbitrarily large M makes the Euclidean distance between the mean actions arbitrarily small relative to their distance to the quasi-referendum's cutoffs. One can show that, for $M > 0$ large enough, the Kullback-Leibler distance between the mean actions is at least smaller than their Kullback-Leibler distance to the cutoffs.

B.4.1 Proof of Proposition 1

We present the proof of Proposition 1 based on the construction of interior fixed points $p_{0,N}^*$ of T_N .

Consider a sequence of interior fixed points $p_{0,N}^*$ of T_N . By definition, for any N and any interior fixed point $p_{0,N}^*$ of T_N , the agents' strategy $\mathbf{p}_N^* = (p_{0,N}^*, p_{1,N}^*(p_{0,N}^*))$ and the principal's mixed strategy given by $z^*(p_{0,N}^*)$ constitute an equilibrium. Since $\mathbf{p}_N^* \in D(\delta)$, the corresponding sequence of agents' strategies is a sequence of δ -approximately truthful strategies. Since the limits of the strategies' mean actions ($N \rightarrow \infty$) lie strictly in the interval I_{j^*} by (29) and (30), an application of the law of large numbers implies $\lim_{N \rightarrow \infty} \Pr(m \in I_{j^*} | \eta_N, N) = 1$. This proves Proposition 1.

B.4.2 Lemma about the Principal's Indifference

The principal will be indifferent if she observes that $k = k^*(p_0) + 1$ out of N agents have chosen the action 1. For any fixed $p_0 \in [\Pr(\omega = 1), 1)$, we define $k^*(p_0) + 1$ as the minimal observed number of actions 1 such that the principal weakly prefers $x = 1$ given $p_1 = q_{\frac{\delta}{2}}$, i.e.,

$$\frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \frac{\Pr\left(k = k^*(p_0) + 1 | \omega = 1; p_1 = q_{\frac{\delta}{2}}, p_0, N\right)}{\Pr\left(k = k^*(p_0) + 1 | \omega = 0; p_1 = q_{\frac{\delta}{2}}, p_0, N\right)} \geq 1. \quad (36)$$

Note that $k^*(p_0)$ equals the best-response cutoff \bar{k} given (p_0, p_1) , so that (23) implies it is not degenerate, i.e. $k^*(p_0) \neq N$.

Lemma 5 shows that, for any candidate cutoff p_0 , we can find a cutoff $p_1^*(p_0)$ such that, given the agents' strategy $\mathbf{p} = (p_0, p_1^*(p_0))$, the principal becomes indifferent at $k^*(p_0) + 1$, the cutoff defined in the preceding section.

Lemma 5 (Indifference of the Principal). *There exists $\delta_2 > 0$ such that for all $\delta \leq \delta_2$ there exists $N_2(\delta) \in \mathbb{N}$ and for all $N_2(\delta) \leq N$, there is a continuous function that maps each $p_0 \in [\Pr(\omega = 1), 1)$ to a number $p_1^*(p_0) \in [q_{\frac{\delta}{4}}, q_{\frac{\delta}{2}}]$ satisfying*

$$\frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \frac{\Pr\left(k = k^*(p_0) + 1 | \omega = 1; p_1 = p_1^*(p_0), p_0, N\right)}{\Pr\left(k = k^*(p_0) + 1 | \omega = 0; p_1 = p_1^*(p_0), p_0, N\right)} = 1. \quad (37)$$

Proof. We show the uniform existence of a number $p_1^*(p_0) \in [q_{\frac{\delta}{4}}, q_{\frac{\delta}{2}}]$ that solves (37). For this we establish the existence of some $\delta_2 > 0$ such that for each $\delta \leq \delta_2$ there is some $N_2(\delta) \in \mathbb{N}$ and

$$\frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \frac{\Pr\left(k = k^*(p_0) + 1 | \omega = 1; p_1 = q_{\frac{\delta}{4}}, p_0, N\right)}{\Pr\left(k = k^*(p_0) + 1 | \omega = 0; p_1 = q_{\frac{\delta}{4}}, p_0, N\right)} < 1 \quad (38)$$

for all $N \geq N_2(\delta)$, $\delta \leq \delta_2$, and $p_0 \in [\Pr(\omega = 1), 1)$.

(note that we fix $p_1 = q_{\frac{\delta}{4}}$ here). Combining (36) and (38) and applying the intermediate value theorem then yields a $p_1^*(p_0) \in (q_{\frac{\delta}{4}}, q_{\frac{\delta}{2}}]$ such that the principal is indifferent given

$p_1^*(p_0)$ —i.e., (37) holds. Now, δ_2 and $N(\delta)$ for $\delta \leq \delta_2$ exist by the following argument. First, the minimality of $k^*(p_0) + 1$ implies a uniform bound $\gamma_2 > 0$ such that

$$\Pr(\omega = 1 | k = k^*(p_0) + 1; \mathbf{p}, N) - \frac{1}{2} \leq \gamma_2 \quad (39)$$

for all N , $\delta \leq \frac{1}{4}$, and $p_0 \in [\Pr(\omega = 1), 1)$. Second, note that the definition of $k^*(p_0)$ equals that of \bar{k} for any $\mathbf{p} \in D(\delta)$ with $p_1 = q_{\frac{\delta}{2}}$. So, Lemma 3 implies

$$\frac{k^*(p_0)}{N} \in \left(q(0; \mathbf{p}) + \gamma_1, q(1; \mathbf{p}) - \gamma_1 \right) \quad (40)$$

for all $\mathbf{p} \in D(\delta)$ with $p_1 = q_{\frac{\delta}{2}}$, any $\delta \leq \frac{1}{4}$, and any $N \geq N_1$; cf. (23). Third, note that $q(0; \mathbf{p})$ and $q(1; \mathbf{p})$ are both strictly decreasing in p_1 . Given this, (40), and the properties of the prior and signal distribution, there is $\delta_2 > 0$ sufficiently small and $\gamma_3 > 0$ so that

$$\frac{k^*(p_0) + 1}{N} \in \left(q(0; \mathbf{p}), q(1; \mathbf{p}) \right)$$

and

$$\frac{\partial}{\partial p_1} \text{KL} \left(\frac{k^*(p_0) + 1}{N}, q(0; \mathbf{p}) \right) - \text{KL} \left(\frac{k^*(p_0) + 1}{N}, q(1; \mathbf{p}) \right) \geq \gamma_3$$

for all $\mathbf{p} \in D(\delta)$, $\delta \leq \delta_2$, and $N \geq N_1$. Jointly, these observations and (25) imply that the posterior

$$\frac{\Pr \left(k = k^*(p_0) + 1 | \omega = 1; \mathbf{p}, N \right)}{\Pr \left(k = k^*(p_0) + 1 | \omega = 0; \mathbf{p}, N \right)}$$

is strictly increasing in p_1 . Further, its derivative is bounded from below by an arbitrarily large number if N is arbitrarily large. Recalling (39) and (36), we see that for any fixed $\delta \leq \delta_2$, there is $N_2(\delta) \in \mathbb{N}$ such that (38) holds uniformly. As argued, (38) implies the uniform existence of $p_1^*(p_0) \in [q_{\frac{\delta}{4}}, q_{\frac{\delta}{2}}]$ solving (37).

Finally, we note that $p_1^*(p_0)$ is unique and continuous. It is unique because the posterior is strictly increasing. The continuity of $p_1^*(p_0)$ in p_0 follows from an application of the implicit function theorem. \square

B.4.3 The Agents' First Fixed Point Equation

Lemma 6 shows that, for any candidate cutoff p_0 , we can find a principal's mixed strategy $z^*(p_0)$ such that

$$\hat{p}(1; \mathbf{p}, z^*(p_0)) = p_1^*(p_0). \quad (41)$$

Lemma 6 (Mixing Lemma). *There exists $\delta_3 > 0$ such that for all $\delta \leq \delta_3$ there is $N_3(\delta) \in \mathbb{N}$ and for all $N \geq N_3(\delta)$, there is a continuous function that maps each $p_0 \in [\Pr(\omega = 1), 1)$ to a number $z^*(p_0) \in [0, 1]$ such that (41) holds.*

Proof. We fix $p_1 = p_1^*(p_0)$ for the duration of the proof of Claim 6. The proof leverages

the principal's indifference between the pure strategies with cutoffs $k^*(p_0)$ and $k^*(p_0) + 1$ (we drop p_0 from its notation in the following). We will derive approximations of the indifference cutoff $\hat{p}(1; \mathbf{p}, z)$ for the agents' best response, given either of the two pure strategies. The key is to establish that for small δ and large N ,

$$\hat{p}(1; \mathbf{p}, z) < p_1^*(p_0) \text{ if } z = 0, \text{ and} \quad (42)$$

$$\hat{p}(1; \mathbf{p}, z) > p_1^*(p_0) \text{ if } z = 1. \quad (43)$$

Since $\hat{p}(1; \mathbf{p}, z)$ is continuous in the probability z —see (3) and (9)—an application of the intermediate value theorem then implies the existence of a principal's mixed strategy $z^* \in (0, 1)$ such that $\hat{p}(1, \mathbf{p}, z^*) = p_1^*(p_0)$.

In the following, we first establish that the inequalities (42) and (43) hold in the double-limit where we take $N \rightarrow \infty$ and then $\delta \rightarrow 0$, i.e., we establish that $\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \hat{p}(1; \mathbf{p}, z) < p_1^*(p_0)$ if $z = 0$ and $\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \hat{p}(1; \mathbf{p}, z) > p_1^*(p_0)$ if $z = 1$. After that, we argue the existence of the uniform bounds δ_2 and $N_2(\delta)$ for $\delta \leq \delta_2$.

We start with the pure strategy where $1 - z = \Pr(k = k^* + 1) = 1$. This means that piv_0 is the event where $k_{-i} = k^*(p_0) + 1$ and

$$\Pr(\text{piv}_0 | \omega = \omega'; \mathbf{p}, N) = \Pr(k_{-i} = k^* + 1 | \omega = \omega'; \mathbf{p}) \text{ for } \omega' \in \{0, 1\}, \quad (44)$$

where $\Pr(k_{-i} = k^* + 1 | \omega = \omega'; \mathbf{p})$ is the posterior conditional on $k^* + 1$ out of $N - 1$ agents choosing the action 1. The principal is indifferent if she observes that $k = k^* + 1$ out of N agents have chosen the action 1, i.e.

$$\frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} = \frac{\Pr(k = k^* + 1 | \omega = 0; \mathbf{p}, N)}{\Pr(k = k^* + 1 | \omega = 1; \mathbf{p}, N)}. \quad (45)$$

Since the strategies given by \mathbf{p} are δ -approximately truthful, as $\delta \rightarrow 0$ we have

$$1 - q(\omega'; \mathbf{p}) \rightarrow \Pr(s_i = 0 | \omega = \omega'), \quad (46)$$

so that

$$\begin{aligned} \frac{\Pr(k_{-i} = k^* + 1 | \omega = 0; \mathbf{p}, N)}{\Pr(k_{-i} = k^* + 1 | \omega = 1; \mathbf{p}, N)} &= \left(\frac{q(0; \mathbf{p})}{q(1; \mathbf{p})} \right)^{k^* + 1} \left(\frac{1 - q(0; \mathbf{p})}{1 - q(1; \mathbf{p})} \right)^{N - k^* - 2} \\ &\rightarrow \frac{\Pr(k = k^* + 1 | \omega = 0; \mathbf{p}, N)}{\Pr(k = k^* + 1 | \omega = 1; \mathbf{p}, N)} \cdot \frac{\Pr(s_i = 0 | \omega = 1)}{\Pr(s_i = 0 | \omega = 0)}. \end{aligned} \quad (47)$$

Recall the property (30). It implies (7), i.e., that in the limit as $N \rightarrow \infty$, an agent cares only about piv_0 . Hence, the condition (9) for the indifference cutoffs $p(s)$ of the best response imply

$$\lim_{N \rightarrow \infty} \frac{\hat{p}(s; \mathbf{p}, z)}{1 - \hat{p}(s; \mathbf{p}, z)} = \frac{\Pr(\text{piv}_0 | \omega = 0; \mathbf{p}, N)}{\Pr(\text{piv}_0 | \omega = 1; \mathbf{p}, N)} \cdot \frac{\Pr(s_i = s | \omega = 0)}{\Pr(s_i = s | \omega = 1)}. \quad (48)$$

If we combine (44)–(47) and take $\delta \rightarrow 0$, this indifference condition becomes

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \hat{p}(1; \mathbf{p}, z) = \underline{p}_1 \quad (49)$$

for $s = 1$, with \underline{p}_1 given by

$$\frac{\underline{p}_1}{1 - \underline{p}_1} = \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \frac{\Pr(s_i = 0 | \omega = 1)}{\Pr(s_i = 0 | \omega = 0)} \cdot \frac{\Pr(s_i = 1 | \omega = 0)}{\Pr(s_i = 1 | \omega = 1)}. \quad (50)$$

Next, consider the pure strategy where $z = \Pr(\tilde{k} = k^*) = 1$. For this strategy,

$$\Pr(\text{piv}_0 | \omega = \omega'; \mathbf{p}) = \Pr(k_{-i} = k^* | \omega = \omega'; \mathbf{p}) \text{ for } \omega \in \{0, 1\}. \quad (51)$$

As $\delta \rightarrow 0$, we have

$$\begin{aligned} \frac{\Pr(k_{-i} = k^* | \omega = 0; \mathbf{p}, N)}{\Pr(k_{-i} = k^* | \omega = 1; \mathbf{p}, N)} &= \left(\frac{q(0; \mathbf{p})}{q(1; \mathbf{p})} \right)^{k^*} \cdot \left(\frac{1 - q(0; \mathbf{p})}{1 - q(1; \mathbf{p})} \right)^{N - k^* - 1} \\ &\rightarrow \frac{\Pr(k = k^* + 1 | \omega = 0; \mathbf{p}, N)}{\Pr(k = k^* + 1 | \omega = 1; \mathbf{p}, N)} \cdot \frac{\Pr(s_i = 1 | \omega = 1)}{\Pr(s_i = 1 | \omega = 0)}. \end{aligned} \quad (52)$$

If we combine (45), (51), and (52), and take $\delta \rightarrow 0$, the indifference condition (9) becomes

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \hat{p}(1; \mathbf{p}, z) = \bar{p}_1 \quad (53)$$

for $s = 1$, with \bar{p}_1 given by

$$\frac{\bar{p}_1}{1 - \bar{p}_1} = \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot \frac{\Pr(s_i = 1 | \omega = 1)}{\Pr(s_i = 1 | \omega = 0)} \cdot \frac{\Pr(s_i = 1 | \omega = 0)}{\Pr(s_i = 1 | \omega = 1)} = \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)}. \quad (54)$$

Now, we combine the approximations (49) and (53) with the requirement (5) on the prior distribution to argue that (42) and (43) hold “in the double-limit.” The requirement (5) implies

$$p_1 < q_{\frac{\delta}{4}} \text{ and } q_{1 - \frac{\delta}{4}} < \Pr(\omega = 1). \quad (55)$$

Combining this with $\Pr(\omega = 1) \leq \bar{p}_1$ and $q_{\frac{\delta}{4}} < p_1^*(p_0) \leq q_{\frac{\delta}{2}}$ shows that

$$\underline{p}_1 < p_1^*(p_0) < \bar{p}_1. \quad (56)$$

Hence, (49) and (53) imply (42) and (43) in the limit, i.e., as $N \rightarrow \infty$ and then $\delta \rightarrow 0$. Note that we used Lemma 5 here, which guarantees the existence of $p_1^*(p_0)$ for $\delta \leq \delta_2$ and $N \geq N_2(\delta)$ (this also explains the order of limits).

Next we show that there are uniform bounds $\delta_3 > 0$ and $N_3(\delta) \in \mathbb{N}$ for $\delta \leq \delta_3$ such that (42) and (43) hold not just in the limit, but for any $\delta \leq \delta_3$, $N_3(\delta) \leq N$, and $\mathbf{p} \in D(\delta)$ with $p_1 = p_1^*(p_0)$. The limit analysis for $\hat{p}(1; \mathbf{p}, z)$ used two approximations, (46) and (48), and we argue that for any $\gamma_4 > 0$ we can find uniform bounds that ensure that both approximations hold up to an error term of at most γ_4 . For (46), this is obvious for some δ small

enough. For (48), this follows from observing that the convergence here is exponential in the difference $\min_{\omega' \in \{0,1\}} \text{KL}\left(\frac{\bar{k}}{N}, q(\omega'; \mathbf{p})\right) - \min_{\omega' \in \{0,1\}} \min_{j>0} \text{KL}\left(m_j, q(\omega'; \mathbf{p})\right)$, and this difference is uniformly bounded from below, given (30) and (23), for any $N \geq N_1$, and $\mathbf{p} \in D(\delta)$ for $\delta \leq \delta_1$ and (with N_1 and δ_1 as in Lemmas 3 and 4). Hence, for any $\gamma_4 > 0$, the likelihood ratio $\frac{\hat{p}(1; \mathbf{p}, z)}{1 - \hat{p}(1; \mathbf{p}, z)}$ is γ_4 -close to its limit when N is above a certain bound $N_3(\delta)$. This way, the inequalities (55) and (56) do not only hold for the limit terms, but for $N \geq N_3(\delta)$,

$$\hat{p}(1; \mathbf{p}, z) < \begin{matrix} q_{\frac{\delta}{4}} \text{ for } z = 0 \text{ and} \\ q_{1-\frac{\delta}{4}} \end{matrix} \quad (57)$$

$$q_{1-\frac{\delta}{4}} < \hat{p}(1; \mathbf{p}, z) \text{ for } z = 1. \quad (58)$$

Thus, (42) and (43) hold, given that $q_{\frac{\delta}{4}} < p_1^*(p_0) \leq q_{\frac{\delta}{2}}$, as stated above.

Finally, since (42) and (43) hold for any $\delta < \delta_3$, $N \geq N_3(\delta)$, and $\mathbf{p} \in D(\delta)$ with $p_1 = p_1^*(p_0)$, for any $p_0 \geq \Pr(\omega = 1)$, an application of the intermediate value theorem yields a $z^*(p_0)$ such that (41) holds. Since $\hat{p}(1; \mathbf{p}, z)$ is strictly monotone in z , the mixed strategy z^* is unique; since $\hat{p}(1; \mathbf{p}, z)$ is continuous in p_0 (this is because p_0 , p_1 , and z affect the likelihood of the pivotal events in a continuous way), z^* is also continuous in p_0 . \square

B.4.4 The Fixed-Point Construction

We use a fixed point argument to establish the existence of p_0^* such that the agents' strategy given by $(p_0^*, p_1^*(p_0))$ is a best response to itself and the principal's mixed strategy $z(p_0^*)$.

Fix any N and δ that satisfy the uniform bounds of Lemma 6. This ensures that the mapping from $p_0 \in [\Pr(\omega = 1), 1)$ to the projection

$$T_N(p_0) := \max\left(\Pr(\omega = 1), \hat{p}(0; \mathbf{p}, z^*(p_0))\right), \text{ with } \mathbf{p} = (p_0, p_1^*(p_0)),$$

is well-defined.

Lemma 7 (Fixed-Point Lemma). *For any N , there is $\gamma_5 > 0$ so that the function T_N is a continuous self-map on the compact interval $[\Pr(\omega = 1), 1 - \gamma_5]$ and has only interior fixed-points.*

Proof. The mapping is continuous in $p_0 \in [\Pr(\omega = 1), 1)$ because $\hat{p}(0; \mathbf{p}, z^*(p_0))$ is continuous in p_0 . (This is because $p_1^*(p_0)$ and $z^*(p_0)$ are continuous in p_0 and all three parameters p_0 , $p_1^*(p_0)$ and $z^*(p_0)$ affect the likelihood of the pivotal events in a continuous way.) For any fixed N , δ , and an agents' information structure, the best response $\hat{p}(0; \mathbf{p}, z^*)$ is uniformly bounded, i.e., $\hat{p}(0; \mathbf{p}, z^*) \leq 1 - \eta_5$ for some $\eta_5 > 0$ and all $\mathbf{p} \in D(\delta)$. So, the projection T_N is a continuous self-map on the compact interval $[\Pr(\omega = 1), 1 - \eta_5]$. An application of Brouwer's fixed point theorem then yields a fixed point $p_N^*(0)$.

We argue that any fixed point $p_N^*(0)$ is interior, i.e., it is strictly greater than $\Pr(\omega = 1)$. To show this, we use (50). The best-response cutoff $\hat{p}(0; \mathbf{p}, z^*)$ relates to $\hat{p}(1; \mathbf{p}, z^*) =$

$p_1^*(p_0)$ via the following equation:

$$\frac{p_1^*(p_0)}{1 - p_1^*(p_0)} = \frac{\hat{p}(0; \mathbf{p}, z^*)}{1 - \hat{p}(0; \mathbf{p}, z^*)} \cdot \frac{\Pr(s_i = 0 | \omega = 1)}{\Pr(s_i = 0 | \omega = 0)} \cdot \frac{\Pr(s_i = 1 | \omega = 0)}{\Pr(s_i = 1 | \omega = 1)}.$$

Comparing this to (50), we see that $\underline{p}_1 < p_1^*(p_0)$ (which holds since N and δ satisfy the uniform bounds of Claim 6; cf. (56) in the proof of this claim) implies $\Pr(\omega = 1) < \hat{p}(0; \mathbf{p}, z^*)$. But this means that the boundary point $p_0 = \Pr(\omega = 1)$ is not a fixed point. \square

C Proof of Proposition 2

For any large enough N , we construct an equilibrium strategy σ_N with the mean action exceeding the highest cutoff in each state,

$$m_R + \gamma < q(0; \sigma_N) < q(1; \sigma_N), \quad (59)$$

for some $\gamma > 0$. An application of the law of large numbers then yields the claim of Proposition 2

The equilibrium strategy is found among a parametric family of cutoff strategies. For each $L > 0$, let σ_L denote the strategy under which, after observing signal $s \in \{0, 1\}$, a non-partisan agent chooses action 1 if and only if

$$p_i \geq p_L(s),$$

where $p_L(s)$ is implicitly defined by

$$L := \frac{\Pr_i(\omega = 1 \mid p_i = p_L(s), s_i = s)}{\Pr_i(\omega = 0 \mid p_i = p_L(s), s_i = s)} = \frac{p_L(s)}{1 - p_L(s)} \frac{\Pr(s_i = s \mid \omega = 1)}{\Pr(s_i = s \mid \omega = 0)}. \quad (60)$$

Thus, higher values of L correspond to more demanding cutoff types. In particular, $p_L(s)$ is strictly increasing in L for each signal realization s , and since signal 1 is more favorable to state 1 than signal 0, we have

$$p_L(1) < p_L(0) \quad \text{for all } L > 0.$$

Hence the induced mean actions satisfy

$$q(0; \sigma_L) < q(1; \sigma_L).$$

We will work on a compact set of parameters $L \in [\underline{L}, \bar{L}]$, where

$$\underline{L} := \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot I^{-1} < \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} \cdot I =: \bar{L}. \quad (61)$$

for $I = \frac{\Pr(s_i=1|\omega=1) \Pr(s_i=0|\omega=0)}{\Pr(s_i=1|\omega=0) \Pr(s_i=0|\omega=1)}$. These bounds are chosen so that, for any principal best response (k, \tilde{x}) to a strategy σ_L , the likelihood ratio conditional on the pivotal event piv_0

lies in the interval $[\underline{L}, \bar{L}]$:

$$\frac{\Pr\left(\text{piv}_0 | \omega = 0; \sigma_L, (\bar{k}, \tilde{x}), N\right)}{\Pr\left(\text{piv}_0 | \omega = 1; \sigma_L, (\bar{k}, \tilde{x}), N\right)} \in [\underline{L}, \bar{L}]. \quad (62)$$

To verify (62), recall that piv_0 is the event that the realized number k_{-i} of other agents choosing action 1 is either \bar{k} or $\bar{k} + 1$. Since

$$\frac{\Pr\left(\omega = 1 | k = \bar{k}; \sigma_L, N\right)}{\Pr\left(\omega = 0 | k = \bar{k}; \sigma_L, N\right)} < 1 \leq \frac{\Pr\left(\omega = 1 | k = \bar{k} + 1; \sigma_L, N\right)}{\Pr\left(\omega = 0 | k = \bar{k} + 1; \sigma_L, N\right)},$$

the likelihood ratio conditional on piv_0 must indeed lie in $[\underline{L}, \bar{L}]$ (we will also make sure that \bar{k} is not degenerate, meaning that $\bar{k} \neq N$, so that the posterior likelihood ratios are well-defined.)

The next lemma packages the fixed-point argument that we will run on the compact parameter set.

Lemma 8 (Self-map lemma). *There exists $\bar{q} \in (m_R, 1 - \rho_0)$ such that, for any distribution of priors satisfying*

$$\rho_1 + (1 - \rho_0 - \rho_1) \left(1 - F(\bar{p})\right) > \bar{q},$$

there are $\bar{N} \in \mathbb{N}$ and $\bar{\delta} > 0$ with the following property: for every $N \geq \bar{N}$, every $L \in [\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}]$, and every principal best response (\bar{k}_N, \tilde{x}_N) to σ_L ,

(i) the induced mean actions satisfy

$$\bar{q} < q(0; \sigma_L) < q(1; \sigma_L),$$

(ii) the agents' best response to σ_L and (\bar{k}_N, \tilde{x}_N) is again a cutoff strategy $\sigma_{L'}$ with

$$L' \in [\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}].$$

Proof. Recall the definition of \bar{p} from the main text; in mathematical notation,

$$\frac{\Pr_i(\omega = 1 | p_i = \bar{p}, s_i = 0)}{\Pr_i(\omega = 0 | p_i = \bar{p}, s_i = 0)} = \frac{\Pr(\omega = 1)}{\Pr(\omega = 0)} I. \quad (63)$$

Fix an information structure satisfying

$$\rho_1 + (1 - \rho_0 - \rho_1) \left(1 - F(\bar{p})\right) > \bar{q}$$

for some $1 - \rho_0 > \bar{q} > m_R$ to be chosen below.

We begin with part (i). Since $q(0; \sigma_L) < q(1; \sigma_L)$ for all L , it remains to obtain a uniform lower bound on $q(0; \sigma_L)$. By continuity of $p_L(s)$ in L and the identity $p_{\bar{L}}(0) = \bar{p}$ (which follows by comparing (60) at $s = 0$ with (63) and the definition (61) of \bar{L}), there

exists $\delta > 0$ such that

$$\rho_1 + (1 - \rho_0 - \rho_1) \left(1 - F(p_{\bar{L}+\delta}(0))\right) > \bar{q}.$$

Now fix any $L \in [\underline{L} - \delta, \bar{L} + \delta]$. Since $p_L(s)$ is increasing in L ,

$$p_L(s) \leq p_{\bar{L}+\delta}(s) \quad \text{for all } s \in \{0, 1\}.$$

Hence, using $p_L(1) < p_L(0)$,

$$\begin{aligned} q(\omega; \sigma_L) &= \rho_1 + (1 - \rho_0 - \rho_1) \sum_{s=0,1} \Pr(s_i = s \mid \omega) \left(1 - F(p_L(s))\right) \\ &\geq \rho_1 + (1 - \rho_0 - \rho_1) \left(1 - F(p_L(0))\right) \\ &\geq \rho_1 + (1 - \rho_0 - \rho_1) \left(1 - F(p_{\bar{L}+\delta}(0))\right) > \bar{q} \end{aligned}$$

for both states $\omega \in \{0, 1\}$. This proves part (i).

Next, fix $L \in [\underline{L} - \delta, \bar{L} + \delta]$ and consider any sequence of principal's best responses (\bar{k}_N, \tilde{x}_N) to σ_L . We claim that

$$q(0; \sigma_L) < \lim_{N \rightarrow \infty} \frac{\bar{k}_N}{N} < q(1; \sigma_L). \quad (64)$$

(which implies $\bar{k}_N \neq N$). Indeed, if one of the inequalities failed, then

$$\lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(0; \sigma_L)\right) \neq \lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(1; \sigma_L)\right),$$

so by (17) the principal's posterior at \bar{k}_N would converge to either 0 or 1. But, by definition, the principal's posterior at \bar{k}_N lies below $\frac{1}{2}$ and her posterior at $\bar{k}_N + 1$ weakly above. Since a single agent's action is only boundedly informative, this directly contradicts the convergence to 0 or 1.

Choose $\bar{q} \in (m_R, 1 - \rho_0)$ large enough so that whenever $q(0; \sigma_L) \geq \bar{q}$, then

$$\text{KL}\left(m_R, q(\omega'; \sigma_L)\right) > \text{KL}\left(m_0, q(0; \sigma_L)\right) \quad \text{for all } \omega' \in \{0, 1\}, \quad (65)$$

where

$$m_0 := \lim_{N \rightarrow \infty} \frac{\bar{k}_N}{N}.$$

Such a choice is possible because when $q(0; \sigma_L)$ is close to $1 - \rho_0$, (64) implies m_0 is close to both mean actions, so the right-hand side of (65) is close to 0, whereas the left-hand side remains strictly positive since $m_R < 1 - \rho_0$.

Combining part (i) with (65), and using (21)- (22), we obtain the key large-deviation implication:

$$\Pr(\text{piv}_0 \mid \text{piv}; \sigma_L, N) \rightarrow 1. \quad (66)$$

That is, conditional on being pivotal, an agent almost certainly affects the principal's preference over policies instead of the feasible policy set.

We now prove part (ii). Fix any $L \in [L - \delta, \bar{L} + \delta]$ and any principal best response (\bar{k}_N, \tilde{x}_N) to σ_L . By (66) and the best-response characterization (9), the agents' best response is again of cutoff form, say $\sigma_{L'_N}$, with

$$L'_N \rightarrow \frac{\Pr(\text{piv}_0 | \omega = 0; \sigma_L, (\bar{k}_N, \tilde{x}_N), N)}{\Pr(\text{piv}_0 | \omega = 1; \sigma_L, (\bar{k}_N, \tilde{x}_N), N)}$$

as $N \rightarrow \infty$. By (62), the limit lies in $[L, \bar{L}]$. Therefore, we can find a bound \bar{N} , for which

$$L'_N \in [L - \delta, \bar{L} + \delta] \quad \text{for all } N \geq \bar{N}.$$

This proves part (ii). □

We can now complete the proof of Proposition 2. For every $N \geq \bar{N}$, let Γ_N denote the correspondence on the compact interval $[L - \delta, \bar{L} + \delta]$ that maps a parameter L to the set of parameters L' generated by the following two-step procedure:

- (a) choose a principal best response (\bar{k}_N, \tilde{x}_N) to σ_L ,
- (b) choose an agents' best response $\sigma_{L'}$ to σ_L and (\bar{k}_N, \tilde{x}_N) .

By Lemma 8, Γ_N maps the interval into itself. Moreover, it has non-empty, compact, convex values and a closed graph. Hence Kakutani's fixed point theorem yields a fixed point L_N^* of Γ_N for every $N \geq \bar{N}$.

The corresponding strategy $\sigma_{L_N^*}$ is an equilibrium strategy. By Lemma 8(i),

$$\bar{q} < q(0; \sigma_{L_N^*}) < q(1; \sigma_{L_N^*}).$$

Since $m_R < \bar{q}$, choose any $\gamma > 0$ such that $m_R + \gamma < \bar{q}$. Then

$$m_R + \gamma < q(0; \sigma_{L_N^*}) < q(1; \sigma_{L_N^*});$$

so the equilibrium sequence satisfies (59), as claimed.

D Proof of Lemma 1

Recall that Proposition 3's proof sketch in the main text established that any equilibrium is informative.

To prove Lemma 1, we show that the average effect cannot equal zero in both states,

$$U(0; \eta_N) \neq 0 \quad \text{or} \quad U(1; \eta_N) \neq 0. \tag{67}$$

Lemma 1 follows from (67) since $U(0; \eta_N)$ and $U(1; \eta_N)$ must have the same sign. If the two average effects had opposite signs, then the best-response inequality (2) would point *all* non-partisan types in the same direction regardless of their signal, producing an uninformative best response and contradicting the informativeness of equilibrium.

Before we establish (67), a preliminary observation: Any principal's equilibrium strategy is necessarily non-constant on average, i.e. $E(x(k))$ is non-constant in k . Suppose it were not. The principal is never indifferent except possibly when observing $k = \bar{k} + 1$,

hence chooses a boundary point of the feasible set otherwise. Constant $E(x(k))$ would mean chosen boundaries of the lower and upper feasible set $Q(0)$ and $Q(1)$ that are equal. Since $\min Q(0) < \max Q(0) < \max Q(1)$ and $\min Q(1) \neq \max Q(0)$, equality implies the choice is $E(x(k)) = \min Q(0) = \min Q(1)$. However, this implies a contradiction: By Bayes-consistency, the prior $\Pr(\omega = 1) > \frac{1}{2}$ implies k with $\Pr(\omega = 1|k) > \frac{1}{2}$ and thus $E(x(k)) \in \{\max Q(0), \max Q(1)\}$, a contradiction.

Now we establish (67): For each k , let $r_N(k) := E(x(k+1)) - E(x(k))$ denote the change in the principal's expected policy when the observed number of actions 1 increases from k to $k+1$. For any given agent, the number k_{-i} of action 1 of the other agents is Binomial, so we can write the average effect of an additional action 1 as a binomial-weighted sum:

$$U(\omega'; \eta_N) = \sum_{k_{-i}=k \in \{\lfloor m_1 N \rfloor, \bar{k}_N, \bar{k}_N + 1\}} \binom{N-1}{k} q_N(\omega'; \sigma_N)^k (1 - q_N(\omega'; \sigma_N))^{N-1-k} r_N(k).$$

Introduce the odds ratio $t := \frac{q_N(\omega')}{1 - q_N(\omega')} \in (0, \infty)$. Since $q^k(1-q)^{N-1-k} = (1-q)^{N-1} t^k$, the condition $U(\omega'; \eta_N) = 0$ is equivalent to a single polynomial equation in t ,

$$P_N(t) := \sum_{k_{-i}=k \in \{\lfloor m_1 N \rfloor, \bar{k}_N, \bar{k}_N + 1\}} \binom{N-1}{k} r_N(k) t^k = 0,$$

where the polynomial's coefficients are the (scaled) policy jumps $\binom{N-1}{k} r_N(k)$. Suppose the coefficients at \bar{k} , $\bar{k}+1$, and $\lfloor m_1 N \rfloor$ are all non-zero. Then, the two jumps at \bar{k}_N and $\bar{k}_N + 1$ have the same sign (if $q(0; \sigma_N) \leq q(1; \sigma_N)$, then $\Pr(\omega = 1|k = \bar{k}; \sigma_N) < \frac{1}{2} \leq \Pr(\omega = 1|k = \bar{k} + 1; \sigma_N)$ and both are weakly positive; if $q(0; \sigma_N) > q(1; \sigma_N)$, both are weakly negative), so the coefficient sequence has *at most one sign change*. This is trivially true if there are at most two non-zero coefficients. Further, since the principal's strategy is non-constant, the polynomial has *some* non-zero coefficients. By Descartes' rule of signs, such a polynomial has at most one positive real root. Since any equilibrium is informative ($q_N(0) \neq q_N(1)$), the corresponding odds ratios t differ across states, hence $U(0; \eta_N)$ and $U(1; \eta_N)$ cannot both be zero.

E Proof of Proposition 3

Take any equilibrium sequence $(\eta_N)_{N \in \mathbb{N}}$ of a quasi-referendum meeting the conditions of Proposition 3. Using Lemma 1, the main text's sketch of proof already established the existence of unique indifferent types $0 < p_N(s) < 1$ for any signal s and any N ; cf. (9).

Suppose that the limit indifferent types are interior, i.e.,

$$0 < \lim_{N \rightarrow \infty} p_N(1) < \lim_{N \rightarrow \infty} p_N(0) < 1. \quad (68)$$

Then the limit of the mean action differs across signals and thus across the two states, i.e., $0 < \lim_{N \rightarrow \infty} q(0; \sigma_N) \neq \lim_{N \rightarrow \infty} q(1; \sigma_N) < 1$. By an application of the law of large numbers, the realized share of actions 1 concentrates around the mean action in each state, implying that the principal learns the state from observing it. Thus, information

aggregates.

It remains to show that information aggregation is also implied by the negation of (93). Note that the negation of (93) implies

$$\lim_{N \rightarrow \infty} p_N(0) = \lim_{N \rightarrow \infty} p_N(1) \in \{0, 1\}. \quad (69)$$

We lead (69) to a contradiction in a generic case (Case 1), and show that it implies information aggregation in the complementary non-generic case (Case 2).

The arguments are based on a detailed analysis of point events for the realized number k_{-i} of actions 1 of the other agents: For any sequence $(m_N)_{N \in \mathbb{N}}$ with $m_N N \in \mathbb{N}$ for all N , we apply (17) to obtain

$$\Pr(k_{-i} = m_N N | \omega = \omega'; \sigma_N, N) = \exp\left(- (N-1) \text{KL}(m_N, q(\omega', \sigma_N)) + o(N)\right),$$

and the left equation in (18) to obtain

$$\begin{aligned} \frac{\Pr(\omega = 1 | k_{-i} = m_N N; \sigma_N, N)}{\Pr(\omega = 0 | k_{-i} = m_N N; \sigma_N, N)} &= \frac{\Pr(\omega = 1)}{1 - \Pr(\omega = 1)} \\ &\cdot \exp\left((N-1) \left(\text{KL}(m_N, q(0; \sigma_N)) - \text{KL}(m_N, q(1; \sigma_N)) \right)\right). \end{aligned} \quad (70)$$

Specifically, we will consider the sequences given by $m'_N = \lfloor \frac{m_1 N}{N} \rfloor$, $m''_N = \frac{\bar{k}_N}{N}$, and $m'''_N = \frac{\bar{k}_N + 1}{N}$, with \bar{k}_N being the unique number satisfying (1). These sequences correspond to the pivotal events piv_1 , $\text{piv}_{0, \bar{k}_N}$, and $\text{piv}_{0, \bar{k}_N + 1}$. We make three preliminary observations: First, as long as $\bar{k}_N \neq N$ for all N large enough, we have

$$\lim_{N \rightarrow \infty} \frac{\Pr(\omega = 1 | k_{-i} = m_N N; \sigma_N, N)}{\Pr(\omega = 0 | k_{-i} = m_N N; \sigma_N, N)} \in \left(\kappa, \frac{1}{\kappa}\right) \text{ for } m_N \in \{m''_N, m'''_N\}, \quad (71)$$

for some $\kappa > 0$, by the defining property (1) of \bar{k}_N (whenever we apply (71), we will rule out the case $\bar{k}_N = N$). Second, (69) implies

$$\lim_{N \rightarrow \infty} \frac{\Pr(\omega = 1 | k_{-i} = m_N N; \sigma_N, N)}{\Pr(\omega = 0 | k_{-i} = m_N N; \sigma_N, N)} \in \{0, \infty\} \text{ for } m_N = m'_N. \quad (72)$$

Otherwise the inference from piv_1 would be bounded as $N \rightarrow \infty$. Then, for any N large enough, we would have either $\bar{k}_N = N$, so that only piv_1 would be relevant for the agents' best response, or, by (71), the inference from piv_0 would also be uniformly bounded. In either case, the indifferent types would be bounded away from 0 and 1.³⁴ Third, (69) implies $\lim_{N \rightarrow \infty} q(0; \sigma_N) = \lim_{N \rightarrow \infty} q(1; \sigma_N)$. We set

$$q^* = \lim_{N \rightarrow \infty} q(\omega'; \sigma_N), \text{ and } \Delta_N = q(1; \sigma_N) - q(0; \sigma_N) \text{ for any } N.$$

³⁴The indifferent types are pinned down by (9). Given (71) and (72), boundedness of the indifferent type would follow from setting $a = \Pr(\text{piv}_0 | \omega = 0; \eta_N, N)$, $b = \Pr(\text{piv}_0 | \omega = 1; \eta_N, N)$, $c = \Pr(\text{piv}_1 | \omega = 0; \eta_N, N)$, $d = \Pr(\text{piv}_1 | \omega = 1; \eta_N, N)$, and using the fact that for any $u, v, a, b, c, d > 0$, we have $\min(\frac{a}{b}, \frac{c}{d}) \leq \frac{au+cv}{bu+dv} \leq \max(\frac{a}{b}, \frac{c}{d})$.

Case 1: $m_1 \neq q^*$ ³⁵

In this case we derive a contradiction to (69). For $\gamma > 0$, let $m_N^+(\gamma) = q^* + \gamma$ and $m_N^-(\gamma) = q^* - \gamma$. For any sequence $(m_N)_{N \in \mathbb{N}}$, when $\lim_{N \rightarrow \infty} m_N \neq q^*$, the linear approximation (20) applies, so that

$$(N-1) \left(\text{KL}(m_N, q(0; \sigma_N)) - \text{KL}(m_N, q(1; \sigma_N)) \right) \approx -(N-1) \left(\frac{1-m_N}{1-q^*} - \frac{m_N}{q^*} \right) \Delta_N.$$

For $m_N = m'_N$, the unbounded inference (72) implies

$$\lim_{N \rightarrow \infty} (N-1) \Delta_N \in \{-\infty, \infty\}.$$

We show case by case that

$$\lim_{N \rightarrow \infty} m''_N = \lim_{N \rightarrow \infty} m'''_N = q^*. \quad (73)$$

First suppose $\lim_{N \rightarrow \infty} (N-1) \Delta_N = \infty$. By definition, this implies $q(0; \sigma_N) < q(1; \sigma_N)$ for large N . Using the above linear approximation for $m_N = m_N^\pm(\gamma)$, we see that for any $\gamma > 0$,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\Pr(\omega = 1 | k_{-i} = \lfloor m_N N \rfloor; \sigma_N, N)}{\Pr(\omega = 0 | k_{-i} = \lfloor m_N N \rfloor; \sigma_N, N)} &= 0 \text{ for } m_N = m_N^-(\gamma), \\ \lim_{N \rightarrow \infty} \frac{\Pr(\omega = 1 | k_{-i} = \lfloor m_N N \rfloor; \sigma_N, N)}{\Pr(\omega = 0 | k_{-i} = \lfloor m_N N \rfloor; \sigma_N, N)} &= \infty \text{ for } m_N = m_N^+(\gamma). \end{aligned}$$

In particular, there exist both collective actions k for which the principal's posterior $\Pr(\omega = 1 | k; \sigma_N, N)$ exceeds $\frac{1}{2}$, and others for which it does not. Hence, $\bar{k}_N \neq N$ for large N . The monotonicity of the posterior further implies $\lim_{N \rightarrow \infty} m''_N \in (m_N^-(\gamma), m_N^+(\gamma))$ for all $\gamma > 0$, from which the claim (73) follows. The case in which $\lim_{N \rightarrow \infty} (N-1) \Delta_N = -\infty$ holds is analogous. Since $m_1 \neq q^*$, the relevant divergences differ in the limit, i.e.,

$$0 = \lim_{N \rightarrow \infty} \text{KL}(m_N, q(\omega'; \sigma_N)) < \lim_{N \rightarrow \infty} \text{KL}(m_1, q(\omega'; \sigma_N))$$

for $m_N = m''_N$ and $m_N = m'''_N$. Thus,

$$\lim_{n \rightarrow \infty} \frac{\Pr(k_{-i} = m_N N | \omega, \sigma_N, N)}{\Pr(k_{-i} = m'_N N | \omega; \sigma_N, N)} = \infty$$

for all ω , $m_N = m''_N$, and $m_N = m'''_N$. Since the inference from each of m''_N and m'''_N is bounded, by (71), this implies interior limit cutoffs (cf. (3) and (9)), contradicting the initial assumption (69).

Case 2: $m_1 = q^*$ This case can be decomposed into several analogous subcases; we present one. Consider an equilibrium sequence such that $\lim_{N \rightarrow \infty} \Pr(\omega = 1 | \text{piv}_1; \sigma_N, N) = 1$, and such that, for any N , a type chooses the action 1 if and only if $p_i \leq p_N(s)$. Then, since $p_N(1) < p_N(0) \in (0, 1)$, it holds that $\rho_1 < q(1; \sigma_N) < q(0; \sigma_N) < 1 - \rho_0$. Recall that $p_N(s) \rightarrow 0$ or $p_N(s) \rightarrow 1$ by the assumption (69). If $p_N(s) \rightarrow 1$, then $q^* = 1 - \rho_0$.

³⁵In the main text we assert that this is a generic case. This is true because (69) implies $q^* \in \{\rho_1, 1 - \rho_0\}$.

However, then $\Pr(\text{piv}_1|\omega = 1; \sigma_N, N) \leq \Pr(\text{piv}_1|\omega = 0; \sigma_N, N)$ and $\lim_{N \rightarrow \infty} \Pr(\omega = 1|\text{piv}_1; \sigma_N, N) = 1$ cannot hold. Thus, $p_N(s) \rightarrow 0$, which implies $q^* = \rho_1$.

We now carefully examine the mean actions in each state,

$$q(\omega'; \sigma_N) - m_1 = (1 - \rho_1 - \rho_0) \left(\sum_{s=0,1} \Pr(s_i = s|\omega = \omega') \left(F \left(p_N(s) \right) \right) \right) + \rho_1 - \rho_1.$$

Using simple algebra,³⁶ we see that

$$\frac{\Pr(s_i = 0|\omega = 1)}{\Pr(s_i = 0|\omega = 0)} \leq \lim_{N \rightarrow \infty} \frac{q(1; \sigma_N) - m_1}{q(0; \sigma_N) - m_1} \leq \frac{\Pr(s_i = 1|\omega = 1)}{\Pr(s_i = 1|\omega = 0)},$$

which implies

$$\lim_{N \rightarrow \infty} \frac{q(\omega'; \sigma_N) - m_1}{-\Delta_N} \in (0, \infty) \quad (74)$$

for all ω' . Using the approximation $q(\omega'; \sigma_N) \approx q^*$, we restate the quadratic approximation (19) for $m = m_N$ and $q = q(\omega'; \sigma_N)$ as follows:

$$\text{KL}(m_N, q(\omega'; \sigma_N)) \approx \frac{\left(m_N - q(\omega'; \sigma_N) \right)^2}{2q^*(1 - q^*)}.$$

This approximation yields the following difference in divergences:

$$\begin{aligned} & (N - 1) \left(\text{KL}(m_N, q(0; \sigma_N)) - \text{KL}(m_N, q(1; \sigma_N)) \right) \\ & \approx \frac{(N - 1)}{2q^*(1 - q^*)} \left(2m_N \Delta_N + q(0; \sigma_N)^2 - q(1; \sigma_N)^2 \right) \\ & \approx \frac{(N - 1)}{2q^*(1 - q^*)} \left(2 \left(m_N - q(0; \sigma_N) \right) \Delta_N - \Delta_N^2 \right). \end{aligned}$$

For $m_N = m'_N$, the unbounded inference (72) together with (74) then implies that $\Delta_N^2 N \rightarrow \infty$. Applying the central limit theorem and denoting by $\frac{k_N}{N}$ the realized share of actions 1 among all N agents, we have

$$\lim_{N \rightarrow \infty} \Pr \left(\left| \frac{k_N}{N} - q(\omega'; \sigma_N) \right| < \frac{1}{4} |\Delta_N| \mid \omega = \omega'; \sigma_N, N \right) = 1.$$

Letting $m_N = \frac{k_N}{N}$, we see that with probability approaching one, as $N \rightarrow \infty$

$$\begin{aligned} 2 \left(m_N - q(0; \sigma_N) \right) \Delta_N &> \frac{3}{2} \Delta_N^2 \text{ if } \omega' = 1, \\ 2 \left(m_N - q(0; \sigma_N) \right) \Delta_N &< \frac{1}{2} \Delta_N^2 \text{ if } \omega' = 0; \end{aligned}$$

³⁶To be precise, we use the fact that for any $u, v, a, b, c, d > 0$, we have $\min\left(\frac{a}{b}, \frac{c}{d}\right) \leq \frac{au+cv}{bu+dv} \leq \max\left(\frac{a}{b}, \frac{c}{d}\right)$.

hence,

$$(N-1) \left(\text{KL} \left(m_N, q(0; \sigma_N) \right) - \text{KL} \left(m_N, q(1; \sigma_N) \right) \right) \rightarrow \begin{cases} \infty & \text{if } \omega = 1, \\ -\infty & \text{if } \omega = 0. \end{cases}$$

Given (70), this means the principal learns the state with probability approaching one, as $N \rightarrow \infty$: $\lim_{N \rightarrow \infty} \Pr(\omega = 1 | k_N; \sigma_N, N) = 1$ if the state is 1 and $\lim_{N \rightarrow \infty} \Pr(\omega = 1 | k_N; \sigma_N, N) = 0$ if the state is 0.

F Proof of Theorem 1: The Remaining Equilibrium Existence Result

Theorem 1's proof in the main text relegated a general equilibrium existence result to the appendix, which clarifies that the payoff guarantee of a quasi-referendum is well-defined: For *any* N , *any* quasi-referendum, and *any* agents' information structure, an equilibrium exists.

To show this result, we lift the fixed-point problem to a finite-dimensional space. We claim that equilibria can be represented as fixed points in the space of vectors $(q(0; \sigma), q(1; \sigma), v)$ with $\sum_l v_l = 1$ and $q(\omega'; \sigma) \in [\rho_1, 1 - \rho_0]$ for $\omega' \in \{0, 1\}$. First, a principal's pure strategy is a mapping from the observed number $k \in \{0, \dots, N\}$ of actions 1 to policy choices $x(k) \in \mathcal{P}$. Thus, we can represent mixed strategies by vectors $v = (v_1, \dots, v_{|\mathcal{P}^{N+1}|-1}) \in [0, 1]^{|\mathcal{P}^{N+1}|-1}$ with $\sum_l v_l = 1$. Second, recall that $q(\omega'; \sigma) \in [\rho_1, 1 - \rho_0]$ for $\omega' \in \{0, 1\}$ given the partisans. We then observe that $(q(0; \sigma), q(1; \sigma), v)$ is a sufficient statistic for the best-response correspondence since the principal's inference only depends on $q(0; \sigma)$ and $q(1; \sigma)$, given (1), and since the pivotal events and their likelihood only depend on (\mathbf{q}, v) ; compare to (2), and (3). Together, the observations imply that any equilibrium induces a fixed point vector (\mathbf{q}, v) and vice versa, hence our representation. This representation will allow us a direct application of Kakutani's fixed point theorem to prove the existence result. Kakutani can be applied since the domain of vectors (\mathbf{q}, v) is non-empty, compact and convex. Further, the correspondence that maps each vector (\mathbf{q}, v) to its best response vectors is non-empty, convex, and has a closed graph.

G Proof of Theorem 2

Fix a generalized quasi-referendum Q . The proof of Theorem 2 proceeds in four steps.

1. We first establish information aggregation and show that in any equilibrium sequence, the state-contingent mean actions $q(0; \sigma_N)$ and $q(1; \sigma_N)$ converge to distinct interior limits (Section G.1).
2. We then prove Lemma 2, which states that under the simple majority gateway referendum, the limit mean action in state 1 exceeds the majority cutoff (Section G.2).
3. This implies that the status quo constraint binds only in state 0; in this way, the principal's payoff is guaranteed to be arbitrarily close to the full-information

benchmark.

4. Finally, we show that any other generalized quasi-referendum fails to achieve this payoff guarantee (Section G.3).

G.1 Information Aggregation

As mentioned in the main text, the same proof as that of Proposition 3 establishes information aggregation. The only difference is that the non-generic case (Case 2) can be ruled out immediately, arguing with the properties of the gateway referendum with a simple majority cutoff:

Proposition 3's proof in the appendix starts with the assumption (69) of non-interior limit indifferent types. This assumption implies that either only the 1-partisans choose the action 1 as $N \rightarrow \infty$ or all except the 0-partisans. Thus $q^* = \lim_{N \rightarrow \infty} q(\omega'; \sigma_N) \in \{\rho_1, 1 - \rho_0\}$. Since the simple majority cutoff is $m_1 = \frac{1}{2}$ and since $\rho_1 < \frac{1}{2} < 1 - \rho_0$ (cf. Section 1) the case $m_1 = q^*$ (Case 2) can be ruled out. The other case (Case 1) leads to a contradiction, as before. We conclude that the limit indifferent types must be interior, which implies that the limit mean actions differ across the states, i.e.

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) \neq \lim_{N \rightarrow \infty} q(0; \sigma_N). \quad (75)$$

Since the realized collective action concentrates around the mean action in each state, the principal learns the state from observing it. That is, information aggregates.

G.2 Proof of Lemma 2

We prove Lemma 2 by contradiction. Suppose that

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}.$$

We begin with a preliminary observation. Given (75), the limit mean actions are strictly ordered. Then the principal's cutoff must lie strictly between them:

$$\lim_{N \rightarrow \infty} q(0; \sigma_N) < \lim_{N \rightarrow \infty} \frac{\bar{k}}{N} < \lim_{N \rightarrow \infty} q(1; \sigma_N) \quad \text{or} \quad \lim_{N \rightarrow \infty} q(1; \sigma_N) < \lim_{N \rightarrow \infty} \frac{\bar{k}}{N} < \lim_{N \rightarrow \infty} q(0; \sigma_N), \quad (76)$$

depending on which ordering in (75) holds. To see this, note that if one of the inequalities failed, then

$$\lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}}{N}, q(0; \sigma_N)\right) \neq \lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}}{N}, q(1; \sigma_N)\right),$$

so by (22) the principal's inference from observing the collective action $\frac{\bar{k}}{N}$ would be unbounded. However, this cannot be, given the definition of \bar{k} ; see (1).

Now, we consider the two possibilities in (75) and start with the second, reverse ordering since ruling it out establishes the first ingredient discussed in the proof sketch in the main text.

Step 1: The reverse ordering cannot arise. Suppose that

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) < \lim_{N \rightarrow \infty} q(0; \sigma_N).$$

We distinguish two subcases.

Step 1a: The case $\lim_{N \rightarrow \infty} q(1; \sigma_N) < \frac{1}{2}$. Under the reverse ordering, the principal's posterior

$$\Pr(\omega = 1 \mid k; \sigma_N, N)$$

is decreasing in the realized number k of actions 1. Hence her constrained best response has the following form: $x(k) = 0$ for $k \leq \lfloor N/2 \rfloor$, it jumps up to some $x \geq \varepsilon$ at $k = \lfloor N/2 \rfloor + 1$, and it is weakly decreasing for larger k , while remaining strictly positive.

Since $\lim_{N \rightarrow \infty} q(1; \sigma_N) < \frac{1}{2}$, in state 1 the number of other agents choosing action 1 is Binomial with mode weakly below $\lfloor N/2 \rfloor$. Among all pivotal events, the event

$$k_{-i} = \left\lfloor \frac{N}{2} \right\rfloor$$

therefore has the highest probability in state 1; all other pivotal events lie further above the mode and are thus strictly less likely. At this event, an additional action 1 induces a discrete upward jump from $x = 0$ to some $x \geq \varepsilon$. By contrast, at all other pivotal events it induces only weakly smaller downward adjustments. Hence the average effect of an additional action 1 in state 1 is strictly positive:

$$U(1; \eta_N) > 0.$$

We claim that this contradicts the reverse ordering. To see this, consider the agents' best-response characterization (9). If $U(0; \eta_N) \leq 0$, then all non-partisans strictly prefer action 1, so

$$q(0; \sigma_N) = q(1; \sigma_N) = 1 - \rho_0,$$

a contradiction. If $U(0; \eta_N) > 0$, there are cutoff types $0 < p_N(1) < p_N(0) < 1$ pinned down by (9), and an agent chooses action 1 if and only if $p_i > p_N(s)$. Since signal 1 is more likely in state 1, this implies

$$q(0; \sigma_N) \leq q(1; \sigma_N),$$

again contradicting the reverse ordering. We conclude that the case

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) < \lim_{N \rightarrow \infty} q(0; \sigma_N) \quad \text{and} \quad \lim_{N \rightarrow \infty} q(1; \sigma_N) < \frac{1}{2}$$

cannot arise.

Step 1b: The case $\lim_{N \rightarrow \infty} q(1; \sigma_N) = \frac{1}{2}$. Suppose now that

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) = \frac{1}{2} \quad \text{and} \quad \lim_{N \rightarrow \infty} q(1; \sigma_N) < \lim_{N \rightarrow \infty} q(0; \sigma_N).$$

Then (76) implies that the principal's cutoff lies strictly above the majority cutoff,

$$\frac{1}{2} = \lim_{N \rightarrow \infty} q(1; \sigma_N) < \lim_{N \rightarrow \infty} \frac{\bar{k}}{N} < \lim_{N \rightarrow \infty} q(0; \sigma_N).$$

The large-deviation calculus in (21) and (22) then implies that an agent becomes certain of state 1 conditional on being pivotal. Consequently, all non-partisans choose action 1 in both states, so that

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) = 1 - \rho_0 > \frac{1}{2},$$

contradicting the case assumption that $\lim_{N \rightarrow \infty} q(1; \sigma_N) = \frac{1}{2}$.

We have therefore ruled out the reverse ordering altogether. Hence one must have

$$\lim_{N \rightarrow \infty} q(0; \sigma_N) < \lim_{N \rightarrow \infty} q(1; \sigma_N). \quad (77)$$

Step 2: Under the ordering (77), the event piv_0 does not affect incentives if $\lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}$. Maintain the contradiction hypothesis

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}.$$

Together with (77), this implies

$$\lim_{N \rightarrow \infty} q(0; \sigma_N) < \lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}.$$

By (76), we then have

$$\lim_{N \rightarrow \infty} \frac{\bar{k}}{N} < \frac{1}{2}. \quad (78)$$

The observation (78) implies that, conditional on piv_0 , the principal is constrained to choose $x = 0$ regardless of the action of the fixed agent i . Hence the event piv_0 does not affect agent i 's incentives. The only relevant pivotal event is therefore the referendum-pivotal event

$$k_{-i} = \left\lfloor \frac{N}{2} \right\rfloor.$$

Step 3: Standard pivotal updating now yields a contradiction. Since (77) and the contradiction hypothesis $\lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}$ imply

$$\lim_{N \rightarrow \infty} \text{KL}\left(\frac{1}{2}, q(0; \sigma_N)\right) > \lim_{N \rightarrow \infty} \text{KL}\left(\frac{1}{2}, q(1; \sigma_N)\right),$$

the standard large-deviation calculus in (21) and (22) yields

$$\lim_{N \rightarrow \infty} \frac{\Pr(\omega = 1 \mid \text{piv}; \eta_N, N)}{\Pr(\omega = 0 \mid \text{piv}; \eta_N, N)} = \infty.$$

That is, conditional on being pivotal, an agent becomes certain that the state is 1. Consequently, all non-partisans choose action 1, so that

$$\lim_{N \rightarrow \infty} q(1; \sigma_N) = 1 - \rho_0 > \frac{1}{2},$$

contradicting the maintained assumption $\lim_{N \rightarrow \infty} q(1; \sigma_N) \leq \frac{1}{2}$.

This contradiction proves Lemma 2.

G.3 Payoff Guarantees

Information aggregation together with Lemma 2 implies that the only potential loss arises in state 0 if the principal is forced to choose some $x > 0$. In this worst case, she then chooses the smallest feasible policy, $x = \varepsilon$. Consequently, the payoff guarantee of the gateway referendum Q^* with the simple majority cutoff has the lower bound

$$B = 1 - \frac{\Pr(\omega = 0)}{\Pr(\omega = 1)} \varepsilon > 1 - \varepsilon$$

What remains to be shown to prove Theorem 2's claim that Q^* has a maximal payoff guarantee is that no generalized quasi-referendum Q has a payoff guarantee that exceeds the lower bound B .

In the following, we assume that $G(Q) > 1 - \frac{\Pr(\omega=0)}{\Pr(\omega=1)} \varepsilon$ for some Q and lead this to a contradiction.

We start with two preliminary observations: First, Theorem 1 implies $G(Q) \leq 1 - \varepsilon$ whenever $Q(m) \neq \{0\}$ for all m . So $Q(m) = \{0\}$ for some m . Second, we argue that, for all $j = 1, \dots, R + 1$, either $Q(m_j) = \{0\}$ or $\{0, 1\} \subseteq Q(m_j)$. If $Q(m_j) \neq \{0\}$, one can construct approximately truthful equilibrium sequences where $Q(m_j)$ becomes almost certainly binding, as $N \rightarrow \infty$, using the identical argument as in Proposition 1's proof sketch. If then $1 \notin Q(m_j)$ or $0 \notin Q(m_j)$, the principal's payoff in these equilibrium sequences is at most B , as $N \rightarrow \infty$. Thus, $Q(m_j) \neq \{0\}$ implies $\{0, 1\} \subseteq Q(m_j)$, under the assumption that $G(Q) > B$.

The rest of the proof distinguishes three cases.

Case 1. Suppose $Q(0) = \{0\}$ and consider an information structure for which $\rho_1 < m_1 < 1 - \rho_0$. By the identical logic as in Section 3, for any large enough N , there is a “deadlock” equilibrium in which the agents use a strategy with $\rho_1 < q(1) < q(0) < m_1$ and so that $\Pr(\omega = 1 | k = \lfloor m_1 N \rfloor + 1; \sigma_N, N) = \frac{1}{2}$ and the principal's best response is 0 after each m . This implies an upper bound of 0 for $G(Q)$, in contradiction to the assumption $G(Q) > 1 - \frac{\Pr(\omega=0)}{\Pr(\omega=1)} \varepsilon$.

Case 2. Suppose $Q(1) = \{0\}$ and consider an information structure for which $\rho_1 < m_R < 1 - \rho_0$. Then, likewise, for any large enough N , there is a “deadlock” equilibrium in which the agents use a strategy with $m_R < q(0) < q(1) < 1 - \rho_0$ and so that $\Pr(\omega = 1 | k = \lfloor m_R N \rfloor + 1; \sigma_N, N) = \frac{1}{2}$ and the principal's best reply is 0 after each m . Again, this implies a contradiction to $G(Q) > 1 - \frac{\Pr(\omega=0)}{\Pr(\omega=1)} \varepsilon$.

Case 3. Suppose $Q(m_j) = \{0\}$ for some $j \in \{2, \dots, R\}$, $Q(0) \neq \{0\}$ and $Q(1) \neq \{0\}$. We argue that for large enough N , there would exist an information structure and an

equilibrium in which the agents choose a strategy σ_N with $q = q(0; \sigma_N) = q(1; \sigma_N)$. Since given such σ , the agents' collective action transmits no information, the principal's best response is her prior-optimal choice 1 if $1 \in Q(m)$ and 0 if $Q(m) = \{0\}$. The principal's payoff from this equilibrium is weakly worse than choosing 1 for all m ; so a sequence of such equilibria would imply an upper bound for her payoff guarantee of $1 - \frac{\Pr(\omega=0)}{\Pr(\omega=1)} < B$, in contradiction to the assumption that $G(Q) > B$.

The following shows that such an equilibrium would indeed exist. Fixing the principal's best response described above (she chooses 1 if $1 \in Q(m_j)$ and 0 otherwise), it is sufficient to show that there is a strategy σ and a q so that $q = q(0; \sigma_N) = q(1; \sigma_N)$ and $U(0; \eta) = U(1; \eta) = 0$. Then, any strategy of the agents is a best response, including σ .

To prepare the argument, let j' denote the minimal number in $\{2, \dots, R\}$ for which $Q(m_{j'}) = \{0\}$ and j'' the minimal number in $\{j', \dots, R\}$ for which $Q(m_{j''+1}) \neq \{0\}$. This guarantees that the principal's choice switches from 1 to 0 when k switches from $\lfloor m_{j'-1}N \rfloor$ to $\lfloor m_{j'-1}N \rfloor + 1$, and it switches from 0 to 1 when k switches from $\lfloor m_{j''}N \rfloor$ to $\lfloor m_{j''}N \rfloor + 1$. So, in the pivotal event $\text{piv}_{j'-1}$, an additional action 1 moves the policy downwards, and in the pivotal event $\text{piv}_{j''}$, an additional action 1 moves the policy upwards. Now, consider an information structure with $\rho_1 < m_{j'-1}$ and $m_{j''} < 1 - \rho_0$; this constraint on the partisans ensures that for any $q \in [m_{j'-1}, m_{j''}]$, there is a strategy σ with $q = q(0; \sigma) = q(1; \sigma)$.

Consider σ so that $q(\omega'; \sigma) = q = m_{j'-1}$. Then, (21) implies $\lim_{N \rightarrow \infty} \Pr(\text{piv}_{j'-1} | \text{piv}; \eta, N) = 1$. Since an additional action 1 moves the policy downwards conditional on $\text{piv}_{j'-1}$, the average effect is negative, i.e. $U(\omega'; \eta) < 0$ for $\omega' \in \{0, 1\}$. Conversely, consider σ is so that $q = m_{j''}$. Then, (21) implies $\lim_{N \rightarrow \infty} \Pr(\text{piv}_{j''} | \text{piv}; \eta, N) = 1$. Since an additional action 1 moves the policy upwards conditional on $\text{piv}_{j''}$, the average effect is positive, i.e. $U(\omega'; \eta) > 0$ for $\omega' \in \{0, 1\}$. Finally, an application of the intermediate value theorem implies there is a strategy σ with $q(\omega'; \sigma) = q \in (m_{j'-1}, m_{j''})$ and $U(0; \eta) = U(1; \eta) = 0$. As explained above, this finishes the proof by contradiction in this last case.

References

- ACHARYA, A. and MEIROWITZ, A. (2017). Sincere voting in large elections. *Games and Economic Behavior*, **101**, 121–131.
- AHN, D. S. and OLIVEROS, S. (2012). Combinatorial voting. *Econometrica*, **80** (1), 89–141.
- ALI, S. N., MIHM, M. and SIGA, L. (2025). The political economy of zero-sum thinking. *Econometrica*, **93** (1), 41–70.
- ALONSO, R. and MATOUSCHEK, N. (2008). Optimal delegation. *The Review of Economic Studies*, **75** (1), 259–293.
- ASHWORTH, S. (2012). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science*, **15** (1), 183–201.
- AUSTEN-SMITH, D. and BANKS, J. S. (1996). Information aggregation, rationality, and the condorcet jury theorem. *American political science review*, **90** (1), 34–45.
- BARELLI, P., BHATTACHARYA, S. and SIGA, L. (2022). Full information equivalence in large elections. *Econometrica*, **90** (5), 2161–2185.

- BARRO, R. J. (1973). The control of politicians: an economic model. *Public choice*, pp. 19–42.
- BATTAGLINI, M. (2017). Public protests and policy making. *The Quarterly Journal of Economics*, **132** (1), 485–549.
- BHATTACHARYA, S. (2013). Preference monotonicity and information aggregation in elections. *Econometrica*, **81** (3), 1229–1247.
- (2018). *Condorcet Jury theorem in a spatial model of elections*. Working paper.
- CASTANHEIRA, M. (2003). Why vote for losers? *Journal of the European Economic Association*, **1** (5), 1207–1238.
- CHEN, K. (2025). *Communication with Multiple Senders*. Working paper.
- COUGHLAN, P. J. (2000). In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *American Political science review*, **94** (2), 375–393.
- CRAMÉR, H. (1938). Sur un nouveau théoreme-limite de la théorie des probabilités. *Colloqué consacré a la théorie des probabilités, Actualités scientifiques et industrielles* 736.
- DAMIANO, E., LI, H. and SUEN, W. (2025). *Informative Voting in Large Elections*. Working paper.
- EKMEKCI, M. and LAUERMAN, S. (2020). Manipulated electorates and information aggregation. *The Review of Economic Studies*, **87** (2), 997–1033.
- ESCHER, F. (1932). On the probability function in the collective theory of risk. *Skand. Aktuarie Tidsskr.*, **15**, 175–195.
- FARQUHARSON, R. (1969). *Theory of Voting*. Yale University Press.
- FEDDERSEN, T. and PESENDORFER, W. (1997). Voting behavior and information aggregation in elections with private information. *Econometrica: Journal of the Econometric Society*, pp. 1029–1058.
- FEREJOHN, J. (1986). Incumbent performance and electoral control. *Public choice*, pp. 5–25.
- FERNANDEZ, R. and RODRIK, D. (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, pp. 1146–1155.
- FEY, M. (1997). Stability and coordination in duverger’s law: A formal model of pre-election polls and strategic voting. *American Political science review*, **91** (1), 135–147.
- HERRERA, H., LLORENTE-SAGUER, A. and MCMURRAY, J. C. (2019). The marginal voter’s curse. *The Economic Journal*, **129** (624), 3137–3153.
- HOLMSTRÖM, B. (1984). On the theory of delegation.

- JOHNSON, N. L., KEMP, A. W. and KOTZ, S. (2005). *Univariate discrete distributions*. John Wiley & Sons.
- KARTIK, N., VAN WEELDEN, R. and WOLTON, S. (2017). Electoral ambiguity and political representation. *American Journal of Political Science*, **61** (4), 958–970.
- KRISHNA, V. and MORGAN, J. (2012). Voluntary voting: Costs and benefits. *Journal of Economic Theory*, **147** (6), 2083–2123.
- LEDYARD, J. O. and PALFREY, T. R. (1999). A characterization of interim efficiency with public goods. *Econometrica*, **67** (2), 435–448.
- LEVIT, D. and MALENKO, N. (2011). Nonbinding voting for shareholder proposals. *The Journal of Finance*, **66** (5), 1579–1614.
- LOHMANN, S. (1994). Information aggregation through costly political action. *The American Economic Review*, pp. 518–530.
- MANDLER, M. (2012). The fragility of information aggregation in large elections. *Games and Economic Behavior*, **74** (1), 257–268.
- MEIROWITZ, A. and SHOTTS, K. W. (2009). Pivots versus signals in elections. *Journal of Economic Theory*, **144** (2), 744–771.
- MORGAN, J. and STOCKEN, P. C. (2008). Information aggregation in polls. *American Economic Review*, **98** (3), 864–896.
- OECD (2020). *Innovative citizen participation and new democratic institutions*. OECD Publications Centre.
- PALFREY, T. R. (2009). Laboratory experiments in political economy. *Annual Review of Political Science*, **12** (1), 379–388.
- PIKETTY, T. (2000). Voting as communicating. *The Review of Economic Studies*, **67** (1), 169–191.
- RAZIN, R. (2003). Signaling and election motivations in a voting model with common values and responsive candidates. *Econometrica*, **71** (4), 1083–1119.
- SELTEN, R. (1988). *Reexamination of the perfectness concept for equilibrium points in extensive games*. Springer.
- WOLINSKY, A. (2002). Eliciting information from multiple experts. *Games and Economic Behavior*, **41** (1), 141–160.

Online Appendix

H Construction of the Agents' Deadlock Equilibrium Strategy

We prove the remaining claim from Section 3 that there are agents' information structures and agents' strategies σ_N with mean actions $0 < q(1; \sigma_N) < q(0; \sigma_N) < m_1$ for which

$$\Pr(\omega = 1 | k = \lfloor m_1 N \rfloor + 1; \sigma_N, N) = \frac{1}{2}.$$

Take any $m_1 \in (0, 1)$, any pair of mean actions $\mathbf{q} = (q(0), q(1))$ with $0 < q(1) \leq q(0) < m_1$, and let $m' = \frac{\lfloor m_1 N \rfloor + 1}{N}$. Then, note that

$$\frac{\Pr(k = m'N | \omega = 1)}{\Pr(k = m'N | \omega = 0)} = \exp\left(-N\left(\text{KL}(m', q(1)) - \text{KL}(m', q(0))\right)\right),$$

by an application of (17).

If $q(1) = q(0)$, the principal learns nothing from her observations and

$$\Pr(\omega = 1 | k = m'N; \mathbf{q}, N) = \Pr(\omega = 1) > \frac{1}{2}.$$

If $q(1) < q(0)$, then $\lim_{N \rightarrow \infty} \text{KL}(m', q(1)) - \text{KL}(m', q(0)) > 0$, so

$$\lim_{N \rightarrow \infty} \Pr(\omega = 1 | k = m'N; \mathbf{q}, N) = 0.$$

By an application of the intermediate value theorem, for any N large enough, there are mean actions $0 < q(1) < q(0) < m_1$ for which

$$\Pr(\omega = 1 | k = m'; \mathbf{q}, N) = \frac{1}{2}.$$

For any such $q(0)$ and $q(1)$, we can always find a strategy σ that induces these mean actions whenever the share of partisans choosing action 1 is not too large, $\rho_1 < q(1)$.

I Optimality for a Given Information Structure

In Section 5.1 of the main text, we stated Propositions 4 and 5. Proposition 4 concerns the case in which the principal's prior is sufficiently high, and its key step is that monotone quasi-referenda with a single cutoff $m_1 = \frac{1}{2}$ admit equilibrium sequences along which the lower policy region $Q(0)$ binds with probability approaching one. We prove this formally here, with an argument analogous to that underlying Proposition 2. Proposition 5 concerns instead a sufficiently low principal's prior, and the corresponding argument is symmetric.

Proposition 6. *Fix the agents' information structure and a monotone quasi-referendum with single cutoff $m_1 = \frac{1}{2}$. If the principal's prior $\Pr(\omega = 1)$ is sufficiently high, then*

there exist equilibrium sequences in which the realized policy set is $Q(0)$ with probability converging to one as $N \rightarrow \infty$.

Sketch. For any large enough N , we construct an equilibrium strategy σ_N with the mean action below the cutoff in each state,

$$q(0; \sigma_N) < q(1; \sigma_N) < \frac{1}{2} - \gamma, \quad (79)$$

for some $\gamma > 0$. An application of the law of large numbers then yields the claim.

We use the same parametric family of cutoff strategies σ_L as in the proof of Proposition 2, defined in (60). As shown there,

$$q(0; \sigma_L) < q(1; \sigma_L) \quad \text{for all } L > 0.$$

We also use the same compact parameter interval $[\underline{L}, \bar{L}]$ from (61). As in the proof of Proposition 2, the posterior likelihood ratio conditional on the pivotal event piv_0 lies in the interval $[\underline{L}, \bar{L}]$; cf. (62).

The critical lemma in the proof of Proposition 2 was Lemma 8, and its analogue is the following.

Lemma 9 (Self-map lemma). *For any $\rho_1 < \bar{q} < \frac{1}{2}$, if $\Pr(\omega = 1)$ is sufficiently high, then there exist $\bar{N} \in \mathbb{N}$ and $\bar{\delta} > 0$ such that, for every $N \geq \bar{N}$, every $L \in [\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}]$, and every principal best response (\bar{k}_N, \bar{x}_N) to σ_L ,*

(i)

$$q(0; \sigma_L) < q(1; \sigma_L) < \bar{q},$$

(ii) *the agents' best response to σ_L and (\bar{k}_N, \bar{x}_N) is again a cutoff strategy $\sigma_{L'}$ with*

$$L' \in [\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}].$$

The proof is analogous to that of Lemma 8, except that part (i) is now based on a uniform *upper* bound on $q(1; \sigma_L)$ rather than a lower bound on $q(0; \sigma_L)$. When $\Pr(\omega = 1)$ is sufficiently high, the interval $[\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}]$ lies far to the right, so the cutoff types $p_L(s)$ are uniformly close to 1 for any L from this interval. Hence, for every $\bar{q} \in (\rho_1, \frac{1}{2})$, all sufficiently high $\Pr(\omega = 1)$, and all sufficiently small $\bar{\delta} > 0$,

$$q(0; \sigma_L) < q(1; \sigma_L) < \bar{q} \quad \text{for all } L \in [\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}].$$

Part (ii) then follows exactly as in Lemma 8: Fix $L \in [\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}]$. First, we show that the principal's cutoff $\frac{\bar{k}_N}{N}$ lies strictly between the two mean actions when N is large, i.e. (64) holds. Second, conditional on being pivotal and as $N \rightarrow \infty$, an agent almost certainly affects the principal's preference over policies instead of the feasible policy set, i.e. (66) holds. Third, there is a uniform bound \bar{N} so that for any $N \geq \bar{N}$, the agents' best response is again a cutoff strategy with parameter in the same interval.

Finally, we conclude exactly as in the proof of Proposition 2. For every $N \geq \bar{N}$, let Γ_N denote the correspondence on $[\underline{L} - \bar{\delta}, \bar{L} + \bar{\delta}]$ that maps a parameter L to the set of parameters L' generated by:

(a) choosing a principal best response (\bar{k}_N, \bar{x}_N) to σ_L ,

(b) choosing an agents' best response $\sigma_{L'}$ to σ_L and (\bar{k}_N, \tilde{x}_N) .

By Lemma 9, Γ_N maps the interval into itself. Further, it has non-empty, compact, convex values and a closed graph. Kakutani's fixed point theorem therefore yields a fixed point L_N^* for every $N \geq \bar{N}$. The corresponding strategy $\sigma_{L_N^*}$ is an equilibrium strategy. By Lemma 9(i),

$$q(0; \sigma_{L_N^*}) < q(1; \sigma_{L_N^*}) < \bar{q}.$$

Choose $\gamma > 0$ such that $\bar{q} < \frac{1}{2} - \gamma$. Then

$$q(0; \sigma_{L_N^*}) < q(1; \sigma_{L_N^*}) < \frac{1}{2} - \gamma,$$

so the equilibrium sequence satisfies (79), as claimed. \square

J Single-Plateau Preferences

We consider the single-plateau setting that we described in Section 5.3.

- All players have common utilities $u(x, \omega) = -c(x) + b(x)\omega$ for $\omega = 0, 1$, with $u(x, \omega)$ continuously differentiable in $[0, 1]$ for $\omega \in \{0, 1\}$;
- $u'(x, 0) < 0$ for all $x \in [0, 1]$;
- there is a unique optimal policy given the prior as well as given certainty that the state is 1;
- $\frac{c'(x)}{b'(x)}$ is strictly increasing. This is the novel assumption, relative to the baseline model.

Theorem 3 shows that the gateway referendum with the simple-majority cutoff achieves a payoff guarantee above $1 - \varepsilon$, as in the baseline setting with linear preferences.

Theorem 3. *In the single-plateau setting, the gateway referendum with the simple-majority cutoff aggregates information across all equilibrium sequences and agents' information structures. It has a payoff guarantee that strictly exceeds $1 - \varepsilon$.*

The proof proceeds as follows.

1. First, we generalize various definitions from the baseline model and characterize the players' best responses (Sections J.1 and J.2).
2. Second, the same argument as in the baseline proof shows that any equilibrium must be informative.
3. Third, we prove an analog of Lemma 1 stated as Lemma 10 below (Section J.3).
4. Fourth, once this analog is established, the information-aggregation and payoff-guarantee arguments follow by the same steps as in the baseline proof (Section J.4).

Lemma 10. *In the single-plateau setting, any equilibrium η_N of the gateway referendum with simple-majority cutoff satisfies*

$$\begin{aligned} & \text{either } U(\omega'; \eta_N) > 0 \text{ for all } \omega' \in \{0, 1\}, \\ & \text{or } U(\omega'; \eta_N) < 0 \text{ for all } \omega' \in \{0, 1\}, \end{aligned}$$

with the average effects defined in (15) and (16).

A remark up front: Given the gateway referendum, the principal will not choose any policy above $x^*(1) = \arg \max_{x \in \mathcal{P}} u(x, 1)$ for any belief about the state. Thus, it is without loss to restrict the policy space to $\mathcal{P}' = \mathcal{P} \setminus \{x : x > x^*(1)\}$. To be consistent with the earlier normalization, we now normalize $x^*(1) = 1$. This way,

$$u(x_j, 0) < u(x_i, 0) \text{ and } u(x_j, 1) > u(x_i, 1) \quad (80)$$

for all $x_i, x_j \in \mathcal{P}'$ with $x_i < x_j$, and $u'(x, 1) > 0$ for all $x \in [0, 1)$.

J.1 Preliminaries: Principal's Best Response

We describe the principal's best response correspondence; first, in terms of belief cutoffs, and then in terms of cutoffs for the observed number k of actions 1

Lemma 11. *In the single-plateau setting, the principal's best response takes an interval form. There are cutoffs $0 < p_{2,1} < p_{3,2} < \dots < p_{l,l-1} < p_{l+1,l} := 1$ so that for $j = 2, \dots, l$,*

1. *The principal is indifferent between x_j and x_{j-1} if she believes the likelihood of state 1 is $p_{j,j-1}$,*
2. *she prefers x_j over all other policies $x \in \mathcal{P}$ if $\Pr(\omega = 1|k; \sigma, N) \in (p_{j,j-1}, p_{j+1,j})$.*

Finally, she prefers x_1 over all other policies $x \in \mathcal{P}$ if $\Pr(\omega = 1|k; \sigma, N) < p_{2,1}$.

Proof. Set

$$p_{j,i} := \frac{c(x_j) - c(x_i)}{b(x_j) - b(x_i)} \text{ for } i < j \text{ with } i, j \in 1, \dots, l$$

A policy x_j is preferred over a policy $x_i < x_j$ at a belief $p = \Pr(\omega = 1)$ if and only if

$$\begin{aligned} \mathbb{E}\left(u(x_j, \omega)|p\right) > \mathbb{E}\left(u(x_i, \omega)|p\right) &\Leftrightarrow -c(x_j) + b(x_j)p > -c(x_i) + b(x_i)p \\ &\Leftrightarrow p > \frac{c(x_j) - c(x_i)}{b(x_j) - b(x_i)}. \end{aligned}$$

Since $u(x_j, 0) < u(x_i, 0)$ and $u(x_j, 1) > u(x_i, 1)$, the cutoffs are interior, i.e. $p_{j,i} \in (0, 1)$, and the principal is indifferent between x_j and x_i if and only if $\Pr(\omega = 1|k; \sigma, N) = p_{j,i}$. Finally, we establish the claimed ordering of the belief cutoffs: For each $j = 2, \dots, l$ and the adjacent pair (x_{j-1}, x_j) , Cauchy's mean value theorem yields some $\xi_j \in (x_{j-1}, x_j)$ such that

$$p_{j,j-1} = \frac{c(x_j) - c(x_{j-1})}{b(x_j) - b(x_{j-1})} = \frac{c'(\xi_j)}{b'(\xi_j)}.$$

Since $\frac{c'}{b}$ is strictly increasing and ξ_j are increasing in j , it follows that

$$p_{2,1} < p_{3,2} < \dots < p_{l,l-1}.$$

Finally, a transitivity argument implies that she prefers x_j over all other policies $x \in \mathcal{P}$ if $\Pr(\omega = 1|k; \sigma, N) \in (p_{j,j-1}, p_{j+1,j})$. \square

Next, we give a parallel description in terms of cutoffs for the observed number k of actions 1: If $q(0; \sigma) \leq q(1; \sigma)$, the posterior $\Pr(\omega = 1|k; \sigma, N)$ is increasing in k . Either the prior-optimal policy is uniquely optimal for all k , and we set $\bar{k}_{j,j-1} = N$ for all $j \in \{2, \dots, l\}$; or, the posteriors cross one of the belief cutoffs. For each $j \in \{2, \dots, l\}$ such that the posterior crosses $p_{j,j-1}$, there is a minimal $-1 \leq \bar{k}_{j,j-1} < N$ so that³⁷

$$\Pr(\omega = 1|\bar{k}_{j,j-1}; \sigma, N) < p_{j,j-1} \leq \Pr(\omega = 1|\bar{k}_{j,j-1} + 1; \sigma, N).$$

Whenever such $\bar{k}_{j,j-1}$ does not exist for some $j \in \{2, \dots, l\}$, we set $\bar{k}_{j,j-1} = N$.

Analogously, if $q(0; \sigma) > q(1; \sigma)$, then, the posterior $\Pr(\omega = 1|k; \sigma, N)$ is decreasing in k . So, either the prior-optimal policy is uniquely optimal for all k , and we set $\bar{k}_{j,j-1} = N$ for all $j \in \{2, \dots, l\}$; or, the posteriors cross one of the belief cutoffs. For each $j \in \{2, \dots, l\}$ such that the posterior crosses $p_{j,j-1}$, there is a minimal $0 \leq \bar{k}_{j,j-1} < N$ so that

$$\Pr(\omega = 1|\bar{k}_{j,j-1}; \sigma, N) > p_{j,j-1} \geq \Pr(\omega = 1|\bar{k}_{j,j-1} + 1; \sigma, N),$$

Whenever such $\bar{k}_{j,j-1}$ does not exist for some $j \in \{2, \dots, l\}$, we set $\bar{k}_{j,j-1} = N$.

J.2 Preliminaries: Agents' Best Response

We start by generalizing several definitions from the baseline model.

First, we set $\text{piv}_0 = \cup_{j=2, \dots, l: \bar{k}_{j,j-1} < N} \cup_{k \in \{\bar{k}_{j,j-1}, \bar{k}_{j,j-1} + 1\}} \text{piv}^k$ where piv^k denotes the event that $k_{-i} = k$ (previously, the principal could only be indifferent at the cutoff \bar{k} , so the only possible pivotal events were \bar{k} and $\bar{k} + 1$), and $\text{piv} = \cup_{j=0, \dots, R} \text{piv}_j$ with piv_j defined as before for $j > 0$. Second, the average effect $U(\omega'; \eta)$ of an additional agent's choosing action 1 in state ω' is defined via (15) and (16).

Now, if agent i has signal $s_i = s$ and type $p_i = p$, then he prefers action 1 if

$$\Pr_i(\omega = 1 | p_i = p, s_i = s) U(1; \eta) - \Pr_i(\omega = 0 | p_i = p, s_i = s) U(0; \eta) > 0, \quad (81)$$

given a strategy profile η .

J.3 Proof of Lemma 10

We show that for any equilibrium η_N and any agents' information structure,

$$U(0; \eta_N) \neq 0 \text{ or } U(1; \eta_N) \neq 0. \quad (82)$$

Lemma 10 follows since $U(0; \eta_N)$ and $U(1; \eta_N)$ must have the same sign. If the two average effects had (weakly) opposite signs, then the best-response inequality (81) would point *all* non-partisan types in the same direction regardless of their signal, producing an

³⁷We abuse notation here and set $\Pr(\omega = 1|k = -1; \sigma, N) = 0$.

uninformative best response and contradicting the informativeness of equilibrium (which can be shown by the same argument as in the baseline proof, as noted).

To prepare the proof of (82), we introduce

$$t = \frac{q}{1-q} \in [0, \infty), \text{ and note } (1-q)^{N-1} = (1+t)^{-(N-1)}.$$

For each ω , define

$$U(1, q) := (1-q)^{N-1} F_1(t), \quad U(0, q) := -(1-q)^{N-1} F_0(t)$$

where

$$F_\omega(t) := \sum_{k_{-i}=0}^{N-1} \binom{N-1}{k_{-i}} t^{k_{-i}} r(k_{-i}, \omega),$$

and

$$r(k, \omega) := \mathbb{E} \left(u(x(k+1), \omega) - u(x(k), \omega) \right).$$

Let

$$q_0 := q(0; \sigma_N), \quad q_1 := q(1; \sigma_N).$$

Then

$$U(0; \eta_N) = U(0, q_0), \quad U(1; \eta_N) = U(1, q_1).$$

Thus it suffices to show that

$$U(0, q_0) \neq 0 \quad \text{or} \quad U(1, q_1) \neq 0.$$

The proof of (82) is by contradiction. Assume that

$$U(0; q_0) = U(1; q_1) = 0. \tag{83}$$

Case 1. $q(0; \sigma_N) \leq q(1; \sigma_N)$

Given the case assumption, the principal's posterior is increasing in k ; thus the same holds for any principal strategy $x(\cdot)$ in the support: $x(k+1) \geq x(k)$ for all k . Since $u(\cdot, 0)$ is strictly decreasing, it follows that $r(k, 0) \leq 0$ for all k . Moreover, $x(\cdot)$ is non-constant (e.g., $x(0) = 0$ and $x(N) \geq \varepsilon$ by definition of the quasi-referendum), hence $x(k+1) > x(k)$ for at least one k , which implies $r(k, 0) < 0$ for that k . Therefore $F_0(t) < 0$ for all $t > 0$, and thus $U(0, q) > 0$ for all $q \in (0, 1)$, in contradiction to (83).

Case 2. $q(1; \sigma_N) < q(0; \sigma_N)$.

The second case requires several auxiliary lemmas. The first shows that $U(1, q)$ has no root around $q = 0$. The second bounds the number of positive roots of $U(\omega; \cdot)$ to be at most 1 for each ω . The third shows that $U(1, \cdot)$ is strictly positive at any root of $U(0, \cdot)$. We derive the contradiction to (83) after the statement of the third lemma. Let $\kappa := \lfloor N/2 \rfloor$.

Lemma 12. $U(1, q) > 0$ for all sufficiently small $q > 0$.

Proof. By definition of the quasi-referendum, $x(k) = 0$ for $k \leq \kappa$ implies $u(x(k+1), 1) - u(x(k), 1) = 0$ for all $k < \kappa$, while $x(\kappa+1) > x(\kappa) = 0$ and strict monotonicity of $u(\cdot, 1)$ imply $u(x(\kappa+1), 1) - u(x(\kappa), 1) > 0$. Thus $F_1(t) = t^\kappa G(t)$ for a polynomial G with

$$G(0) = \binom{N-1}{\kappa} \left(\mathbb{E}(u(x(\kappa+1), 1)) - u(0, 1) \right) > 0,$$

so $F_1(t) > 0$ for all sufficiently small $t > 0$, and hence $U(1, q) > 0$ for all sufficiently small $q > 0$. \square

Lemma 13. For each $\omega \in \{0, 1\}$, the function $U(\omega, q)$ has at most one positive root (counting multiplicities).

Proof. Since $(1-q)^{N-1} > 0$ for $q \in (0, 1)$, strictly positive roots $q \in (0, 1)$ of $U(\omega, \cdot)$ correspond to strictly positive roots $t > 0$ of F_ω .

For every principal's strategy $x(\cdot)$ in the support, we have

$$x(k) = 0 \text{ for all } k \leq \kappa, \quad x(\kappa+1) > x(\kappa) = 0, \quad \text{and } x(k+1) \leq x(k) \text{ for all } k \geq \kappa+1.$$

Since $u(\cdot, 1)$ is strictly increasing while $u(\cdot, 0)$ is strictly decreasing, it follows

$$u(x(k+1), 1) - u(x(k), 1) = \begin{cases} 0 & \text{for } k < \kappa, \\ > 0 & \text{for } k = \kappa, \\ \leq 0 & \text{for } k > \kappa, \end{cases} \quad u(x(k+1), 0) - u(x(k), 0) = \begin{cases} 0 & \text{for } k < \kappa, \\ < 0 & \text{for } k = \kappa, \\ \geq 0 & \text{for } k > \kappa. \end{cases}$$

Taking expectations preserves these (weak/strict) inequalities, so for each $\omega \in \{0, 1\}$ the coefficient sequence $(r(k, \omega))_{k=0}^{N-1}$ of F_ω has at most one sign change (after omission of zero coefficients) and is non-constant. By Descartes' rule of signs, each F_ω has at most one strictly positive real root (counting multiplicities), and therefore also $U(\omega, \cdot)$. \square

Lemma 14. If $U(0, q_0) = 0$, then, $U(1, q_0) > 0$.

We postpone the proof of Lemma 14. First we show how the three lemmas imply a contradiction to (83) as follows. Recall our notation

$$q_0 = q(0; \sigma_N), \quad q_1 = q(1; \sigma_N),$$

so that in Case 2 we have $0 < q_1 < q_0 < 1$. By Lemmas 12 and 14, $U(1, \cdot)$ is positive near 0 and also at q_0 . We claim that this implies $U(1, q) > 0$ for all $q \in (0, q_0]$ and thus $U(1, q_1) > 0$, which contradicts (83).

Suppose not. Then either:

- (i) $U(1, \cdot)$ becomes strictly negative somewhere on $(0, q_0)$, in which case continuity implies that $U(1, \cdot)$ has at least two distinct positive roots; or
- (ii) $U(1, \cdot)$ stays nonnegative on $(0, q_0]$ but equals zero at some $q \in (0, q_0)$, in which case that root has even multiplicity.

Both possibilities contradict Lemma 13. Therefore $U(1, q) > 0$ for all $q \in (0, q_0]$.

Now, we prove Lemma 14. The proof of Lemma 14 is based on certain algebraic identities implied by $U(0, q_0) = 0$, which we will state in Claim 4. To state the identities, we introduce two central objects (δ_k and r_k) and record their monotonicity properties.

Fix $q \in (0, 1)$ and recall that $t := \frac{q}{1-q}$. For $k \in \{\kappa, \kappa + 1, \dots, N - 1\}$ define the (binomial) likelihood ratio

$$\delta_k := \frac{\binom{N-1}{k} t^k}{\binom{N-1}{\kappa} t^\kappa} = \frac{\Pr(\text{Bin}(N-1, q) = k)}{\Pr(\text{Bin}(N-1, q) = \kappa)}. \quad (84)$$

Note that $\delta_\kappa = 1$ and

$$\frac{\delta_{k+1}}{\delta_k} = \frac{\binom{N-1}{k+1} t^{k+1}}{\binom{N-1}{k} t^k} = \frac{N-1-k}{k+1} t,$$

which is strictly decreasing in k . Hence $(\delta_k)_{k=\kappa}^{N-1}$ crosses 1 at most once.

Next, for every $k > \kappa$ with $r(k, 0) > 0$ we define

$$r_k := \frac{r(k, 0) - r(k, 1)}{r(k, 0)} \quad (85)$$

and show the following monotonicity property.

Claim 2. *If $k' > k > \kappa$ and $r(k, 0), r(k', 0) > 0$, then $r_{k'} \geq r_k$.*

Proof. Fix $k' > k > \kappa$ with $r(k, 0), r(k', 0) > 0$. By definition, $u(x, \omega) = -c(x) + b(x)\omega$, so that

$$r(j, 0) = \mathbb{E}(c(x(j)) - c(x(j+1))), \quad r(j, 0) - r(j, 1) = \mathbb{E}(b(x(j)) - b(x(j+1))),$$

and therefore

$$r_j = \frac{\mathbb{E}(b(x(j)) - b(x(j+1)))}{\mathbb{E}(c(x(j)) - c(x(j+1)))} \quad \text{for } j > \kappa.$$

Since $x(j) \geq x(j+1)$ almost surely, we have

$$c(x(j)) - c(x(j+1)) \geq 0.$$

Moreover, $r(j, 0) > 0$ implies

$$\mathbb{E}(c(x(j)) - c(x(j+1))) > 0.$$

Therefore,

$$r_j = \frac{\mathbb{E}\left(\left(c(x(j)) - c(x(j+1))\right) \frac{b(x(j)) - b(x(j+1))}{c(x(j)) - c(x(j+1))} \mathbf{1}_{\{x(j) > x(j+1)\}}\right)}{\mathbb{E}\left(c(x(j)) - c(x(j+1))\right)}.$$

Hence r_j is a weighted average of the realized secant slopes

$$\frac{b(x(j)) - b(x(j+1))}{c(x(j)) - c(x(j+1))}$$

over those realizations for which $x(j) > x(j+1)$, with weights proportional to $c(x(j)) - c(x(j+1))$.

Now let $\bar{x}(j)$ and $\underline{x}(j)$ denote the highest and lowest policies in the support of $x(j)$. By Lemma 11, the principal is indifferent between at most two policies at each j , so these are well-defined; further, if $j' > j$ all principal-optimal policies at j' lie weakly below all principal-optimal policies at j . Thus every policy in the support of $x(j')$ is weakly below $\underline{x}(j)$. Hence

$$\bar{x}(j) \geq \underline{x}(j) \geq \bar{x}(j') \geq \underline{x}(j') \quad \text{for all } j' > j.$$

Consider any realization of the principal's strategy with $x(k) > x(k+1)$ and any realization with $x(k') > x(k'+1)$. Set

$$x_1 := x(k), \quad x_2 := x(k+1), \quad x_3 := x(k'), \quad x_4 := x(k'+1).$$

Then $x_1 > x_2$ and $x_3 > x_4$, and the support ordering above implies

$$x_1 \geq x_3, \quad x_2 \geq x_4.$$

Therefore, by the Claim 3 below,

$$\frac{b(x_1) - b(x_2)}{c(x_1) - c(x_2)} \leq \frac{b(x_3) - b(x_4)}{c(x_3) - c(x_4)}.$$

Since the two realizations were arbitrary, we see that every realized secant slope entering r_k is weakly below every realized secant slope entering $r_{k'}$. Because r_k and $r_{k'}$ are weighted averages of their respective realized secant slopes, it follows that

$$r_k \leq r_{k'}.$$

This proves the claim. □

Claim 3 (Ordered secants). *Suppose $b, c \in C^1([0, 1])$, $c'(x) > 0, b'(x) > 0$ for $x \in [0, 1]$ and $\frac{c'}{b}$ is strictly increasing. Let $x_1 > x_2$ and $x_3 > x_4$, with $x_1 \geq x_3$ and $x_2 \geq x_4$. Then*

$$\frac{b(x_1) - b(x_2)}{c(x_1) - c(x_2)} \leq \frac{b(x_3) - b(x_4)}{c(x_3) - c(x_4)}.$$

Proof. Define $h := c \circ b^{-1}$. Then

$$h'(u) = \frac{c'(b^{-1}(u))}{b'(b^{-1}(u))},$$

which is strictly increasing because $\frac{c'}{b'}$ is strictly increasing. Hence h is strictly convex.

For $u > v$, define the secant slope

$$s_h(u, v) := \frac{h(u) - h(v)}{u - v}.$$

Since h is convex, $s_h(u, v)$ is increasing in both u and v . Setting $u_i = b(x_i)$ for $i = 1, 2, 3, 4$, we have $u_1 \geq u_3$, $u_2 \geq u_4$, and

$$\frac{c(x_1) - c(x_2)}{b(x_1) - b(x_2)} = s_h(u_1, u_2) \geq s_h(u_3, u_4) = \frac{c(x_3) - c(x_4)}{b(x_3) - b(x_4)}.$$

Taking reciprocals yields the claim. \square

The next part rewrites $U(0, q_0) = 0$ as a ‘‘cost-balance’’ identity (86) and shows that, given this identity, $U(1, q_0) > 0$ is equivalent to the displayed inequality in (87).

Claim 4. *Suppose $U(0, q_0) = 0$, and evaluate δ_k at q_0 (equivalently, at $t_0 := \frac{q_0}{1-q_0}$). Then*

$$\sum_{k=\kappa+1}^{N-1} (\delta_k - 1) r(k, 0) = \mathbb{E}(c(x(N))) - c(0), \quad (86)$$

and $U(1, q_0) > 0$ is equivalent to

$$\sum_{k=\kappa+1}^{N-1} (\delta_k - 1) \mathbb{E}(b(x(k)) - b(x(k+1))) < \mathbb{E}(b(x(N))) - b(0). \quad (87)$$

Proof. Since $U(0, q) = -(1-q)^{N-1} F_0(t)$, the identity $U(0, q_0) = 0$ is equivalent to $F_0(t_0) = 0$, i.e.

$$0 = \binom{N-1}{\kappa} t_0^\kappa r(\kappa, 0) + \sum_{k=\kappa+1}^{N-1} \binom{N-1}{k} t_0^k r(k, 0).$$

Dividing by $\binom{N-1}{\kappa} t_0^\kappa > 0$ yields

$$\sum_{k=\kappa+1}^{N-1} \delta_k r(k, 0) = -r(\kappa, 0). \quad (88)$$

Next, a telescoping argument gives

$$-\sum_{k=\kappa+1}^{N-1} r(k, 0) = \mathbb{E}(c(x(N)) - c(x(\kappa+1))). \quad (89)$$

Finally, since $x(\kappa) = 0$ we have $-r(\kappa, 0) = \mathbb{E}\left(c(x(\kappa + 1))\right) - c(0)$. Combining this with (88)–(89) yields (86).

For the second statement, note that $F_1(t_0) = F_1(t_0) - F_0(t_0)$ since $F_0(t_0) = 0$. Writing out both expressions gives

$$F_1(t_0) = \binom{N-1}{\kappa} t_0^\kappa \left(r(\kappa, 1) - r(\kappa, 0) + \sum_{k=\kappa+1}^{N-1} \delta_k (r(k, 1) - r(k, 0)) \right).$$

Thus $U(1, q_0) > 0$ is equivalent to

$$r(\kappa, 1) - r(\kappa, 0) + \sum_{k=\kappa+1}^{N-1} \delta_k (r(k, 1) - r(k, 0)) > 0.$$

Rearranging,

$$\sum_{k=\kappa+1}^{N-1} \delta_k \left(r(k, 0) - r(k, 1) \right) < r(\kappa, 1) - r(\kappa, 0).$$

Substituting

$$r(k, 0) - r(k, 1) = \mathbb{E}\left(b(x(k)) - b(x(k+1))\right), \quad r(\kappa, 1) - r(\kappa, 0) = \mathbb{E}\left(b(x(\kappa+1)) - b(0)\right),$$

gives

$$\sum_{k=\kappa+1}^{N-1} \delta_k \mathbb{E}\left(b(x(k)) - b(x(k+1))\right) < \mathbb{E}\left(b(x(\kappa+1)) - b(0)\right).$$

Subtracting the telescoping identity

$$\sum_{k=\kappa+1}^{N-1} \mathbb{E}\left(b(x(k)) - b(x(k+1))\right) = \mathbb{E}\left(b(x(\kappa+1)) - b(x(N))\right)$$

from both sides yields equivalence of $U(1, q_0) > 0$ to (87). □ □

Now, we combine the single-crossing of $\delta_k - 1$ with the monotonicity of r_k to verify (87) at q_0 , and thereby conclude $U(1, q_0) > 0$.

Define

$$\hat{k} := \max\{k \in \{\kappa + 1, \dots, N - 1\} : \delta_k \geq 1 \text{ and } r(k, 0) > 0\}.$$

Note that $\sum_{k=\kappa+1}^{N-1} (\delta_k - 1) r(k, 0) = \mathbb{E}\left(c(x(N))\right) - c(0) > 0$ implies that there exists some $k > \kappa$ with $\delta_k > 1$ and $r(k, 0) > 0$, so \hat{k} is well-defined. By monotonicity of r_k ,

$$\begin{aligned} r_k &\leq r_{\hat{k}} && \text{for all } \kappa + 1 \leq k \leq \hat{k} \text{ with } r(k, 0) > 0, \\ r_k &\geq r_{\hat{k}} && \text{for all } k \geq \hat{k} \text{ with } r(k, 0) > 0. \end{aligned}$$

Using (85),

$$\mathbb{E}\left(b(x(k)) - b(x(k+1))\right) = r_k r(k, 0) \quad (\text{and both sides are 0 if } r(k, 0) = 0).$$

Thus, Claim 2 and the single-crossing sign pattern of $\delta_k - 1$ imply³⁸

$$\begin{aligned} \sum_{k=\kappa+1}^{N-1} (\delta_k - 1) \mathbb{E}\left(b(x(k)) - b(x(k+1))\right) &= \sum_{k=\kappa+1}^{N-1} (\delta_k - 1) r_k r(k, 0) \\ &\leq r_{\hat{k}} \sum_{k=\kappa+1}^{N-1} (\delta_k - 1) r(k, 0) \\ &= r_{\hat{k}} \left(\mathbb{E}\left(c(x(N))\right) - c(0)\right), \end{aligned} \quad (90)$$

where the last equality uses (86). To conclude (87), it therefore suffices to show

$$r_{\hat{k}} < \frac{\mathbb{E}\left(b(x(N))\right) - b(0)}{\mathbb{E}\left(c(x(N))\right) - c(0)}. \quad (91)$$

This is proved by the same secant slope-monotonicity argument used above to establish the monotonicity asserted in Claim 2. To establish (91), note first that $r_{\hat{k}}$ is a weighted average of the realized secant slopes

$$\frac{b(x(\hat{k})) - b(x(\hat{k} + 1))}{c(x(\hat{k})) - c(x(\hat{k} + 1))}$$

over realizations with $x(\hat{k}) > x(\hat{k} + 1)$. Hence

$$r_{\hat{k}} \leq \max \left\{ \frac{b(x(\hat{k})) - b(x(\hat{k} + 1))}{c(x(\hat{k})) - c(x(\hat{k} + 1))} : x(\hat{k}) > x(\hat{k} + 1) \right\}.$$

Likewise,

$$\frac{\mathbb{E}\left(b(x(N))\right) - b(0)}{\mathbb{E}\left(c(x(N))\right) - c(0)}$$

is a weighted average of the secant slopes

$$\frac{b(z) - b(0)}{c(z) - c(0)}$$

³⁸The inequality (90) is the standard “single-crossing weights + monotone ratios” comparison. It can be viewed as a rearrangement inequality, compare to Hardy, Littlewood and Pólya (1952, *Inequalities*, Ch. 10).

over z in the support of $x(N)$ with $z > 0$, and therefore

$$\frac{\mathbb{E}\left(b(x(N))\right) - b(0)}{\mathbb{E}\left(c(x(N))\right) - c(0)} \geq \min \left\{ \frac{b(z) - b(0)}{c(z) - c(0)} : z \in \text{supp}\left(x(N)\right), z > 0 \right\}.$$

Now fix any realization with $x(\hat{k}) > x(\hat{k} + 1)$ and any $z \in \text{supp}\left(x(N)\right)$ with $z > 0$. By monotonicity of the principal's best reply,

$$x(\hat{k}) \geq x(\hat{k} + 1) \geq z > 0.$$

Applying Claim 3 with

$$(x_1, x_2) = \left(x(\hat{k}), x(\hat{k} + 1)\right), \quad (x_3, x_4) = (z, 0),$$

yields

$$\frac{b(x(\hat{k})) - b(x(\hat{k} + 1))}{c(x(\hat{k})) - c(x(\hat{k} + 1))} < \frac{b(z) - b(0)}{c(z) - c(0)},$$

where the inequality is strict because $x(\hat{k} + 1) > 0$ (so $x_2 > x_4$). This holds for every such pair and we obtain

$$\max \left\{ \frac{b(x(\hat{k})) - b(x(\hat{k} + 1))}{c(x(\hat{k})) - c(x(\hat{k} + 1))} : x(\hat{k}) > x(\hat{k} + 1) \right\} < \min \left\{ \frac{b(z) - b(0)}{c(z) - c(0)} : z \in \text{supp}(x(N)), z > 0 \right\}.$$

Combining with the previous displayed inequalities yields (91).

J.4 Remainder of Theorem 3's Proof

The remainder of Theorem 3's proof is almost verbatim to that in the context of the baseline model, with only smaller notational changes (mostly, replacing the argument for the single cutoff \bar{k} with analogous ones for the multiple cutoffs $\bar{k}_{j,j-1}$). Precisely, mimicking Proposition 3's proof establishes information aggregation, and mimicking Theorem 2's proof establishes the lower bound on the payoff guarantee. The details of the remainder are thus omitted.

K Heterogeneous Ex-Post Preferences

We consider a variation of the baseline model from Section 1 in which the voters have a private preference type, and payoffs can depend on the state in a general way. Except for this additional type dimension, the baseline model is unchanged.

Formally, each agent i is a (non-strategic) partisan for $a \in \{0, 1\}$ and chooses $a_i = a$, with probability $0 < \rho_a < \frac{1}{2}$, as before. A non-partisan agent has a private type, which is a prior $p_i \in [0, 1]$ and a pair $\mathbf{t}_i = (t_i(0), t_i(1)) \in [0, 1]^2$, describing the type's constant marginal benefit from the policy choice in the two states. Types t_i are drawn from a distribution G and independently from priors, signals, and the state, and across voters.

A type $(p_i, t_i(0), t_i(1))$'s payoff from x in ω is

$$x(t_i(\omega) - c),$$

Given a strategy profile η , a type prefers the action 1 if and only if

$$\begin{aligned} & p_i(t_i(1) - c) \frac{\Pr(s_i = s | \omega = 1)}{\Pr_i(s_i = s)} U(1; \eta) \\ & - (1 - p_i)(c - t_i(0)) \frac{\Pr(s_i = s | \omega = 0)}{\Pr_i(s_i = s)} U(0; \eta) \geq 0. \end{aligned} \quad (92)$$

We assume that the mass of types for which $p_i(t_i(1) - c) = 0$ and $(1 - p_i)(c - t_i(0)) = 0$ is zero and ignore these types in the following, without loss of generality. Finally, we generalize the notion of an agents' information structure π to mean the pair of a signal and a type distribution.

K.1 The Aggregate Preference Function Φ

We show that the equilibrium set in this generalized setting depends on the type distribution only through the *aggregate preference function*

$$\begin{aligned} & \Phi(U(0; \eta), U(1; \eta), l) \\ & = \rho_1 + (1 - \rho_1 - \rho_0) \Pr\left(\{(p_i, \mathbf{t}_i) : p_i(t_i(1) - c)U(1; \eta) \geq (1 - p_i)(c - t_i(0)) \cdot l \cdot U(0; \eta)\}\right) \end{aligned}$$

for $l := \frac{\Pr(s_i = s | \omega = 0)}{\Pr(s_i = s | \omega = 1)}$, via two observations:

First, equilibria are equivalently characterized by a principal's strategy (\bar{k}, \tilde{x}) and a mean action pair $\mathbf{q} = (q(0), q(1))$ so that (\bar{k}, \tilde{x}) and \mathbf{q} are best replies to (\bar{k}, \tilde{x}) and \mathbf{q} . To make sense of this, note that, for any strategy profile $\eta = (\sigma, (\bar{k}, \tilde{x}))$, the mean action pair $\mathbf{q}(\sigma) = (q(0; \sigma), q(1; \sigma))$ pins down the set of principal's best replies; \mathbf{q} and (\bar{k}, \tilde{x}) together pin down the average effects $U(\omega'; \eta)$; ³⁹ and the average effects are a sufficient statistic for the agents' best reply, given (92). In conclusion, $\mathbf{q}(\sigma)$ and (\bar{k}, \tilde{x}) are a sufficient statistic for the best reply correspondence, which yields the claimed equilibrium characterization.

Second, multiplying (92) by $\frac{\Pr(s_i = s)}{\Pr(s_i = s | \omega = 1)}$ shows the best reply correspondence's mean action pairs depend on the type distribution only through Φ . Consequently, the same is true for the equilibrium set, as claimed.

Before we proceed to deriving analogs of the main results for the heterogeneous-preference setting, we recall from the main text that the value of Φ at $(U(0; \eta), U(1; \eta), l)$ only depends on

$$z_1 := \frac{U(0; \eta)}{U(1; \eta)} \cdot l.$$

whenever $U(1; \eta) \neq 0$, and that an agents' information structure has *monotone preferences* if Φ is continuously differentiable in z_1 and $\partial\Phi/\partial z_1$ has the same non-zero sign for all

³⁹Cf. (1) and (3).

$z_1 \in (0, \infty)$.

K.2 Analog of Proposition 3

An analog of Proposition 3's statement about information aggregation holds in the heterogeneous-preference setting.

Proposition 3'. *Consider any monotone quasi-referendum with a single cutoff and any agents' information structure with monotone preferences. Information aggregates in all equilibrium sequences if the quasi-referendum has no balance and $\max Q(0) < \max Q(1)$.*

The proof of Proposition 3' closely follows that of Proposition 3. We therefore do not repeat the full argument, but only describe the necessary modifications.

In the baseline proof, the first step establishes Lemma 1. This lemma continues to hold in the heterogeneous-preference setting, with an identical proof.

The baseline argument then shows that, for each signal $s \in \{0, 1\}$, Lemma 1 implies the existence of a unique indifferent type $p_N(s) \in (0, 1)$. The proof proceeds by distinguishing two complementary cases:

$$0 < \lim_{N \rightarrow \infty} p_N(1) < \lim_{N \rightarrow \infty} p_N(0) < 1, \quad (93)$$

or

$$\lim_{N \rightarrow \infty} p_N(1) = \lim_{N \rightarrow \infty} p_N(0) \in \{0, 1\}, \quad (94)$$

and shows that each case leads either to a contradiction or to information aggregation.

In the heterogeneous-preference setting, equilibrium strategies are no longer characterized by two indifferent types. Instead, the analogous dichotomy is expressed in terms of the ratio of average effects. We distinguish the cases

$$\lim_{N \rightarrow \infty} \frac{U(0; \eta_N)}{U(1; \eta_N)} \in (0, \infty), \quad (95)$$

and

$$\lim_{N \rightarrow \infty} \frac{U(0; \eta_N)}{U(1; \eta_N)} \in \{0, \infty\}. \quad (96)$$

In the first case, monotonicity of preferences implies that the mean actions differ across states, i.e. $q(0; \sigma) \neq q(1; \sigma)$ (cf. (14)). Since the realized collective action concentrates around the mean in each state, the principal learns the state from the observed outcome, and information aggregates.

In the second case, note that in the baseline model (96) is equivalent to (94) via (9). Hence, the baseline proof can be adapted almost verbatim: (96) either yields a contradiction (generic Case 1) or implies information aggregation (non-generic Case 2). The required modifications are as follows.

1. In the baseline proof, (72) follows from the fact that the indifferent types are bounded away from 0 and 1. In the present setting, the same conclusion follows directly from (96): if (72) failed, then the ratio $\frac{U(0; \eta_N)}{U(1; \eta_N)}$ would remain bounded away from 0 and ∞ , contradicting (96).

2. In Case 2 of the baseline proof, subcases are distinguished according to:

- (a) whether $\lim_{N \rightarrow \infty} \Pr(\omega = 1 \mid \text{piv}_1; \sigma_N, N)$ equals 0 or 1, and
- (b) whether types choose action 1 for $p_i \leq p_N(s)$ or for $p_i \geq p_N(s)$.

In the baseline model, (b) is equivalent to the sign of $U(1; \eta_N)$: if $U(1; \eta_N) < 0$, a type chooses 1 iff $p_i \leq p_N(s)$, and if $U(1; \eta_N) > 0$, a type chooses 1 iff $p_i \geq p_N(s)$. Since the object $U(1; \eta_N)$ is also well-defined in the heterogeneous-preference setting and given this equivalence, the same case distinction and proof can be reused.

3. The first paragraph of Case 2 is replaced by:

“This case can be decomposed into several analogous subcases; we present one. Consider an equilibrium sequence such that $\lim_{N \rightarrow \infty} \Pr(\omega = 1 \mid \text{piv}_1; \sigma_N, N) = 1$ and $U(1; \eta_N) < 0$ for all N . Monotonicity of preferences then implies $\rho_1 < q(1; \sigma_N) < q(0; \sigma_N) < 1 - \rho_0$. By (96), either $\frac{U(0; \eta_N)}{U(1; \eta_N)} \rightarrow \infty$ or $\rightarrow 0$. If the ratio diverges to ∞ , then $q^ = 1 - \rho_0$, which contradicts $\lim_{N \rightarrow \infty} \Pr(\omega = 1 \mid \text{piv}_1; \sigma_N, N) = 1$. Hence the ratio converges to 0, implying $q^* = \rho_1$.”*

4. In the subsequent analysis, the expression $F(p_N(s))$ is replaced by $\Phi(z_1(s), z_2)$, where

$$z_1(s) = \frac{U(0; \eta_N) \Pr(s_i = s \mid \omega = 0)}{U(1; \eta_N) \Pr(s_i = s \mid \omega = 1)}, \quad z_2 = \text{sign}(U(1; \eta_N)).$$

K.3 Analogs of Theorem 1 and 2

The analogs of Theorem 1 and 2 hold, provided monotone preferences and the condition that a majority of types ranks policy 0 highest in state 0 and policy 1 highest in state 1. Formally,

$$\rho_1 + (1 - \rho_1 - \rho_0) \Pr(\{\mathbf{t}_i : t_i(1) - c > 0\}) > \frac{1}{2}, \quad \text{and} \quad (97)$$

$$\rho_0 + (1 - \rho_1 - \rho_0) \Pr(\{\mathbf{t}_i : c - t_i(0) > 0\}) > \frac{1}{2}. \quad (98)$$

Theorem 1’. *Any collective veto of the maximal policy has a payoff guarantee of $1 - \varepsilon$ across all agents’ information structures with monotone preferences and satisfying (97) and all equilibrium sequences. This maximizes the payoff guarantee across all quasi-referenda.*

Theorem 2’. *The gateway referendum with the simple majority cutoff $m_1 = \frac{1}{2}$ has a payoff guarantee larger than $1 - \varepsilon$ across all agents’ information structures with monotone preferences and satisfying (97) and all equilibrium sequences. This maximizes the payoff guarantee across all generalized quasi-referenda.*

The proofs of Theorem 1’ and Theorem 2’ are almost the same as those of Theorem 1 and Theorem 2, whose proofs relied on Proposition 1 and Proposition 3.

Proposition 1 showed the existence of an information structure and a corresponding inefficient equilibrium sequence. It remains true since we only expanded the set of feasible information structures by allowing for heterogeneous ex-post preferences. We already proved an appropriate analog of Proposition 3 in Section K.2.

With these, the same proofs as before can be mimicked. In the proof of Theorem 2', one has to take care to replace and invoke (97) instead of its baseline model version condition $\rho_1 < \frac{1}{2} < 1 - \rho_0$ whenever the latter one was needed in an argument.

K.4 Agent-Optimal Quasi-referenda

We provide sufficient conditions for the quasi-referenda defined by (12) and also the generalized quasi-referenda defined by (13) to be also agent-optimal in their respective settings. That is, they maximize the agents' payoff guarantee. By this we mean the percentage of the full-information payoff achieved in the worst-case scenario and as $N \rightarrow \infty$, by a social planner who has full information about the state and maximizes the agent's ex-ante welfare

$$\int_{p_i=0}^1 p_i E_G(t_i(1) - c) + (1 - p_i) E_G(t_i(0) - c) dF(p_i).$$

The first condition is

$$0 \leq E_G(t_i(0)) < c, \text{ and } 1 \geq E_G(t_i(1)) > c$$

and means that, when the state is known, policies are ranked in the same way whether considering the principal's or the agents' ex-ante welfare (lower policies are strictly preferred in state 0 and higher ones in state 1). This means the agents' full information payoff is

$$E_F(p_i) \left(E_G(t_i(1)) - c \right)$$

and the agents' payoff guarantee is

$$\hat{G}(Q) := \inf_{(\eta_N)_{N \in \mathbb{N}, \pi}} \left(\liminf_{N \rightarrow \infty} E(x \mid \omega = 1; \eta_N) - \frac{E_F(1 - p_i) \left(c - E_G(t_i(0)) \right)}{E_F(p_i) \left(E_G(t_i(1)) - c \right)} E(x \mid \omega = 0; \eta_N) \right)$$

It differs from the principal's payoff guarantee simply by replacing $\Pr(\omega = 0)$ with $E_F(1 - p_i) \left(c - E_G(t_i(0)) \right)$ and $\Pr(\omega = 1)$ with $E_F(p_i) \left(E_G(t_i(1)) - c \right)$.

The second condition is

$$\frac{1 - E_F(p_i)}{E_F(p_i)} \cdot \frac{c - E_G(t_i(0))}{E_G(t_i(1)) - c} \geq \varepsilon. \quad (99)$$

and the relevant implication is that choosing $x = 1$ in both states yields an agent's ex-ante payoff smaller than $1 - \varepsilon$ times the full-information payoff.⁴⁰

Given the two conditions, mimicking Section 4.1's proof of Theorem 1 shows that the quasi-referenda (12) maximize the agents' payoff guarantee across all quasi-referenda.

⁴⁰The payoff from choosing $x = 1$ in both states divided by the full-information payoff is $1 - \frac{E_F(1 - p_i) \left(c - E_G(t_i(0)) \right)}{E_F(p_i) \left(E_G(t_i(1)) - c \right)}$, which is smaller than $1 - \varepsilon$ if (99) holds.

Similarly, mimicking the proof of Theorem 2 in Appendix G shows that the quasi-referenda (13) maximize the agents' payoff guarantee across all generalized quasi-referenda.

L Existence of Efficient Equilibria for Quasi-referenda with a Single Cutoff

Many quasi-referenda with a single cutoff have ex-post efficient equilibrium sequences for *any* agents' information structure for which the mass of one side of the partisans does not trivially exceed the cutoff, i.e. $\rho_1 < m_1 < 1 - \rho_0$. We establish a sufficient condition for this property, which is that the minimum and maximum policies are increasing, i.e., $\min Q(0) < \min Q(1)$ and $\max Q(0) < \max Q(1)$, and that the ex-post optimal policies are not excluded, i.e., $\min Q(0) = 0$ and $\max Q(1) = 1$.

Proposition 7. *Take any quasi-referendum with a single cutoff m_1 . If $\min Q(0) < \min Q(1)$, $\max Q(0) < \max Q(1)$, $\min Q(0) = 0$, and $\max Q(1) = 1$, given any agents' information structure with $\rho_1 < m_1 < 1 - \rho_0$, there exists an equilibrium sequence $(\eta_N)_{N \in \mathbb{N}}$ for which*

$$\lim_{N \rightarrow \infty} \Pr(x = \omega | \eta_N, N) = 1.$$

Proof. The equilibrium strategies are found among a parametric set of candidate strategies σ_L where an agent with signal $s \in \{0, 1\}$ chooses $a_i = 1$ if and only if $p_i \geq p_L(s)$ with $p_L(s)$ solving

$$L = \frac{\Pr_i(\omega = 1 \mid p_i = p_L(s), s_i = s)}{\Pr_i(\omega = 0 \mid p_i = p_L(s), s_i = s)}.$$

Note that $p_L(s)$ is increasing in L . Hence, the mean action in each state, $q(\omega'; \sigma_L)$ is strictly decreasing in L . We consider a compact set of parameters $L \in [\underline{L}_N, \bar{L}_N]$, with the parameter bounds implicitly given by the equations

$$q(0; \sigma_{\underline{L}_N}) = \frac{\lfloor m_1 N \rfloor}{N}, \text{ and} \tag{100}$$

$$q(1; \sigma_{\bar{L}_N}) = \frac{\lfloor m_1 N \rfloor}{N}. \tag{101}$$

A preliminary result prepares the formal fixed-point argument that constructs the equilibrium sequence.

Claim 5. *Take any sequence $(L_N)_{N \in \mathbb{N}}$ with $L_N \in [\underline{L}_N, \bar{L}_N]$ for all $N \in \mathbb{N}$. The sequence of the cutoffs $(\bar{k}_N)_{N \in \mathbb{N}}$ of the principal's best response to $(\sigma_{L_N})_{N \in \mathbb{N}}$ satisfies $\lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(0; \sigma_{L_N})\right) = \lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(1; \sigma_{L_N})\right) > 0$.*

Proof. The bound $L_N \in [\underline{L}_N, \bar{L}_N]$ implies $0 < \lim_{N \rightarrow \infty} q(0; \sigma_{L_N}) < \lim_{N \rightarrow \infty} q(1; \sigma_{L_N}) < 1$. An application of the law of large numbers thus shows that information aggregates. In particular, this rules out the boundary case $\bar{k}_N = N$ because the principal's posterior crosses her indifference belief $\frac{1}{2}$ as a function of k .

Now, we suppose that

$$\lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(0; \sigma_{L_N})\right) \neq \lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(1; \sigma_{L_N})\right) \quad (102)$$

and derive a contradiction. Given (102), an application of (17) yields $\lim_{N \rightarrow \infty} \Pr(\omega = 1 | \bar{k}_N + 1; \sigma_{L_N}, N) \in \{0, 1\}$, but this contradicts the minimality of $\bar{k}_N + 1$.

Finally, $\lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(0; \sigma_{L_N})\right) = \lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(1; \sigma_{L_N})\right)$ and $\lim_{N \rightarrow \infty} q(0; \sigma_{L_N}) < q(1; \sigma_{L_N})$ together imply $\lim_{N \rightarrow \infty} \text{KL}\left(\frac{\bar{k}_N}{N}, q(\omega'; \sigma_{L_N})\right) > 0$ for $\omega' \in \{0, 1\}$. \square

We describe the fixed-point correspondence f . It is a mapping on the parameter space $[\underline{L}_N, \bar{L}_N]$: Take any $L \in [\underline{L}_N, \bar{L}_N]$, any principal's best response (\bar{k}, \tilde{x}) to σ_L , and the strategy profile $\eta = (\sigma_L, (\bar{k}, \tilde{x}))$. We claim that $U(\omega'; \eta) > 0$ for any ω' and large enough N , and prove this momentarily. Given (9), the agents' best response to η is then the strategy $\sigma_{L'_N}$ with $L'_N = \frac{U(0; \eta)}{U(1; \eta)} \in (0, \infty)$. We consider the correspondence f that maps L to the set consisting of the projections $\min\left(\max(L_N, L'_N), \bar{L}_N\right)$ of all such best responses L'_N .

The proof of the claim $U(\omega'; \eta) > 0$ relies on our assumptions for the quasi-referendum. Note that Claim 5 implies that $0 < \lim_{N \rightarrow \infty} q(0; \sigma_{L_N}) < \lim_{N \rightarrow \infty} \frac{\bar{k}_N}{N} < \lim_{N \rightarrow \infty} q(1; \sigma_{L_N}) < 1$. This implies, (a) all collective actions $k \in \{0, \dots, N\}$ are on path, (b) the principal's posterior $\Pr(\omega = 1 | k; \eta_N, N)$ is weakly increasing, when N is large enough. The assumptions $\min Q(0) < \min Q(1)$ and $\max Q(0) < \max Q(1)$ then imply that (c) under any principal's best reply, the choice $x(k)$ is weakly increasing in k , and (d) $x(\bar{k}) < x(\bar{k} + 2)$. Together (a), (c), and (d) imply the claim.

Now we provide a fixed-point argument that constructs efficient equilibrium sequences. Given the defining equation (100) for \underline{L}_N , for any N and $L_N = \underline{L}_N$,

$$\text{KL}\left(\frac{\lfloor m_1 N \rfloor}{N}, q(0; \sigma_{L_N})\right) = 0 < \text{KL}\left(\frac{\lfloor m_1 N \rfloor}{N}, q(1; \sigma_{L_N})\right).$$

Claim 5 together with the expressions (21) and (22) for the pivotal likelihoods implies $\lim_{N \rightarrow \infty} \Pr\left(\omega = 1 | \text{piv}; \sigma_L, (\bar{k}_N, \tilde{x}_N), N\right) = 0$ for any sequence of principal's best responses (\bar{k}_N, \tilde{x}_N) to σ_L . Consequently,

$$L'_N > \underline{L}_N \text{ for all } L'_N \in f(\underline{L}_N) \quad (103)$$

and N large enough. Conversely, for $L = \bar{L}_N$,

$$\text{KL}\left(\frac{\lfloor m_1 N \rfloor}{N}, q(1; \sigma_L)\right) = 0 < \text{KL}\left(\frac{\lfloor m_1 N \rfloor}{N}, q(0; \sigma_L)\right).$$

Claim 5 together with (21) and (22) implies $\lim_{N \rightarrow \infty} \Pr\left(\omega = 1 | \text{piv}; \sigma_L, (\bar{k}_N, \tilde{x}_N), N\right) = 1$ for any sequence of principal's best replies (\bar{k}_N, \tilde{x}_N) to σ_L . Consequently,

$$L'_N < \bar{L}_N \text{ for all } L'_N \in f(\bar{L}_N) \quad (104)$$

and N large enough. Finally, an application of Kakutani's fixed point theorem yields a sequence of fixed points $\underline{L}_N < L_N^* < \bar{L}_N$ for which $L_N^* \in f(L_N^*)$.

Finally, we argue that any sequence of fixed points corresponds to an equilibrium sequence that achieves the full-information payoffs. First, since

$$\underline{L}_N < L_N^* < \bar{L}_N, \quad (105)$$

any fixed point L_N^* is interior when N is sufficiently large, so the corresponding sequence $\sigma_{L_N^*}$ is a sequence of equilibrium strategies.

Second, the ordering (105) also holds in the limit as $N \rightarrow \infty$,

$$\lim_{N \rightarrow \infty} \underline{L}_N < \lim_{N \rightarrow \infty} L_N^* < \lim_{N \rightarrow \infty} \bar{L}_N. \quad (106)$$

Suppose, for example that $\lim_{N \rightarrow \infty} \underline{L}_N = \lim_{N \rightarrow \infty} L_N^*$. Then, since the Kullback-Leibler divergence is continuous, the same argument leading to (103) would imply $\lim_{N \rightarrow \infty} \underline{L}_N < \lim_{N \rightarrow \infty} L_N^*$, a contradiction.

Given (100) and (101), this limit ordering implies

$$\lim_{N \rightarrow \infty} q(0; \sigma_{L_N^*}) < m_1 < \lim_{N \rightarrow \infty} q(1; \sigma_{L_N^*}). \quad (107)$$

for all N large enough. Thus, information aggregates and the principal's choice is $x = \min Q(0) = 0$ in state 0 and $x = \max Q(1) = 1$ in state 1, with probability converging to 1. We conclude that the equilibrium sequence achieves the efficient full-information payoffs. \square

M Interior Mean Actions Imply Trembling-hand Perfection

Given the partisans, any equilibrium $\eta = (\sigma, (\bar{k}, \hat{x}))$ has interior mean actions, i.e. $q(\omega'; \sigma) \in (0, 1)$ for $\omega' \in \{0, 1\}$. We show this implies trembling-hand perfection (Selten, 1988). By definition, η is trembling-hand perfect if there exists a sequence of completely mixed strategy profiles $(\eta_k)_{k \in \mathbb{N}}$ with agent strategies $(\sigma_k)_{k \in \mathbb{N}}$ and the following properties:

- (i) $(\eta_k)_{k \in \mathbb{N}}$ converges to η ,
- (ii) every player's strategy in η is a best reply to η_k for all k .

It is easy to verify that interior mean actions imply that there is a completely mixed sequence $(\eta_k)_{k \in \mathbb{N}}$ for which (i) and

- (iii) $q(\omega'; \sigma_k) = q(\omega'; \sigma)$ and $U(\omega'; \eta_k) = U(\omega'; \eta)$ for all $\omega' \in \{0, 1\}$,

hold. Since the vector $\left(q(\omega'; \sigma), U(\omega'; \eta) \right)_{\omega' \in \{0, 1\}}$ is a sufficient statistic for the best-response correspondence if the mean actions in each state are interior, given (1) and (2), (ii) also holds. To conclude, η is perfect.