Social Choice under Gradual Learning*

Yiman Sun† Caroline Thomas‡ Takuro Yamashita§

2 June 2023

Abstract

This paper combines dynamic mechanism design with collective experimentation. Agents are heterogeneous in that some stand to benefit from a proposed policy reform, while others are better off under the status quo policy. Each agent's private information regarding her preference type accrues only gradually, over time. A principal seeks a mechanism that maximises the agents' joint welfare, while providing incentives for the agents to truthfully report their gradually acquired, private information. The first-best policy may not be incentive compatible, as uninformed agents may have an incentive to prematurely vote for a policy instead of waiting for their private signal. Under the second-best policy, the principal can incentivise truth-telling by setting a deadline for experimentation, delaying the implementation of the policy reform, and keeping agents in the dark regarding others' reports.

JEL Classification Number: *C72, C73, D82, D83.*

Keywords: Dynamic mechanism design, social choice, voting, decentralised learning

---

*Thanks to Willie Fuchs, Meg Meyer, Max Stinchcombe, Jean Tirole, Tom Wiseman, Juuso Välimäki, and seminar audiences at Austin, Indiana University, Northwestern, Notre Dame, NYUAD, University of Pennsylvania, Pennsylvania State University, Seminar in Economic Theory (online), Virtual Seminar in Economic Theory (online), and Toulouse for helpful comments. Yiman Sun and Takuro Yamashita gratefully acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) and from the European Research Council (Starting Grant #714693).

†CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, yiman.sun@cerge-ei.cz

‡Department of Economics, University of Texas at Austin and CEPR, caroline.thomas@austin.utexas.edu

§Osaka School of International Public Policy, Osaka University, yamashitat@osipp.osaka-u.ac.jp

1 Introduction

A policy reform is proposed and an election will determine the collective's joint decision to either adopt the reform or to maintain the status quo. To illustrate, think of the 2016 referendum on the withdrawal of the United Kingdom from the European Union (Brexit). Some UK citizens stand to benefit from the reform. Others would be better off under the status quo. The implications of the policy reform can be complex. Thus, each citizen privately learns her preference position only gradually, over time. The government is a utilitarian social planner seeking optimally to translate the agents' learning into a social choice. Can it elicit the citizens' private information? What is the constrained optimal voting system?

The interaction takes place in continuous time. Agents' preference types—whether they are better off under the policy reform or the status quo—are independently and identically drawn by nature at the outset of the game, and remain fixed thereafter. An agent's private learning takes the following form: at an exponentially distributed random time, she receives a private signal that conclusively reveals her preference type. The arrival times of signals are independently and identically distributed across the electorate. Thus, until she receives her private signal, the agent perceives no information regarding her preference type, and her posterior equals her prior. We assume that learning is exogenous and that there is no moral hazard.

At any point in time, the planner can abandon the status quo policy and implement the reform, according to some decision rule. We assume that this policy switch is irreversible. At each instant, agents cast a vote in favour of either policy, or they abstain. This is equivalent to a direct mechanism in which, at each instant, the principal solicits reports by the agents regarding their private information. Truthful reporting consists of abstaining until the arrival of one's private signal, followed by immediate, truthful reporting of the realised signal. We distinguish two cases according to whether the history of votes is publicly observed or only observed privately by the planner.

If all signals regarding preferences arrived publicly, the planner could implement a decision rule that maximises joint payoffs. We show that this first-best decision rule cannot always be implemented when signals arrive privately and the planner relies on agents' reports. Typically, the binding constraint will come from uninformed agents who can benefit from falsely reporting a signal at voting histories where the planner would like these agents to keep waiting for their signal. There are two reasons why the planner's and agents' incentives to wait for the agents' signals are misaligned. First, the planner also cares about the payoff of agents who have already cast their vote. Second, uninformed agents exert a negative externality on one another, as each agent can end up trapped in an undesired policy by another agent's report. In summary, our setup is characterised by gradual, costless learning, private signals, and misaligned incentives due to collective decision-making.

We begin by analysing the scenario where agents make reports about their private signals in continuous time. We derive the first-best policy, corresponding to the benchmark in which all agents' signals are publicly observed upon arrival. The first-best policy is a Markov policy with respect to the state variable consisting of the number of agents who have discovered that they are better off under the status quo policy and the number of agents who have discovered that they are better off under the policy reform. (All remaining agents are uninformed.) Given a state, the first-best policy is a threshold policy with respect to the belief (on being the type who favours the status quo) of the uninformed agents who have not yet learnt their preferences. The fist-best policy implements the reform once the number of supporters is sufficiently large compared to the number of status-quo supporters and the number of uninformed agents.

When learning is private, so that the planner relies on the agents' reports to implement the first-best policy, truthful reporting may not be incentive compatible, when . In particular, an uninformed agents may want to misreport that she has received a signal. To illustrate, consider a two-agent example. Suppose that one agent has already voted for the status quo and that the benefit of the policy reform is relatively high, so that the remaining uninformed agent is pivotal under the first-best policy: as long as she abstains, the planner waits; once she reports a preference type, the planner implements her preferred policy forever. Moreover, suppose the belief of the uninformed agents is relatively low (below the single-agent experimentation threshold) so that waiting for her signal is not worth her while and she prefers immediately switching to the reform. The planner would agree with her if the other agent were also uninformed. However, because the other agent has already made a report to support the status quo policy, the planner is biased in favour of the status quo and wants to learn the last uninformed agent's type before making a final decision. Thus, the planner and the uninformed agent disagree on whether to continue the status quo policy. If the first-best policy is implemented, the uninformed agent will simply misreport her type and vote in favor of the reform.[1]

The above example illustrates that a single pivotal uninformed agent may have an incentive to misreport. When there are multiple pivotal uninformed agents, one agent's report imposes a negative externality on the other agents' learning, and produces an additional incentive to prematurely report a preference type.

Given that the first-best policy is not incentive compatible, how can the planner induce truth-telling? As monetary transfers are not feasible, it seems that there is not much the planner can do. We consider two examples of our model, where we completely characterise the second

[1]We find that the other direction is also possible. That is, given the first-best policy, when the belief of the uninformed agents is sufficiently high (above the single agent myopic threshold), they may prematurely vote for the status quo.

best policy, i.e. the policy maximising joint payoffs subject to truthful reporting being incentive compatible.

The first example (described above) consider the case with two agents, one of whom has already learned that she is in favor of the status quo policy. Thus, the problem consists in incentivising one pivotal uninformed agent. We assume reports are made publicly so that all players have common knowledge on the history of reports. We find that the second-best mechanism distorts the first-best using deadlines and delays. A deadline means that, at a voting history where it would be welfare-maximising to continue experimenting with the status-quo policy, the planner commits to switch to the reform at a pre-determined, fixed date. A delay means that, at a voting history where it would be welfare-maximising immediately to implement the reform, the planner remains in the status quo policy and switches to the reform after a (possibly stochastic) delay. As the uninformed agent has an incentive to induce the reform, a deadline reduces her cost of waiting in the status quo policy, which improves her truth-telling incentive. Moreover, delays also reduce her misreporting incentive as she cannot induce the reform as fast as she would want. Those two tools can be used together to make the uninformed agent indifferent between truth-telling and misreporting at any time before the deadline. This requires delays to depend on the reporting time. In particular, delays are monotonic in the reporting time. That is, if an agent reports that her preference is in favor of the reform at an earlier time, then the time at which the policy switches to the reform is also earlier. Moreover, the policy switching date after a delay never exceeds the deadline. Those properties guarantee that an agent who has truly learned her preference has the proper incentive to report truthfully and immediately. In this example, there is no scope for improving welfare using private reporting (where reports are only observed by the planner), as the fact that the planner has not yet implemented the reform allows the uninformed agent to infer that she is, in fact, pivotal. Thus, public and private reporting coincide.

In general, private reporting provides minimal information to agents and thus minimises opportunities for them to misreport a signal. The planner cannot do worse with private reporting. We consider private communication in our second example, where there are two uninformed agents and the first-best policy is to switch to the reform as soon as one agent supports the reform. Again, as the prior belief is low, uninformed agents want to induce the reform as soon as possible, which conflicts with the planner's first-best policy. We can still use the same tools—deadlines and delays—to induce truth-telling when no one has learned any signal. However, this is not enough to restore incentive compatibility. When a first agent votes for the reform before the deadline, this triggers the delay. However, there is still another uninformed agent in the game, and she, too, wants to induce the reform as soon as possible. What should the planner do to restore the second agent's truth-telling incentive? Although it is efficient to listen to her report to make the final decision, listening to her report also provides an opportunity for her to manipulate the policy. We show that the second-best policy has to distort efficiency at three voting histories. While both agents abstain, a deadline is set for the status quo policy. Before that deadline, when one agent supports the reform, a delay is set to switch to the reform. The switching time is the same as the deadline, regardless when the first agent votes for the reform. Meanwhile, during the delay, the planner still listens to the second agent's reports. If the second agent votes for the status quo before the end of the delay, the status quo policy will be implemented forever. If the second agent votes for the reform, then another delay will be imposed. Moreover, this mechanism delivers higher joint payoffs when agents are kept in the dark about each other's report so that they are only indifferent in expectation between truth-telling and misreporting. Private communication here improves efficiency, since otherwise under public communication the efficiency has to be distorted further away from the first-best policy.

Last, we also study the single election problem, where the planner has only one chance to collect information from the agents. Given the election date, the first-best policy is static—it maximizes the myopic social welfare by assigning different weights to different types of agents according to their contribution to social welfare. Since an uninformed agent's preference is not as strong as an informed agent, the weight of the uninformed agent is small. However, unless the prior belief equals the single-agent myopic threshold, her preference toward the two policies is strict. Thus, under the first-best policy, the uninformed agents will exaggerate their preference by misreporting that they are informed. In particular, supposing the prior belief is low,[2] they will misreport that they are in favor of the reform. Thus, the first-best policy is not incentive compatible. To restore incentive compatibility, it is necessary to give the uninformed agents and the agents who support the reform the same option in expectation, in the interim stage. We show that the second-best policy in fact does not distinguish these two types ex-post. In other words, preference intensities are not taken into account. Essentially, the policy is to remain in the status quo if there is enough support. But how much support is enough is determined by the election time. We show that more support is required if the election is held at a later date, since as time passes, an agent who is against the status quo is more likely to be an agent who has discovered her preference in favor of the reform, and thus should be assigned more weight. We also compare the optimal election dates between the first-best and the second-best policies, and show that the relative importance between pre-time-zero-learning and post-time-zero-learning determines whether the election date in the second-best policy should be held earlier.

Literature Our paper relates to the voting literature. One strand of the literature, starting from impossibility results of Gibbard (1973) and Satterthwaite (1975), examines preference aggregation in a static setting and is concerned with the strategy-proofness of a social choice function.

[2]The case with a high prior belief is similar.

While most papers require dominant strategy implementation, a small literature advocates a less demanding requirement—ordinally Bayesian incentive compatibility (Majumdar and Sen (2004)) where the social choice function is restricted to be ordinal, or Bayesian incentive compatibility (Azrieli and Kim (2014); Kim (2017); Ehlers et al. (2020)) where the social choice function can take cardinal information (i.e., preference intensities) into consideration. A cardinal rule may achieve higher utilitarian social welfare than any ordinal rule with three alternatives (Kim (2017)). However, with only two alternatives, the incentive compatible social choice function that maximizes the social welfare is a weighted majority rule and only takes ordinal information into account (Azrieli and Kim (2014)). This is in line with our findings in the single-election scenario, where a voter's problem is essentially static. Given our dynamic environment, we also show how the assignment of voting weights depends on the election date through learning and the implications on the optimal election date.

There is also a literature on voting with costly participation. As voting is costly, the voter turnout rate may be a concern and compulsory voting appears to be a remedy. However, the literature identifies, under majority rule, different reasons to support voluntary voting (Börgers (2004)), or compulsory voting (Ghosal and Lockwood (2009)), or even subsidies and fines to encourage participation (Krasa and Polborn (2009)). Krishna and Morgan (2015) and Chakravarty et al. (2018) further find that costly voting benefits utilitarian social welfare under majority rule, as the cost can deter voters who have low preference intensities to vote so that only those high-preference-intensities voters have an incentive to vote and they are the ones who contribute to the social welfare the most. Voting costs would also help our incentive problem, since a direct truthful mechanism requires the uniformed agents to abstain. We assume away such a cost and focus on other tools to restore incentive compatibility. Finally, Grüner and Tröger (2019) also study costly voting in a static environment and find a linear voting rule is utilitarian-optimal. In the single-election scenario of our paper, without voting costs, we find that the first-best and the second-best policies are both linear voting rules,[3] as they can be represented by a line that separates the state space for different policy implementations. But when agents vote in continuous time, the first-best policy does not need to have this linear property, but it has a (weaker) monotonic property. The second-best policy is more complex and depends on the general history, rather than just the state.

Another strand of the voting literature studies information aggregation, starting from the rediscovery of the Condorcet Jury Theorem by Black (1958) with sincere voting. A growing literature (Feddersen and Pesendorfer (1997); Dekel and Piccione (2000); Bhattacharya (2013); Barelli et al. (2020) for formal elections, and Battaglini (2016); Ekmekci and Lauermann (2019) for informal elections) shows that in an environment where voters' preferences are common or depend on

[3]In particular, according to Grüner and Tröger (2019), the second-best policy is a one-sided voting rule.

some common fundamentals, such information can be aggregated asymptotically through (super-)majority voting even when agents vote strategically. The result is referred to as Full Information Equivalence, as (super-)majority voting produces the outcome that would be produced if all private information were public. Ali et al. (2017) identify a failure of information aggregation in a static and private values environment, when voters' payoffs have certain negative correlation. In our setup, agents' values are private and independent, and we do not study information aggregation in large elections with a fixed voting rule. Instead, we fix the size of the electorate and seek a welfare maximizing mechanism when voters have to learn their private values gradually. But in the same spirit, we compare public learning and private learning, and show the efficiency loss due to information asymmetries.

In contrast to the above static voting literature, a small literature considers a dynamic environment where voters' private preference may change over time. Fernandez and Rodrik (1991) and Gersbach (1993) examine how majority voting makes a voter biased in favor of the status quo policy, in an environment where the individual benefit of the alternative policy is uncertain and is only revealed to voters when it is implemented or it is feasible to be implemented. The risk that a voter is a part of the minority group explains their bias. Albrecht et al. (2010) and Compte and Jehiel (2010) study a collective search problem by a committee with private values. Albrecht et al. (2010) compare search by committee by (super-)majority rule to the corresponding single-agent problem and show that each committee member is less picky and more conservative than the single agent, due the externalities imposed by other committed members. Compte and Jehiel (2010) analyze which member have more impact on the decision under majority rule and the degree of randomness of the decision. Messner and Polborn (2012) and Moldovanu and Rosar (2021) examine the option value of waiting when the reform policy is irreversible and analyze different (super-)majority rules. Although an individual voter's type may change over time, they assume either such a change is independent of the voter's past type, or the distribution of voters' types stays constant over time. In such a stationary environment, they find some super-majority rule is optimal. In particular, Moldovanu and Rosar (2021) show that a utilitarian planner's optimal policy can be implemented by a carefully chosen super-majority rule. In our non-stationary environment, an uninformed agent will gradually become and stay informed about his preference, but not vice versa. Thus, the optimal mechanism in our setting is complex and simple mechanisms such as (super-)majority rules are sub-optimal.

A few papers go beyond (super-)majority rules. Casella (2005) considers a repeated collective choice problem where voters' private preference may change over time. She proposes a dynamic voting rule where, for each election, a majority rule applies but voters are allowed to store their vote for future use, and shows that using storable votes can achieve a more efficient outcome than non-storable votes. This idea relates to the linking mechanism in dynamic mechanism design without money. Chan et al. (2018) study sequential information acquisition by a committee where, at each date, agents decide whether to adopt one of two options or to continue information acquisition. They show that majority voting and super-majority voting have their own advantages, as they assume agents are heterogeneous not only in their preferences over the options, but also in how patient they are. Thus, they propose a two-step decision rule in which the first step is to vote on whether to stop collecting information and the second step is to vote on which option to adopt. Our paper also examines sequential information acquisition. Instead of studying public learning about some common fundamentals as in their model, we consider private learning about private values.

More closely related to our work, Strulovici (2010) also studies collective experimentation and decision-making in an environment where agents learn their preferences gradually.[4] That paper adopts a different signal structure—it embeds the exponential bandit of Keller et al. (2005)—the main differences being that, first, an agent can only conclusively learn that she prefers the status quo to the reform (the complementary event cannot be learnt) and, second, an uninformed agent's posterior belief that the status-quo benefits her drifts down in the absence of a signal. Nevertheless, as in our setup, an individual's lack of control over the collective decision makes her less keen to experiment than in the single-player benchmark. A more fundamental difference with our paper is that Strulovici (2010) either assumes that signals are publicly observed when comparing the social welfare induced by different voting rules, or focuses on the majority voting rule when considering private signals. It does not address the question at the center of our paper: what is the welfare-maximizing voting rule when signals are privately observed? Moreover, the signal structure in Strulovici (2010) implies that at each time there are only two types of agents. In such a setting, it can be shown that the optimal policy under public learning and private learning coincide, as agents have the right incentive to report their types truthfully even if learning were private, as long as the optimal policy with public learning is implemented.[5] With three types as in our setting, the optimal policy with public learning fails to be incentive compatible under private learning, which is the central issue in our paper.

Our paper also relates to the literature of dynamic mechanism design without money. Grenadier et al. (2015) and Guo (2016) study a dynamic delegation problem where an agent has private information only at the start of game. When agents acquire private information over time, Guo and Hörner (2020), Jackson and Sonnenschein (2007), and Balseiro et al. (2019) consider how a

[4] Callander (2011) examines the experimentation pattern in repeated elections with majority voting. As he considers myopic and non-strategic voters, the problem is essentially a single-agent (the median voter) experimentation problem.

[5] Given that the planner implements the optimal policy under public learning: (1) if the interests of the two types are aligned, then it coincides with the planner's policy; (2) if there is a conflict of interests between the two types, then the best they can do is to report their type truthfully to make the planner side with them.

principal, without using monetary transfers, should optimally or asymptotically optimally allocate resources to agents repeatedly when they have private information. As the allocation problem occurs repeatedly, the idea is to link allocations intertemporally (via a fixed quota mechanism in Jackson and Sonnenschein (2007), and a path-dependent "quantified entitlement" mechanism in Guo and Hörner (2020)) as a tool to elicit private information efficiently.

In contrast, Escobar and Zhang (2020), as well as our paper, study a problem where only one single irreversible decision is to be made. They consider a dynamic delegation problem where a principal relies on an agent's private information to learn the profitability of an investment. Similar to our model, the principal and the agent disagree on when to stop experimentation to invest in the project. They find that the principal's optimal delegation delays the investment upon receiving conclusive good news until some deadline, in order to relax the agent's incentive to misreport such news. Our second-best mechanisms also feature delays in implementing the policy toward which an uninformed agent is biased and deadlines in experimentation, which distorts the efficient implementation of the first-best policy but induces truth-telling by the uninformed agent. The main difference is that we consider multiple agents. The new challenge we face is that the conflict of interest between the planner and an agent not only depends on the agent's discovery of his preference, but also depends on other agents' discoveries of their preferences. This feature also provides us an additional tool—private reporting—to elicit information. We show that efficiency may be improved if the planner keeps an agent in dark about other agents' reports.

The feature of delays and deadlines in our paper also appears in Damiano et al. (2012) and Damiano et al. (2021). They study collective decision-making in a committee with two members in a class of mechanisms called "delay mechanisms," in which a simple majority rule determines the collective decision if at least one member votes for his opponent's favorite alternative, but otherwise a costly delay is incurred before the committee vote again. They assume private information is obtained at the beginning of the game so that the decision itself is static, but their model resembles a war of attrition game so that a dynamic mechanism involves delays as a collective punishment for disagreement and improves efficiency.

Finally, we apply the revelation principle in multistage games (Forges (1986); Myerson (1986); Sugaya and Wolitzky (2020)) in our private communication scenario (i.e., agents' reports are only observable by the planner, not by other agents). In particular, we apply Forges (1986) and use Bayes Nash Equilibrium as our solution concept, as there is no observable deviations in our setting. We also study direct truthful public mechanisms, where the history of reports is made public to all players and we require that truth-telling is a best response regardless of the prior reports of other players. This is related to periodic ex post incentive compatibility in Bergemann and Välimäki (2010). It is with loss to consider public mechanisms.[6] We see public mechanisms as an important benchmark. The planner may have legal obligations to disclose voting information as otherwise the social choice outcome is subject to endless dispute. In some situations, such as in protests, it is also infeasible to conceal such information as "reports," i.e., the number of people who are on the street supporting the status quo policy or militating for a policy reform, are public by nature. Moreover, from the point of view of exposition, we use public mechanisms to illustrate the incentive problem first.

The paper is organised as follows. Section 2 describes the model. Section 3 describes the first-best mechanism when agents make reports in continuous time. We derive conditions under which the first-best is not incentive compatible. Those failure of incentive-compatibility are best illustrated in the public communication setup. In Section 4, we derive the second-best mechanism in an example with one pivotal voter, under public communication. We show that incentive compatibility requires inefficient delays and deadlines. In Section 5, we derive the second-best mechanism in an example with two pivotal voters, but this time under private communication. Section 6 considers the problem when there is a single election. Section 7 concludes and outlines open questions and work in progress. Appendix A contains proofs. Appendix B provides a formal analysis in the discrete time framework for Section 5. Appendix C proves a revelation principle relevant to our continuous-time problem.

2 The model

Time $t \geq 0$ is continuous and payoffs are discounted at rate r. There is one principal, or social planner (he), and $n \geq 2$ agents (she). Each agent's preference type $\theta_i \in \{A, B\}$ is independently and identically distributed, where $\Pr(\theta_i = A) = p$. The "default" policy is policy A, generating i's (unobservable) flow payoff of $\alpha > 0$ if $\theta_i = A$ and -1 if $\theta_i = B$. At each time t, the principal either continues policy A, or irreversibly switches to policy B. If policy B is implemented, then i's payoff is 0 regardless of θ_i.

Individual, private learning: Learning is exogenous and costless (hence, no moral hazard). At a random time, agent i privately observes a signal that conclusively reveals θ_i. The agents' signal arrival times are identically and independently distributed according to an exponential distribution with parameter $\lambda > 0$. The absence of a signal is uninformative regarding an agent's type. Therefore, the agent's posterior belief regarding her type remains constant, equals to the prior, p, until she receives a signal. We let the random variable $p_t^i \in \{0, p, 1\}$ denote agent i's

[6] Except for the single-election problem studied in Section 6, where agents make simultaneous and private reports and the usual revelation principle applies.

posterior belief at t, with $p_t^i = 1$ ($p_t^i = 0$) signifying that the agent has leant that her preference type is A (B). We allow learning to start at some date $-\underline{T} < 0$, prior to the beginning of the game. Consequently, agents are heterogeneous at the outset of the game, with some agents having already learnt their preference type. This assumption is pertinent in applications where the policy reform addresses an old issue over which part of the population already has clear ideas.

Single agent benchmark: Consider an uninformed single agent who experiments with A, i.e. who switches to B upon receiving a B signal but maintains policy A otherwise. If $\theta_i = A$, the agent never receives a B signal, so she collects a flow payoff α in perpetuity. If $\theta_i = B$, the agent collects a flow payoff of -1 under policy A until the arrival of a B signal, when she switches to policy B and gets a continuation payoff of zero. Thus, the single agent's payoff from experimenting is

$$v^\star(p) := p\alpha + (1-p)\int_0^\infty e^{-\lambda t}\lambda(1-e^{-rt})(-1)dt = p\alpha - (1-p)(1-g_1),$$

where $g_1 := \lambda/(r+\lambda)$ is the expected discounted time until a signal arrives. Experimenting with A is better than immediately switching to B if and only if $v^\star(p) \geq 0$ or, equivalently,

$$p \geq \frac{1-g_1}{1+\alpha-g_1} =: p^\star, \tag{1}$$

where $p^\star$ denotes the single agent (experimentation) threshold.

Now suppose that the agent must decide, once and for all, wether to switch to policy B or commit to policy A. Adhering to policy A induces a payoff $v^M(p) := p\alpha - (1-p)$, and is better than switching to policy B if and only if $v^M(p) \geq 0$ or, equivalently,

$$p \geq \frac{1}{1+\alpha} =: p^M, \tag{2}$$

where p^M denotes the single agent myopic threshold. Observe that $v^\star(p) > v^M(p)$ for every prior $p \in (0,1)$, as being able to wait for one's signal before committing to a policy is valuable. Consequently, $p^M < p^\star$.

Social choice: The principal seeks to optimally translate the agents' learning into a social choice. His goal is to maximize the expected discounted sum of the agents' payoffs. He designs a mechanism that specifies a decision rule mapping the agents' reports into a social choice. We focus on a direct mechanism where the principal asks voter i's information, i.e. her posterior belief, at each $t \in T$, where T denotes the set of feasible reporting times. We will consider two possible setups. The first one is when reports are made in continuous time, so $T = [0, \infty)$. This corresponds to *protests*, where agents can join a movement in support of their favoured policy at any point in time. The second one is when reports are made at a single election date, which corresponds to a *referendum*, so $T \in [0, \infty)$.

We assume first all reports are public. At each date $t \in T$, agent i submits a report $\mu_t^i \in \{0, p, 1\}$. The principal and all agents observe the history of past and current reports, $\boldsymbol{\mu}_t^i := (\mu_x^i)_{x \leq t, x \in T}$. We also consider private reports later. That is, the history of reports is only observed by the principal. Agents, beyond their private information, only observe the calendar time and the history of policies, which is essentially whether the policy has been switched to B yet given that policy B is irreversible.

A direct mechanism specifies, a date at which policy B is irreversibly implemented, as a function of the history of reports. Formally, it is a stopping time, $\xi \in T$, with respect to the filtration induced by the history of reports, $\{\boldsymbol{\mu}_t^i\}_{i=1}^n$. We say that agent i *reports truthfully* if $\mu_t^i = p_t^i$ at each $t \in T$. We study truthful direct mechanisms, where all agents reports truthfully.

Without loss of optimally, we restrict attention to decision rules that ignore an agent's messages subsequent to her announcing that she has observed a signal. Specifically, we restrict attention to decision rules in the following class: Let $t^i := \inf\{t \in T : \mu_t^i \in \{0, 1\}\}$ denote the first date at which agent i reports a signal. The decision rule then only conditions on the history of agent i's messages up to and including date t^i. One implication is that (mis-)reporting a signal cannot later be overridden with a different message.

3 Continuous-time problem and first-best mechanism

In this section, we begin by deriving the planner's first-best policy in continuous time, assuming the agents report their information truthfully. We then derive sufficient conditions on the parameters α and p for truth-telling not to be incentive-compatible under the planner solution. We argue that these conditions are permissive enough to conclude that the planner solution is not implementable for any "reasonable" parametrisation of the model. For now, we focus on setups where there is at least one history of reports at which an uninformed agent has an incentive to misreport a B signal, bearing in mind that there also are setups where an uninformed agent might prefer to misreport an A signal. In addition, we momentarily abstract from pre-time-zero learning, so that our results are applicable regardless of the duration $\underline{T}$ of the learning period. Indeed, if for some realisation of past reports, the planner solution is not incentive compatible when $\underline{T} = 0$, then it will be vulnerable to misreporting when $\underline{T} > 0$, as the realised prior learning and induced reports at date 0 necessarily correspond to a history reached with positive probability over the course of play in the absence of prior learning.

3.1 The public reporting game:

At each $t \geq 0$, each agent i makes a public report $\mu_t^i \in \{0, p, 1\}$ regarding her private information. To allow agents to instantaneously react to other agents' reports, we model the interaction as a multistage game, as in Murto and Välimäki (2013).[7] Each stage is characterised by the number of agents having previously reported A or B signals. An agent's strategy prescribes a date at which she reports a signal, as a function of her private information and of the number of previous A and B votes.

Formally, we consider a multistage game with at most n stages. For each stage $k \geq 1$, let $s_k \in [0, \infty]$, $s_k \leq s_{k+1}$, denote the date at which stage $k-1$ ends and stage k begins. In stage $k \geq 1$, let (a_k, b_k) denote the number of agents having reported A, B signals in stages $0, \ldots, k-1$, and let $m_k := n - a_k - b_k$ denote the associated number of undeclared agents. Let $s_0 := 0$ and $(a_0, b_0, m_0) := (0, 0, n)$. In stage $k \geq 0$, a strategy for each undeclared agent i has two elements: a stopping time $\sigma_k^i(a_k, b_k) \in [s_k, \infty]$ with respect to the natural filtration induced by player i's private signal process, and the report $\rho_k^i(a_k, b_k, p_{\sigma^i}^i) \in \{0, 1\}$ that the agent makes at date $\sigma_k^i(a_k, b_k)$, with the interpretation that agent i reports being uninformed at each $t \in [0, \sigma_k^i(a_k, b_k))$ and reports an A or B signal at date $\sigma_k^i(a_k, b_k)$. The first date in $[s_k, \infty]$ at which at least one of the m_k undeclared agents reports a signal, A or B, determines s_{k+1}.

3.2 Planner solution

We assume that all agents report their information truthfully, and derive the policy maximising the joint payoff. Observe that, under truthful reporting, the agent's strategies are invariant to other agents' reports. Consequently the first-best policy is optimal both under public and under private reporting. At each point in time when the principal makes a policy decision, let $a \in \{0, 1, \ldots, n\}$ denote the number of declared A-type agents, $b \in \{0, 1, \ldots, n\}$ the number of declared B-type agents, and $m = n - a - b \geq 0$ denote the number of undeclared agents. Since the signal arrival processes of the agents are stationary, it is without loss of generality to let the date be $t = 0$.

Let $V(a, b, p)$ denote the value function in the planner problem. For every (a, b, p), it satisfies

$$V(a, b, p) = \max\{V^A(a, b, p), V^B(a, b, p)\}, \tag{3}$$

where $V^B(a, b, p) = 0$ denotes the joint payoff from irrevocably switching to B, and

$$V^A(a, b, p) = (1 - g_m)\, u(a, b, p) + g_m \left[p\, V(a+1, b, p) + (1-p)\, V(a, b+1, p)\right] \tag{4}$$

[7] See also Stinchcombe (1992) and Khan and Stinchcombe (2015).

denotes the payoff from experimenting with A where $u(a,b,p) := \alpha\, a - b + m\, v^M(p)$ denotes the flow payoff from implementing policy A in state (a,b,p), and where $g_m := \mathbb{E}[e^{-r\tau_m}] = \frac{m\lambda}{r+m\lambda}$ is the expected discounting until the date τ_m at which the next signal is observed by one of the m remaining uninformed agents. Observe that g_m is strictly increasing in m: with more uninformed agents the next signal will arrive sooner.

The next lemma allows us to compare the two value functions in the second term of (4). Fix the number $m \in \{0,\dots,n-1\}$ of uninformed agents. Changing the type of one informed agent from A to B strictly decreases the joint payoff from policy A, causing a weak decrease in the planner value function V.

Lemma 1. *For every pair (a,b) such that $1 \le a+b \le n$ and for every $p \in (0,1)$ we have that $V^A(a+1,b,p) > V^A(a,b+1,p)$.*

Next, we show that for every number $m \ge 1$ of uninformed agents, the joint payoff from policy A strictly increases with the probability p with which each of these uninformed agents is an A type.

Lemma 2. *Fix (a,b) such that $a+b<n$. Then $V^A(a,b,p)$ is continuous and increasing in $p \in [0,1]$ and differentiable at almost every $p \in [0,1]$.*

We are now ready to state the main result, that the planner policy can be expressed as a threshold policy with respect to the prior belief p. That is, for every pair (a,b), there exists a threshold belief $p^*(a,b)$ such that $V^A(a,b,p) \ge 0$ if and only if $p \ge p^*(a,b)$. We show that $p^*(a,b)$ is decreasing in a and increasing in b. Intuitively, this means that the planner switches to B once the number of declared B types is sufficiently large relative to the number of declared A types and the number of uninformed agents.

Proposition 1. *For every pair (a,b) such that $a+b<n$, there exists a unique $p^*(a,b) \in [0,1]$ satisfying $V^A(a,b,p) \ge 0$ if and only if $p \ge p^*(a,b)$. If $p^*(a,b) \in (0,1)$, then (i) $p^*(a,b) > p^*(a+1,b)$, and (ii) $p^*(a,b) < p^*(a,b+1)$.*

Next, we state some properties of the planner solution. Consider the term $u(a,b,p)$ in (4). It represents the payoff from never switching to B. Since $m(1+\alpha) > 0$, this payoff strictly increases with p whenever $m \ge 1$, i.e. whenever there is at least one uninformed agent. Consequently, $\alpha a - (b+m) \le u(a,b,p) \le \alpha(a+m) - b$, where the lower (upper) bound is the payoff from adhering to policy A forever if all m uninformed agents turn out to be B (A) types.

The next lemma shows that if $\alpha a - (b+m) > 0$ or, equivalently, if $a > n/(1+\alpha)$, then there are sufficiently many A types that committing to policy A forever is optimal. Conversely, if $\alpha(a+m) - b < 0$ or, equivalently, if $b > n\alpha/(1+\alpha)$, then there are sufficiently many B types that switching to policy B is optimal.

Lemma 3. *(i) If $a > n/(1+\alpha)$, then $p^*(a,b) = 0$ so that $V(a,b,p) = V^A(a,b,p)$ for every $p \in (0,1)$. (ii) If $b > n\alpha/(1+\alpha)$, then $p^*(a,b) = 1$ so that $V(a,b,p) = V^B(a,b,p)$ for every $p \in (0,1)$.*

We let $\bar{a} := \lfloor n/(1+\alpha) \rfloor$ and $\bar{b} := \lfloor n\alpha/(1+\alpha) \rfloor$, and let $S_n := \{(a,b,m) \in \mathbb{N}^3 : a+b+m = n\}$ denote the set of all possible states, and $S_n^* := \left\{(a,b,m) \in S_n | a \leq \bar{a}, b \leq \bar{b}\right\}$ denote the set of feasible states in which there are not yet enough A or B types to conclusively determine the optimal policy, so that the planner threshold is strictly between 0 and 1.

For each state $(a,b,m) \in S_n^*$, there exists a unique "myopic" threshold belief $p^M(a,b) \in [0,1]$ such that $u(a,b,p^M(a,b)) = 0$. Observe that we necessarily have that $p^*(a,b) < p^M(a,b)$. This is because, for each $(a,b,m) \in S_n^*$, if the prior equals $p^*(a,b)$ and one of the uninformed agents learns that she is an A type, then continuing to experiment with A is strictly better than switching to B. That is, we have $V^A(a+1,b,p^*(a,b)) > 0$, implying that the continuation value in state (a,b) is strictly positive. Therefore, $V^A(a,b,p^*(a,b)) = 0$ necessitates a negative flow payoff, i.e. $u(a,b,p^*(a,b)) < 0$.

Finally, we let $S_n^{**} \subset S_n^*$ be the set of states where $p > p^*(a,b)$, so that continuing with A is strictly preferred under the planner policy, and $S_n^{***} \subset S_n^{**}$ be the set of states in which A is optimal and in which one more vote in favour of B makes the planner switch to B. Thus S_n^{***} can be thought of as the set of states in which a B vote is pivotal.

3.3 Example: Majority rule

In this example, we specialise the model as follows: $\alpha = 1$, $p = p^*$, $\lambda \leq r$ and $n < \bar{n}(r,\lambda)$, where $\bar{n}(r,\lambda)$ is a strictly increasing function of r and strictly decreasing function of λ, defined below

(5) $$\bar{n}(r,\lambda) := 1 - \frac{1}{v^M(p^\star)} = \frac{2(r+\lambda)}{\lambda}$$

Under this parametrisation, the first-best policy corresponds to the majority rule: the report policy is implemented the first time that the number of votes in favour of the reform is no less than the number of votes in favour of the status quo policy. It is illustrated in Figure (1). Here, $\bar{a} = \bar{b} = \frac{n-1}{2}$. Thus, in the state $(\bar{a}, \bar{b})$, there is one uninformed agent, and she is pivotal. The set of states S_n^{***} in which a B vote is pivotal (i.e. it induces the planner to implement the reform) has $a = b+1$.

Here is an intuition for why, under these conditions, the planner solution corresponds to the majority rule. The formal proof is in Appendix A.6. Consider the undiscounted limit where $r \to \infty$. Then $p^\star \uparrow p^M$. Moreover, $g_m \downarrow 0$, so that $V^A(a,b,m,p) \to u(a,b,m,p)$. It is then immediate that $V^A(a,a,n-2a,p) \to (n-2a)v^M(p)$, where $v^M(p)$ denotes the myopic payoff.

$a \backslash b$	0	1	2	$\cdots$		$\frac{n\alpha}{1+\alpha}$			$\cdots$			n
0	B	B	B	B	B	B	B	B	B	B	B	B
1	A	B	B	B	B	B	B	B	B	B	B	
2	A	A	B	B	B	B	B	B	B	B		
$\vdots$	A	A	A	B	B	B	B	B	B			
	A	A	A	A	B	B	B	B				
$\frac{n}{1+\alpha}$	A	A	A	A	A	B	B					
	A	A	A	A	A	A						
	A	A	A	A	A							
$\vdots$	A	A	A	A								
	A	A	A									
	A	A										
n	A											

Figure 1: First best policy when $n = 11$, $\alpha = 1$, $\lambda = 2$, $r = 10$, $p = p^\star$.

Since $v^M(p^\star) < 0$ by the definition of $p^\star$, we have that $\lim_{r\to\infty} V^A(a, a, n-2p, p^\star) < 0$. Similarly, $V^A(a, a-1, n-2a+1, p) \to 1 + (n-2a-1)v^M(p)$. Since $1 + (n-2a-1)v^M(p^\star)$ is strictly increasing in a, a sufficient condition for $\lim_{r\to\infty} V^A(a, a-1, n-2a+1, p^\star) > 0$ is to have $\lim_{r\to\infty} V^A(1, 0, n-1, p^\star) > 0$ or, equivalently, $n < \bar{n}(r, \lambda)$.

This parametrisation is convenient because it allows us to derive closed-form expressions for the agents' payoffs under the first-best policy.

Lemma 4. *Suppose that S_n^{***} is the diagonal $\{(a, b, m)|a = b + 1 \leq \bar{a}\}$. Then, for each state $(a, b, m) \in S^{**}$, the joint payoff under the planner policy is given by*

$$V^A(a, b, m) = u(a, b, m) - \sum_{k=a}^{\bar{a}} C(a, b, k)\, p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{n-a-b} g_j \right) u(k, k, n-2k), \tag{6}$$

where

$$C(a, b, k) = \binom{2k-a-b-1}{k-a} \frac{a-b}{k-b}. \tag{7}$$

To understand these expressions, observe that $C(a, b, k)$ is the number of paths from state $(a, b, m) \in S_n^{**}$ to state $(k, k, n-2k)$, $k \in \{a, \ldots, \bar{a}\}$ that only visit states within the planner policy's continuation region S_n^{**} or, equivalently, that never visit the diagonal $\{(a, b)|a = b \leq \bar{a}\}$ that forms the boundary of the stopping region (except when reaching the state (k, k)). Observe that $C(a, a-1, k) = \binom{2(k-a)}{k-a}\frac{1}{k-a+1}$ is the $(k-a)^{\text{th}}$ Catalan number.[8]

[8] The n^{th} Catalan number, $C_n := \binom{2n}{n}\frac{1}{n+1}$, is the number of lattice paths in $\mathbb{Z}^2$ from $(0, 0)$ to (n, n) that are monotonic, i.e. with steps (1,0) and (0,1), such that the path never rises above the line $y = x$. See Stanley (2015).

3.4 Failures of Incentive Compatibility

We begin by describing the reporting game induced by the planner policy. Lemma 5 shows that, in a state where there is a unique uninformed agent who is pivotal, that uninformed agent has an incentive to misreport a B signal whenever the prior p lies below the single agent experimentation threshold $p^\star$. Incentives to misreport under the planner policy can also arise in states with more than one uninformed agent.

3.4.1 A pivotal event

Suppose that there are $n \geq 2$ agent. Fix $\alpha > 0$ such that $\frac{n}{1+\alpha}$ is not an integer. (This condition is generically satisfied.) Then, $\bar{a} + \bar{b} = n - 1$. Thus, in state $(\bar{a}, \bar{b})$, there is only one uninformed agent. Moreover, that agent is pivotal under the planner policy, as her reporting an A signal makes it socially optimal to stick with policy A forever ($u(\bar{a}+1, \bar{b}, p) > 0$), while her reporting a B signal makes it socially optimal to switch to policy B ($u(\bar{a}, \bar{b}+1, p) < 0$).

We now argue that truth-telling is optimal for the pivotal uninformed agent if and only if the prior p is no lower than the single agent threshold, $p^\star$. Because, under the planner policy, the pivotal uninformed agent's report determines the planner's policy choice, she effectively faces the single-agent experimentation problem. Her payoff from reporting truthfully equals the single-agent payoff from experimenting with A, $v^\star(p)$. This is weakly preferred to misreporting B if and only if $v^\star(p) \geq 0$ or, equivalently, $p \geq p^\star$. (It is strictly preferred to misreporting A, since the induced, myopic payoff, $v^M(p)$, is strictly below the payoff from reporting truthfully.)

We now consider $p^*(\bar{a}, \bar{b})$, the threshold belief above which it is optimal for the planner to continue with policy A in state $(\bar{a}, \bar{b})$. The planner's payoff from doing so is given by (4) and can be rewritten as

$$V^A(\bar{a}, \bar{b}, p) = \left(\alpha\bar{a} - \bar{b}\right)(1 - (1-p)\, g_1) + v^\star(p).$$

The second is the pivotal uninformed agent's payoff. The first term reflects the uninformed agent's externality on the remaining agents: the planner will maintain policy A forever, unless the uninformed agent receives a B signal, which happens with probability $1-p$ after an expected discounted duration g_1. If $\alpha\bar{a} - \bar{b} > 0$, the externality is positive, and the planner has a greater incentive to continue experimenting with A than the single agent, who does not internalise his positive effect on social welfare. We then have $p^*(\bar{a}, \bar{b}) < p^\star$. Indeed, it is easy to see that

$$p^*(\bar{a}, \bar{b}) := \frac{(1 - g_1 - (1-g_1)(\alpha\bar{a} - \bar{b})}{1 + \alpha - g_1 + g_1(\alpha\bar{a} - \bar{b})} \tag{8}$$

strictly decreases with $\alpha\bar{a} - \bar{b}$, and coincides with the single-agent threshold $p^\star$ if and only of $\alpha\bar{a} - \bar{b} = 0$.

Suppose $p^*(\bar{a},\bar{b}) < p^\star$. Then, for $p \in (p^*(\bar{a},\bar{b}), p^\star)$ there exists a conflict of interest between the planner and the uninformed agent: the planner wants to experiment with A until the uninformed agent reports her signal. However, the uninformed agent prefers misreporting a B signal immediately. For every $g_1 \in (0,1)$, the condition $p^*(\bar{a},\bar{b}) < p^\star$ holds if and only if $\alpha > \frac{\bar{b}}{\bar{a}}$. The next lemma summarises the above arguments.

Lemma 5. *For every $\alpha > \frac{\bar{b}}{\bar{a}}$ and $p \in (p^*(\bar{a},\bar{b}), p^\star)$, the planner solution is not incentive compatible in state $(\bar{a},\bar{b})$, as the pivotal uninformed agent has an incentive to misreport a B signal.*

Observe that the right-hand side, $\bar{b}/\bar{a}$, also depend on α. The next lemma expresses this condition in terms of the primitive, α.

Lemma 6. *Fix $n \geq 2$. Then $p^*(\bar{a},\bar{b}) < p^\star$ if and only if $\alpha \in \left(\frac{n-(k+1)}{k}, \frac{n-k}{k}\right)$, $k \in \{1,2,\ldots,n-1\}$.*

The proof can be summarised as follows. Fix an integer $k \in \{1,2,\ldots,n-1\}$, and recall that we exclude the values $\alpha \in \{\frac{n-k}{k}\}_{k=1}^{n-1}$, as these are the non-generic values at which the thresholds $\frac{n}{1+\alpha}$ and $\frac{n\alpha}{1+\alpha}$ from Lemma 3 are integers. For $\alpha \in \left(\frac{n-(k+1)}{k+1}, \frac{n-k}{k}\right)$ we have $\lfloor \frac{n}{1+\alpha} \rfloor =: \bar{a} = k$ and $\lfloor \frac{n\alpha}{1+\alpha} \rfloor =: \bar{b} = n-(k+1)$. Consequently, the externality imposed by the pivotal uninformed agent is positive if and only if $\alpha\bar{a} - \bar{b} > 0$ or, equivalently, $\alpha > \frac{n-(k+1)}{k}$.

3.4.2 Majority rule

Consider the parametrisation of Section (3.3). We begin by giving expressions for an agent's payoff udner the planner policy. We then show that the IC constraint is violated in every state in S_n^{***}, i.e. in every state where one B report induces the planner to implement the reform.

Informed agents: We begin by considering an agent, i, who has already observed (and truthfully reported) her preference type. She cannot make an additional report, so whether and when the planner switches to B depends entirely on the reports of the m remaining uninformed agents. Let $v(\theta^i) = \mathbb{1}\{\theta^i = A\} - \mathbb{1}\{\theta^i = B\}$ denote type θ^i's flow payoff under policy A.

Lemma 7. *Suppose that S_n^{***} is the diagonal $\{(a,b,m) | a = b+1 \leq \bar{a}\}$. Then, for each state $(a,b,m) \in S^{**}$, type θ^i's payoff under the planner policy, assuming all remaining m uninformed agents report truthfully, is given by*

$$\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\left(1 - \mathbb{E}\left[e^{-r\tau(a,b,m,p)}\right]\right), \tag{9}$$

where

$$\mathbb{E}\left[e^{-r\tau(a,b,m,p)}\right] = \sum_{k=a}^{\bar{a}} C(a,b,k)\ p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j \tag{10}$$

is the expected discount factor applied to the continuation payoff from switching to policy B. The expectation is taken over the random variable, $\tau(a,b,m,p)$, denoting the time at which the planner switches to B when the initial state is (a,b,m).

Recall that $C(a,b,k)$, defined in (7), is the number of distinct paths from state (a,b,m) to state $(k,k,n-2k)$ that do not cross into the stopping region $S_n^* \setminus S_n^{**}$ (except at the last step, to reach state $(k,k,n-2k)$). Equation (9) then says that an uninformed agent i collects a flow payoff of $v(\theta^i)$ from policy A up until the date $\tau(a,b,m,p)$ (possibly infinite) at which the planner switches to policy B. For every state $(a,b,m) \in S_n^* \setminus S_n^{**}$ in the stopping region, $\mathfrak{u}^{\theta^i}(a,b,m,p) = 0$.

Uninformed agents: Next, we consider an uninformed agent, say i, who has not yet learnt her type. She has not yet made a report, so whether and when the planner switches to B depends on her report and that of the $m-1$ other remaining uninformed agents.

Lemma 8. *Suppose that S_n^{***} is the diagonal $\{(a,b,m)|a = b+1 \le \bar{a}\}$. Then, for each state $(a,b,m) \in S^{**}$, uninformed agent i's payoff under the planner policy, assuming all remaining m uninformed agents, including agent i, report truthfully, is*

$$(11)\quad \mathfrak{u}_i(a,b,m,p) = p\left(1 - \mathbb{E}\left[e^{-r\tau(a,b,m,p)}|\theta^i = A\right]\right) + (1-p)\,(-1)\left(1 - \mathbb{E}\left[e^{-r\tau(a,b,m,p)}|\theta^i = B\right]\right),$$

where

$$(12)\quad \mathbb{E}\left[e^{-r\tau(a,b,m,p)}|\theta^i = A\right] = \sum_{k=a}^{\bar{a}} C(a,b,k)\,\frac{1}{m}\,[k-a+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j$$

and

(13)

$$\mathbb{E}\left[e^{-r\tau(a,b,m,p)}|\theta^i = B\right] = \sum_{k=a}^{\bar{a}} C(a,b,k)\,\frac{1}{m}\,[k-b+(n-2k)(1-p)]\,p^{k-a}(1-p)^{k-b-1} \prod_{j=n-2k+1}^{m} g_j.$$

Here, $\mathbb{E}[e^{-r\tau(a,b,m,p)}|\theta^i = A]$ denotes the expectation taken over the random time $\tau(a,b,m,p)$ at which the planner switches to B when the initial state is (a,b,m), conditional on uninformed agent i's type being, in fact, A, and where $C(a,b,m)$ is given in (7). This expectation differs from the unconditional expectation in (10), as it accounts for the fact that any B signal cannot have come from agent i, as i's signal would surely be A. Observe that

$$p\,\mathbb{E}\left[e^{-r\tau(a,b,m,p)}|\theta^i = A\right] + (1-p)\,\mathbb{E}\left[e^{-r\tau(a,b,m,p)}|\theta^i = B\right] = \mathbb{E}\left[e^{-r\tau(a,b,m,p)}\right].$$

Equation (11) says that if the uninformed agent i has in fact type $\theta^i = A$ ($\theta^i = B$), then she will expect to collect a payoff of 1 (-1) until the date $\tau(a,b,m,p) \in [0,\infty]$ at which the planner switches to policy B, where agent i's expectation over $\tau(a,b,m,p)$ takes account of the fact that she herself can only report an A (B) signal. For every state $(a,b,m) \in S_n^* \setminus S_n^{**}$ (in the stopping region), $\mathfrak{u}_i(a,b,m,p) = 0$.

Failure of IC: We focus on the IC constraint that says that an uninformed agent prefers reporting truthfully to misreporting B news. In state $(a, b, m) \in S_n^{***}$, misreporting B news induces the planner to switch to B, and the agent's payoff is zero. Therefore, the IC constraint is

$$\mathfrak{u}_i(a, b, m, p) \geq 0. \tag{14}$$

The next lemma shows that this IC constraint is violated for every state in S_n^{***}, where one more B vote induces the planner so switch to policy B.

Lemma 9. *For each $(a, b, m) \in S_n^{***}$ and $p = p^\star$, the IC constraint* (14) *is violated. Moreover, the payoff $\mathfrak{u}_i(a, a-1, n-2a+1, p)$ from reporting truthfully increases with a.*

The proof shows that, in state $(\bar{a}, \bar{a}-1)$, two agents are pivotal and have a strict incentive to misreport B news. Indeed, the agents import a negative externality on each other, as the first agent to report might trap the other agent in an undesired state. Consequently, their payoff $\mathfrak{u}_i(\bar{a}, \bar{a}-1, p)$ is strictly less than the single agent experimentation payoff $v^\star(p)$, which equals zero at $p = p_\star$. The payoff $\mathfrak{u}_i(a, a-1, n-2a+1, p)$ decreases with the number of uninformed agents, as the externality just described becomes more acute.

4 Second-best mechanism under public communication

In this section, we derive the second-best mechanism with two agents under public communication. That is, the mechanism maximising the ex-ante expected sum of agents' payoffs, subject to the constraint that it is optimal for each agent to report her private information truthfully at every date $t \geq 0$.

We focus on the setup where $\alpha < 1$. In this case, $\bar{a} = 1$ and $\bar{b} = 0$, so that, under the first-best policy, it is optimal to switch to B immediately as soon as there is any B report, and it is better to stay in A forever in case both agents report A. Moreover, we assume that, as a result of pre-time-zero learning, one agent, the "expert", has already learnt that her type is A, whereas the other agent, the "novice", has not yet observed a signal, and we assume that both agents have reported their private information truthfully in stage 0. That is, we focus our analysis on stage 1 of the game, in the state $(a, b) = (1, 0)$ where the novice is pivotal. Since $\alpha \in (0, 1)$ and $n = 2$ satisfy the conditions of Lemma 6, we have that $p^*(1, 0) < p^\star$. We choose a prior $p \in (p^*(1, 0), p^\star)$, so that there is a conflict of interest between the planner and the pivotal novice: the planner wants to wait for the novice's signal before committing to a policy, whereas the novice strictly prefers switching to policy B immediately. The first-best policy is illustrated in Figure 2. Observe that this is a special case of our "majority rule" parametrisation.

A brief aside on mechanisms with private reports: the pivotal novice's incentive problem is relevant also in the first-best mechanism with private communication (i.e., voters' reports are private, and a voter is only informed as to whether the policy has switched to B or not yet), because a novice voter at time 0 knows that the only case where the principal can stay in A at time 0 is when the other voter reports an A signal.

a \ b	0	1	2
0	B	B	B
1	A	B	
2	A		

Figure 2: First-best policy. Truthful reporting is not incentive compatible in state $(1, 0)$.

We now derive the optimal mechanism which is public, direct, incentive compatible in state $(1, 0)$. A mechanism, F, consists of three cdfs: F_A, F_B and F_N, where $F_A(t'; t)$ ($F_B(t'; t)$) denotes the cdf of switching to policy B by time t' in case the novice reports A (B) at time t, conditional on policy A being in place at time t, and $F_N(t')$ denotes the cdf of switching to policy B by time t' in case the novice has been reporting no signal until then.

First, we describe the principal's problem. To lighten notation, we set $r = 1$. Let $A(t) := 1 - \int_{t'=t}^{\infty} e^{-(t'-t)} dF_A(t'; t)$ and $B(t) := 1 - \int_{t'=t}^{\infty} e^{-(t'-t)} dF_B(t'; t)$. This corresponds to the expected continuation payoff of a voter with flow payoff 1 in policy A (and 0 in policy B), conditional on policy A being in place at time t and on the voter reporting either A or B at time t.

The principal objective is a constant plus:

$$
\begin{aligned}
2\alpha p(1-q) \int_{t=0}^{\infty} \lambda e^{-\lambda t} [(\int_{\tau_N=0}^{t} (1 - e^{-\tau_N}) dF_N(\tau_N) + (1 - F_N(t)) A(t)] dt \\
-(1-p)(1-\alpha)(1-q) \int_{t=0}^{\infty} \lambda e^{-\lambda t} [(\int_{\tau_N=0}^{t} (1 - e^{-\tau_N}) dF_N(\tau_N) + (1 - F_N(t)) B(t)] dt \\
-(1-p)(1-\alpha) q B(0),
\end{aligned}
$$

where $q := (1 - e^{-\lambda \underline{T}})$ is the probability of the novice observing a signal by time 0.

Among the novice's incentive compatibility constraints, we focus on the ones ensuring an uninformed novice does not benefit from misreporting a B signal. We later verify that, under the second-best mechanism, the novice has no incentives to otherwise misreport. For each τ:

$$
\begin{aligned}
p\alpha \int_{t=\tau}^{\infty} \lambda e^{-\lambda(t-\tau)} [(\int_{\tau_N=\tau}^{t} (1 - e^{-(\tau_N - \tau)}) \frac{dF_N(\tau_N)}{1 - F_N(\tau)} + \frac{1 - F_N(t)}{1 - F_N(\tau)} A(t)] dt \\
-(1-p) \int_{t=\tau}^{\infty} \lambda e^{-\lambda(t-\tau)} [(\int_{\tau_N=\tau}^{t} (1 - e^{-(\tau_N - \tau)}) \frac{dF_N(\tau_N)}{1 - F_N(\tau)} + \frac{1 - F_N(t)}{1 - F_N(\tau)} B(t)] dt \\
\geq (p\alpha - (1-p)) \frac{1}{1 - F_N(\tau)} B(0),
\end{aligned}
$$

or equivalently:

$$
\begin{aligned}
p\alpha \int_{t=\tau}^{\infty} \lambda e^{-\lambda(t-\tau)}[(\int_{\tau_N=\tau}^{t} (1-e^{-(\tau_N-\tau)})dF_N(\tau_N) + (1-F_N(t))A(t)]dt \\
-(1-p)\int_{t=\tau}^{\infty} \lambda e^{-\lambda(t-\tau)}[(\int_{\tau_N=\tau}^{t} (1-e^{-(\tau_N-\tau)})dF_N(\tau_N) + (1-F_N(t))B(t)]dt \\
+(1-p-p\alpha)B(0) \geq 0.
\end{aligned}
$$

The following proposition describes the second-best mechanism. It makes the novice indifferent, at each t, between reporting truthfully and misreporting a B signal by delaying the resulting switch to policy B.

Proposition 2. *The solution to the principal's problem has a critical time t^* such that*

(i) $F_N^*(t) = 0$ for $t < t^*$, and $F_N^*(t) = 1$ for $t > t^*$;

(ii) F_B^* is such that there is finite delay for $t < t^*$, and no delay for $t > t^*$;

(iii) $F_A^*(t';t) = 0$ for all (t',t). (B is never implemented after A news.)

The proof verifies that above mechanism solves the principal's constrained problem. Delaying the switch to B after an A signal increases the principal's objective while simultaneously relaxing the novice's IC constraint at each $\tau \geq 0$. Therefore, it is optimal to set $A(t) = 1$. Conversely, F_B^* satisfies the IC with equality. Finally, we show that the Lagrangian associated with the problem is linear in $F_N^*(t)$ for each t, and that the coefficient $\phi(t)$ on $F_N^*(t)$ is positive for large t, establishing that $F_N^*(t) = 1$ is optimal for $t > t^*$ (possibly $t^* = 0$). If $\phi(t)$ has multiple roots, we use fact that $F_N^*(t)$ must be non-decreasing to conclude that $F_N^*(t) = 0$ for $t < t^*$.

The optimal plan above can be interpreted both as a deterministic delay policy, where the switching date $\tau_B(t)$ induced by the novice reporting a B signal at date t satisfies $\tau_B(t) \leq t^*$. Or, it can be implemented as a stochastic policy, where the novice reporting a B signal at date t induces a switch to B at t with probability $\sigma_B(t)$ and a switch at t^* with complementary probability.

Thus, the solution features two kinds of distortions to the first-best policy. In state $(1,0)$, the first-best policy remains in policy A—it is efficient for the planner to continue experimenting on the remaining uninformed agent's type. However, the second-best policy sets a *deadline* t^* for experimentation. As the planner and the uninformed agent have a conflict of interests on whether continue experimentation, setting a deadline is a compromise and it increases the agent's truth-telling incentive. In state $(1,1)$, the first-best policy switches to policy B immediately. However, the planner scarifies the efficiency by delaying the policy switch. Although this is inefficient for both the planner and the B-type agent, it reduces the uninformed agent's incentive to misreport B. Combining the deadline and the delays in those two states, it makes the uninformed agent always indifferent between truthful reporting and misreporting. Moreover, it can be shown that although the policy switch has a delay, but the switching time is increasing in the reporting time and is always before the deadline. Thus, the B-type agent also has the right incentive to immediately make a B report.

We consider public reports in this section and start our analysis under the assumption that the state $(1, 0)$ is common knowledge. As the uncertainty comes only from one single uninformed agent, public communication and private communication coincide. We can show further that this policy remains optimal when the state at time 0 is uncertain. In particular, in states where no A reports are made, the planner switches to policy B immediately. Thus, in this special case with two agents, there is no advantage to use private communication, as the uninformed agent can infer the current state when she is still in the game.

5 Second-best mechanism under private communication

In this section, we derive the second-best mechanism in an example with two pivotal voters. This time, we allow private communication. Observe that this example also fits into our "majority rule" parametrisation. As in the previous sections, we present the problem and results in the continuous-time framework. This is mainly to keep the continuity with the previous sections, but its cost is that we can only treat the problem in an informal and heuristic manner. In Appendix B, we (formally) prove the corresponding claim in the discrete-time analog of the model. In this sense, one may think of the material in the main text as the "continuous-time limit" of what we have with discrete time.

5.1 Environment and Notation

We assume that there are two uninformed voters $i = 1, 2$. The principal's flow payoff at time t is denoted by $w(a_t, b_t) = \alpha a_t - b_t - (1 - p - p\alpha)(2 - a_t - b_t) + \alpha_P$, where a_t (b_t) denotes the number of voters who have become A (B) type by time t, and α_P may be interpreted as representing the principal's own preference (e.g., his "bias" toward one policy; the cost of staying in A or switching to B; the payoffs of those who are already informed and hence have reported before the game starts[9]).

Assumption 1. *Assume that $\alpha < 1$, $\lambda < \frac{r}{2}\frac{1-\alpha}{1+\alpha}$, and that p is less than the single-agent experimentation threshold (i.e., $p\alpha - (1-p)\frac{r}{r+\lambda} < 0$).*

Also, the parameters are such that the first-best mechanism looks as follows, as a function of (a_t, b_t).

[9]Observe that this example corresponds to the majority rule setup from Section 3.3 for $n = 3$.

	$b_t = 0$	$b_t = 1$	$b_t = 2$
$a_t = 0$	A	B	B
$a_t = 1$	A	A	
$a_t = 2$	A		

The principal's second-best problem is given as follows:

$$\begin{aligned} \max_{q} \quad & \int_{t=0}^{\infty} e^{-rt} \left[\mathbb{E}_{(a_{\le t}, b_{\le t})} [q_t(a_{\le t}, b_{\le t}) w(a_t, b_t)] \right] \\ \text{sub. to} \quad & IC_t^N, \ \forall t \end{aligned}$$

where $q_t(a_{\le t}, b_{\le t})$ denotes the probability that we stay in policy A at time t, if the history of reports have been $(a_{\le t}, b_{\le t}) = ((a_\tau)_{\tau \le t}, (b_\tau)_{\tau \le t})$ by then; and IC_t^N stands for an uninformed voter's incentive compatibility constraint at time t, more explicitly given below. Although not explicitly written, a feasible plan q must satisfy the natural monotonicity requirement: If $t < t'$ and $(a_{\le t'}, b_{\le t'}) = ((a_{\le t}, (a_\tau)_{\tau \in (t,t']}, b_{\le t}, (b_\tau)_{\tau \in (t,t']})$, then $q_t(a_{\le t}, b_{\le t}) \ge q_{t'}(a_{\le t'}, b_{\le t'})$ due to the irreversibility of switching to policy B.

The uninformed voter's incentive compatibility at time t, IC_t^N, is as follows:

$$\mathbb{E}^{n-1}_{(a_{\le t}, b_{\le t})} [V_t^N(a_{\le t}, b_{\le t}) - V_t^B(a_{\le t}, (b_{<t}, b_t + 1)) \frac{p\alpha - (1-p)}{(-1)}] \ge 0,$$

where $V_t^i(a_{\le t}, b_{\le t})$ is type i's continuation payoff from t on given the number of A and B reports by time t (including his own), $(a_{\le t}, b_{\le t})$. Given that he is uninformed, under his truth-telling, $(a_{\le t}, b_{\le t})$ follow the probability distribution of the other voters' types, and thus, $\mathbb{E}^{n-1}$ represents an expectation operator with respect to $n-1 (=1)$ random variables. If at time t he misreports being a B-type, then his continuation payoff is as if he indeed were a B-type at t, $V_t^B(a_{\le t}, (b_{<t}, b_t + 1))$, times $\frac{p\alpha-(1-p)}{(-1)} (= 1 - p - p\alpha)$, reflecting the fact that the uninformed type's flow payoff is $p\alpha - (1-p)$, while the B type's is -1.

To the extent that we only consider a particular kind of incentive compatibility (and ignore the others, e.g., the informed type's "hiding" it and pretending to be yet uninformed), this problem should be interpreted as a relaxed problem. The ignored constraints are verified later.

5.2 Second-best mechanism

Here, we describe the form of the second-best mechanism. The second-best mechanism has the following feature: There exists t^* such that

1. For any $t < t^*$: (i) unless $(a_t, b_t) = (0, 2)$, stay in A; (ii) if $(a_t, b_t) = (0, 2)$, then either delay or probabilistic switch to B.

2. At $t = t^*$: (i) if $a_t = 0$, then immediately switch to B; (ii) otherwise, stay in A.

So far, this may be seen as the description of a *class* of mechanisms in the sense that I do not fully specify what happens at $t < t^*$ if $(a_t, b_t) = (0, 2)$. Our main claim is that there is at least one mechanism in this class that satisfies all the IC_t^N for $t < t^*$ with equality, and that is the second-best mechanism.

Again, the formal argument based on the discrete-time framework is in the appendix. Here, we informally explain the existence part: there exists a mechanism in the above class that satisfies all the IC_t^N for $t < t^*$ with equality. It is based on the following three observations.

- Observation 1: Consider an (extreme) version of the above mechanism where we stay at policy A for sure if $(a_t, b_t) = (0, 2)$ at any $t < t^*$ (and switch to policy B at $t = t^*$). This mechanism is feasible and satisfies all IC_t^N with $t < t^*$ with strict inequality.

  This observation is straightforward. In the above mechanism, the policy is A until $t = t^*$ with probability one. At $t = t^*$, the policy switches to B if and only if no voter reports A by then. Thus, it is strictly optimal for the uninformed voter to be truthful: Even if he misreports B at $t < t^*$, he has no influence on his continuation payoff, while he loses strictly in case he finds (after misreporting) that he is actually A type.

- Observation 2: Consider another (extreme) version of the above mechanism where we immediately switch to B for sure if $(a_t, b_t) = (0, 2)$ at any $t < t^*$. This mechanism violates the uninformed voter's IC, at any $t < t^*$.

  The argument is slightly more complicated. First, from an uninformed voter's viewpoint, at any $t < t^*$, there are three possible cases (given that he is uninformed): $(a_t, b_t) = (0, 0)$, $(a_t, b_t) = (1, 0)$, $(a_t, b_t) = (0, 1)$. If $(a_t, b_t) = (1, 0)$, he is completely indifferent between truth-telling and misreporting. If $(a_t, b_t) = (0, 1)$, then he is strictly better off by misreporting B because p is less than the single-agent experimentation threshold. If $(a_t, b_t) = (0, 0)$, then again, he is strictly better off by misreporting B for small enough λ.[10] Intuitively, this is because, with λ small, the t^* in the optimal mechanism must be large (see Equation (B.49) in Appendix B for the exact expression; it can be seen there that, as $\lambda \to 0$, we have $t^* \to \infty$). Given very large t^*, for small t, it is as if an uninformed agent is in the situation where switching to B can happen only if both report B, in which case his incentive of misreporting B is very strong; for large t (i.e., t much closer to t^*), an uninformed agent

[10] More specifically, $\lambda < \frac{r}{2} \frac{1-\alpha}{1+\alpha}$.

must assign almost probability 0 to $(a_t, b_t) = (0, 0)$, and hence the incentive given the other case $(a_t, b_t) = (0, 1)$ dominates. In either case, he prefers misreporting.

- Observation 3: Therefore, "by continuity", there exists a way to specify $q_t((0, t_1 t_2)_t)$ for each $t_1, t_2 < t^*$ so that all the ICs are to be binding.

This is quite informal, but again, the actual claim and proofs are in the appendix in the discrete-time setting.

Here, we do not go into the verification part, which shows that the above mechanism is indeed a second-best mechanism. Instead, we informally explain how the proof works in the appendix.

Our proof idea is based on a Lagrangian (weak) duality argument. Consider the Lagrangian of our problem that has the IC_t^N of all t in the objective with some non-negative multipliers. It is easy to show that the value of the Lagrangian is always weakly higher than the objective of the original problem given any feasible (i.e., incentive compatible) mechanism. More importantly, it also implies that, if we can find the pair of a non-negative multiplier vector and a mechanism such that (i) that mechanism is optimal in the Lagrangian problem with that multiplier vector, and moreover, (ii) under that mechanism, the value of the Lagrangian coincides with the value of the original objective (which essentially means that the complementary slackness is satisfied), then that mechanism is a second-best mechanism.

Guessing the "right" multiplier vector is relatively straightforward, if we recall the shape of the candidate second-best mechanism: It is associated with some interior probabilities of switching in case two voters report B types. In the Lagrangian problem, this interiority is essentially equivalent to the principal's indifference,[11] which completely determines the possible Lagrange multiplier. Once the multiplier is given, we can also verify that the above mechanism satisfies properties (i) and (ii) above, which completes the proof.

5.3 Second-best mechanism, and its implementation

We close this section by discussing how this mechanism can be implemented from a more realistic perspective. Three components are crucial: (i) Until t^*, the policy stays in A unless there is a strong opposition; (ii) Until t^*, information should be hidden; (iii) Once time t^* arrives, the policy stays in A (forever) if and only if at most one voter reports A type. Although the literal implementation may be "an election every moment in time", it is obviously difficult to implement in practice. Another potential way to implement it may be a combination of "petitions" and "referendum". First, the principal may commit to have a single election ("referendum"), at

[11] Of course, the principal is *not* indifferent in terms of the original objective, though.

the latest at time t^*. By then, the voters can express their opinions ("petitions") via private communication with the principal: If there are enough petitions collected for switching to policy B (two in our case), then the principal holds a referendum after some time ("delay", as specified in the second-best policy). The voting outcome will be to switch to B, so we switch to B. However, if not enough petitions are collected by t^*, then the principal holds the referendum at that moment. Then we stay in A if and only if there is at least one vote for staying.

6 Single election

In this section, we consider a simple benchmark where the planner holds a single election at one date. The planner has the same objective, i.e., maximising the ex-ante expected social payoff. However, he has only one chance to collect information from agents and makes a collective decision based on that information.

Such a scenario can be thought of a "referendum", as referendums are often held once or a very few times. We consider such a model, since elections are not free in real life. It can be very costly to organize an election. Thus, this section studies the scenario when it is only affordable for the planner to hold one election.

6.1 First-best policy

We first consider the first-best benchmark where agents report their information truthfully but the planner holds only one election at a date he commits to.

Given the election date t, the planner's problem is whether to switch to policy B once and for all. It is immediate that the planner optimally switches to policy B if and only if that is myopically optimal, i.e., when $u(a_t, b_t, p) < 0$.

Thus, the planner's problem is reduced to find a date $t_{\mathcal{U}} \in [0, \infty]$[12] that maximises the ex-ante expected social payoff from holding the election and making a myopically optimal decision on the election date t:

$$\mathbb{E}\left[(\alpha\, x_A - (n - x_A))(1 - e^{-rt}\, \mathbb{1}_{\{u(a_t, b_t, p) < 0\}})\right],$$

where x_A is the number of agents, informed or not, for whom $\theta^i = A$.

The benefit of delaying the election is that the planner's information is improved, since it gives the agents more time to learn their types. However, delaying the election is costly if policy B is actually socially optimal. It is clear that it is optimal to have an election at some finite date, which depends on the pre-learning period $\underline{T}$.

[12] The case $t_{\mathcal{U}} = \infty$ corresponds to the principal committing to policy A at date 0, without ever holding an election.

Is the above first-best policy incentive compatible if the agents must report their private types? According to the policy, voting for A (B) increases the expected probability of implementing policy A (B) thereafter. At the election date, types A and B are clearly better off voting for A and B respectively. The violation of IC comes from the uninformed types. If the prior belief $p < p^M$ ($p > p^M$), it is a strictly dominant strategy to vote for B (A), as the uninformed types is also strictly biased to policy B (A). If $p = p^M$, novices are indifferent between reporting A, B or p, so the first-best can be achieved. Thus, the first-best is achievable if and only if $p = p^M$.

6.2 Second-best policy

The IC failure described in the previous section arises at every date $t \geq 0$ at which the election could be held if the prior belief is not p^M. We assume $p < p^M$ in this section.[13]

Let us fix the election date t. A policy $\mathcal{D}^t$ is a mapping $S_n \to [0, 1]$. $\mathcal{D}^t(a, b, m)$ denote the probability that the principal continues with policy A upon having observed a and b reports for A and B types respectively, and $m = n - a - b$ reports for the uninformed type (i.e. reports that $p_t^i = p$), which can be interpreted as "abstentions". Let $\pi_n^t(a, b, m)$ denote the planner's belief (before receiving any reports) that, at date t, among all n agents, a (b) have observed A (B) news, and that the remaining m agents are uninformed, where $(a, b, m) \in S_n$. Thus, the planner chooses $\mathcal{D}^t$ to maximise the expected social payoff

$$\sum_{(a,b,m)\in S_n} \pi_n^t(a, b, m)\, \mathcal{D}^t(a, b, m)\, u(a, b, p), \tag{15}$$

subject to the truth-telling constraints for all agents.

Let $\pi_{n-1}^t(a, b, m)$ denote agent i's belief that, at date t, among the other $n - 1$ agents, a (b) agents have observed A (B) news, and that the remaining m agents are uninformed, where $(a, b, m) \in S_{n-1}$. Because agents are symmetric ex ante, $\pi_{n-1}^t(a, b, m)$ does not depend on the identity of the particular agent. Moreover, it is also the planner's belief.

Suppose all other agents report truthfully, given the policy $\mathcal{D}^t$, the agent with belief $q \in \{0, p, 1\}$ makes a report $\mu \in \{0, p, 1\}$ so as to maximise her expected payoff

$$\sum_{(a,b,m)\in S_{n-1}} \pi_{n-1}^t(a, b, m)\, \mathcal{D}^t(a + \mathbb{1}_{\{\mu=1\}}, b + \mathbb{1}_{\{\mu=0\}}, m + \mathbb{1}_{\{\mu=p\}})\, v(q),$$

where $v(q) = \alpha q - (1 - q)$ is the agent's expected payoff from policy A thereafter if her belief is $q \in \{0, p, 1\}$. A-type agents have a positive payoff $v(1) = \alpha$ from policy A and will choose a report that maximises the expected probability of continuing policy A. B-type and the uninformed

[13]The other case $p > p^M$ is symmetric.

agents have negative payoffs $v(0) = -1$ and $v(p) < 0$ from policy A respectively, and will choose a report that minimises the expected probability of continuing policy A.

The interests of B-type and the uninformed agents are perfectly aligned. Thus, incentive compatibility requires that truthful reporting gives both B-type and the uninformed agents the same expected probability of continuing policy A, that is

$$\sum_{(a,b,m)\in S_{n-1}} \pi_{n-1}^t(a,b,m)\ \left[\mathcal{D}^t(a,b+1,m) - \mathcal{D}^t(a,b,m+1)\right] = 0. \tag{16}$$

The constraint (16) means that the policy $\mathcal{D}^t$ does not distinguish reports from B-type and the uninformed agents in expectation, if all remaining agents report truthfully. It is an *ex ante* constraint at the reporting stage when agents have no information on other agents beyond their prior information. Thus, the expectation is taking over all possible realizations of other agents' types. It does not mean that the policy $\mathcal{D}^t$ does not distinguish reports from B-type and the uninformed agents at the *ex post* stage for a particular realization of a state $(a, b, m) \in S_n$.

However, our next result show that, there actually exists an optimal policy in which the reports from B-type and the uninformed agents are not distinguished even at the ex post stage.

Proposition 3. *Assume $p < p^M$. For a fixed election date t, there exists a solution $\mathcal{D}^t$ to the planner's problem such that $\mathcal{D}^t(a,b,m) = \mathcal{D}^t(a,b',m')$ for any $(a,b,m),(a,b',m') \in S_n$. Moreover, there exists a threshold $\bar{a}^t \in \{1,2,\cdots,\bar{a}+1\}$ that is increasing in t, such that $\mathcal{D}^t(a,b,m) = 1$ if $a \geq \bar{a}^t$ and $\mathcal{D}^t(a,b,m) = 0$ if $a < \bar{a}^t$.*

Incentive compatibility requires that the optimal policy does not distinguish B-types and the uninformed agents ex ante. Optimality further implies that ex post the policy does not need to make such a distinction either. Therefore, the planner only makes efficient use of the information about A-types and disregards the preference intensities of B-types and the uninformed agents. This result is similar to Azrieli and Kim (2014), who show that the optimal policy is a weighted majority rule and only cardinal information needs to be taken into account. Comparing to their static model, our dynamic setting further allows us to examine how the voting rule is determined by the election date.

As the optimal policy depends only on the number of A-type agents, there exists a threshold number $\bar{a}^t$ of A-type agents, above which the optimal policy will continue with policy A. The optimal policy treats the B-type and the uninformed agents equally as agents who are against policy A. Although the belief (on being type A) of an uninformed agent is constant over time, the belief of an average agent who is against policy A decreases over time. This is because conditional on being an agent who is against policy A, she is more likely to be a B-type agent as time passes. This means that the threshold number $\bar{a}^t$ is increasing over time. In terms of a weighted majority rule, A-type agents have decreasing weights over time, while B-type agents and the uninformed agents have the same and increasing weights over time. If the election is held on a later date, B-types and the uninformed agents have more power over A-types in the election, and it is optimal to continue policy A only if more A-votes have been casted.

6.3 First-best and second-best election dates

The second-best optimal policy does not elicit all private information. Does that mean the planner will have an early election, comparing to the first-best benchmark, as less information is aggregated and waiting is costly? This section we use a simplified model to show that this is not true in general.

Consider the following model where the single election has to be held at either $t = 0$ or $t = 1$. The benefit of delaying the election is that the planner will have better information, but the cost is that he may be stuck with the potentially suboptimal policy too long.

In the first-best benchmark, the planner knows the state $s_t \in S_n$ at the election date $t \in \{0, 1\}$, but in the second-best scenario, he only has corser information $\hat{s}_t \in \hat{S}_n := \{(a, a^c) \in \mathbb{N}^2 : a + a^c = n\}$, i.e., the number of agents a_t who are in favor of policy A and the number of agents a_t^c who are against policy A.

For the first-best policy, on the election date $t \in \{0, 1\}$, the planner optimally continues with policy A with probability one if the expected payoff from policy A, conditional on the state s_t, is positive, i.e., $\mathbb{E}[\alpha\, x_A - (n - x_A)|s_t] > 0$; we denote this event as $A^F(t)$. In the complementary event, i.e., $\mathbb{E}[\alpha\, x_A - (n - x_A)|s_t] \le 0$, denote it as $B^F(t)$, the planner optimally switches to policy B with probability one. However, for the second-best policy, the planner only knows $\hat{s}_t \in \hat{S}_n$ on the election date t. He optimally continues with policy A with probability one if conditional on $\hat{s}_t$, it is optimal to do so, i.e., $\mathbb{E}[\alpha\, x_A - (n - x_A)|\hat{s}_t] > 0$; we denote this event as $A^S(t)$. Similarly, in the complementary event, i.e., $\mathbb{E}[\alpha\, x_A - (n - x_A)|\hat{s}_t] \le 0$, denote it as $B^S(t)$, the planner optimally switches to policy B with probability one.

In the first-best benchmark, the expected payoff of having the election at time 0 is

$$\mathbb{E}\left[(\alpha\, x_A - (n - x_A))\, \mathbb{1}_{A^F(0)}\right],$$

and the expected payoff of having the election at time 1 is

$$(1 - \delta)\mathbb{E}\left[(\alpha\, x_A - (n - x_A))\right] + \delta \mathbb{E}\left[(\alpha\, x_A - (n - x_A))\, \mathbb{1}_{A^F(1)}\right],$$

where $\delta = e^{-r}$ is the discounting factor from 0 to 1.

When the policy only uses information about the number of A-type agents, the expected payoffs of having the election at time 0 and at time 1 have similar expressions, with $A^F(0)$ and $A^F(1)$ replaced by $A^S(0)$ and $A^S(1)$.

We show that the planner's optimal election date, in both the first-best and the second-best policies, depends on how patient he is. In both scenarios, the later the election is held, the better the information the planner acquires, which also incurs a higher delaying cost. Thus, a planner will optimally hold the election on date 1 if and only if he is patient enough.

Proposition 4. *Assume $p < p^M$ and the choice set of the election date is $\{0, 1\}$. There exists thresholds $\delta^F, \delta^S \in (0, 1)$ such that the planner's optimal election date is 1 in the first (second) best policy if and only if his discount factor $\delta > \delta^F$ ($\delta > \delta^S$).*

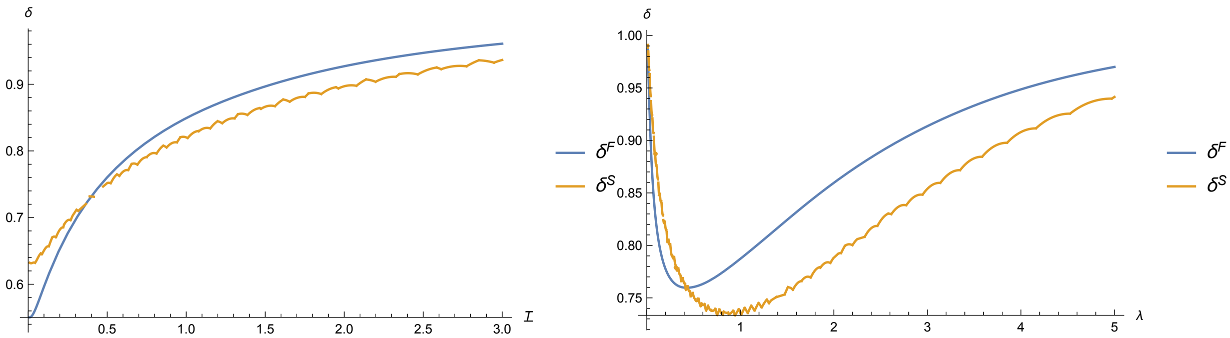
An interesting question is whether the election is held earlier in the second-best policy, comparing to the first-best benchmark. A straightforward argument that supports an early election is that learning obtains less information in the second-best scenario, however, it does not take into account that the information learnt at time 0 (from the pre-learning period) in the second-best scenario is also less than that in the first-best scenario.

Thus, which election should be held earlier in the two scenarios depends on the relative importance of learning, which is evaluated as the incremental value from learning, before time 0, call it "pre-learning", and in between time 0 and 1, call it "post-learning", in each scenario.

When pre-learning dominates, for example when the issue is old or when the learning rate is fast, the planner already has much information by time 0 in both the first-best and the second-best scenarios. However, the planner acquires much more information in the first-best scenario than he does in the second-best scenario. Thus, post-learning, although is less important, enables the planner to acquire more incremental value in the second-best scenario than it does in the first-best scenario. Therefore, the planner's optimal election date in the second-best scenario is later than the optimal election date in the first-best scenario.

When post-learning dominates, e.g. when the issue is new or when the learning rate is slow, the planner does not have much information by time 0 in both the first-best and the second-best scenarios. In this case, the incremental value from post-learning is much higher in the first-best scenario, compared to the second-best scenario. Therefore, the planner's optimal election date in the second-best scenario is earlier than the optimal election date in the first-best scenario.

Figure (3a) and Figure (3b) illustrate the dependence of the thresholds of the discounting factor on the length of the pre-learning period $\underline{T}$ and on the learning rate λ. In both the first-best and the second-best scenarios, a very impatient planner with a small discounting factor will optimally hold the election on date 0, and a very patient planner with a large discounting factor will optimally hold the election on date 1. However, there exists an interval of the discounting factors between the two curves such that the optimal election dates are different in the two scenarios. When the issue is new (a small $\underline{T}$) or the learning rate is slow (a small λ), there exists an interval of the discounting factors between which the optimal election date in the first-best scenario is



(a) $\lambda = 0.5$ (b) $\underline{T} = 0.5$

Figure 3: The thresholds of the discounting factor: $n = 100, \alpha = 1, p = 0.49$

date 1 and the optimal election date in the second-best scenario is date 0. Thus, the insufficient information elicitation due to the incentive problem makes the planner hold an earlier election. However, when the issue is old (a large $\underline{T}$) or the learning rate is fast (a large λ), there exists an interval of the discounting factors between which the optimal election date in the first-best scenario is date 0 and the optimal election date in the second-best scenario is date 1. In this case, the insufficient information elicitation due to the incentive problem delays the election.

7 Conclusion and open questions

We consider the problem of a planner seeking a welfare-maximising social choice rule in a setting where agents have heterogeneous preferences, and where their private information regarding their preferences accrues gradually, over time. We have shown that, regardless whether the principal solicits reports continually, or once, the first-best mechanism is not incentive-compatible. We have also provided a set of tools, including deadlines, delays and private communication, to restore the incentive compatibility in an efficient way.

A number of open questions remain and are the object of ongoing work. First, what is the second-best mechanism in continuous time when there are more than three voters? Additional difficulties arise as not only pivotal agents have incentives to misreport. Thus, agents' incentive constraints are linked in non-trivial ways.

Second, private communication in general also affects an agent' belief about the history of reports in a non-trivial way. This is, on one hand, interesting as it may further help reduce the incentive issue, but on the other hand, challenging for the analysis as the belief about the history affects the agents' incentive and the optimal design of the mechanism, which in turn determines that belief.

A Appendix

A.1 Deriving the expression (4) for V^A

The above expression for V^A can be derived as follows. Suppose that k of the uninformed agents are in fact A-types, and the remaining $m-k$ are B-types. The probability of this event is $\binom{m}{k}p^k(1-p)^{m-k}$. In this case, the joint expected payoff is

$$\text{(A.17)} \quad \int_0^\infty e^{-(k\lambda_A+(m-k)\lambda_B)t}\Big[k\lambda_A\left[(1-e^{-rt})\left[\alpha(a+k)-(b+m-k)\right]+e^{-rt}V(a+1,b,p)\right] \\ +(m-k)\lambda_B\left[(1-e^{-rt})\left[\alpha(a+k)-(b+m-k)\right]+e^{-rt}V(a,b+1,p)\right]\Big]dt,$$

which can be understood as follows. If the first signal comes from one of the k uninformed A-types at date t, then the number of declared A-types increases by 1 and the number of uninformed agents decreases by one. The policy A is implemented until t, resulting in a payoff of 1 for all $a+k$ A-types, regardless of whether they have yet learnt their type. The payoff in case the first signal comes from on of the $m-k$ uninformed B-types (second line of (A.17)) is derived in the same way. Rearranging,

$$\int_0^\infty e^{-(k\lambda_A+(m-k)\lambda_B)t}\Big[\left[k\lambda_A+(m-k)\lambda_B\right](1-e^{-rt})\left[\alpha(a+k)-(b+m-k)\right] \\ +e^{-rt}\left[k\lambda_A\,V(a+1,b,p)+(m-k)\lambda_B\,V(a,b+1,p)\right]\Big]dt,$$

and using $\lambda_A=\lambda_B$ gives

$$\int_0^\infty e^{-m\lambda t}\Big[m\lambda\,(1-e^{-rt})\left[\alpha(a+k)-(b+m-k)\right]+e^{-rt}\lambda\left[k\,V(a+1,b,p)+(m-k)\,V(a,b+1,p)\right]\Big]dt,$$

which yields

$$\frac{r}{r+m\lambda}\left[\alpha(a+k)-(b+m-k)\right]+\frac{\lambda}{r+m\lambda}\left[k\,V(a+1,b,p)+(m-k)\,V(a,b+1,p)\right]$$

Thus, we have that

$$\begin{aligned}
V^A(a,b,m,p) &= \sum_{k=0}^m\left\{\binom{m}{k}p^k(1-p)^{m-k}\left[\frac{r}{r+m\lambda}\left[\alpha(a+k)-(b+m-k)\right]\right.\right. \\
&\qquad\left.\left.+\frac{\lambda}{r+m\lambda}\left[k\,V(a+1,b,p)+(m-k)\,V(a,b+1,p)\right]\right]\right\} \\
&= \frac{r}{r+m\lambda}\sum_{k=0}^m\binom{m}{k}p^k(1-p)^{m-k}\left[\alpha(a+k)-(b+m-k)\right] \\
&\quad+\frac{\lambda}{r+m\lambda}\,V(a+1,b,p)\sum_{k=0}^m\binom{m}{k}p^k(1-p)^{m-k}k \\
&\quad+\frac{\lambda}{r+m\lambda}\,V(a,b+1,p)\sum_{k=0}^m\binom{m}{k}p^k(1-p)^{m-k}(m-k) \\
&= \frac{r}{r+m\lambda}\left[\alpha(a+mp)-(b+m(1-p))\right] \\
&\quad+\frac{\lambda}{r+m\lambda}\,V(a+1,b,p)\,mp \\
&\quad+\frac{\lambda}{r+m\lambda}\,V(a,b+1,p)\,m(1-p).
\end{aligned}$$

Simplifying the latter yields (4).

A.2 Proof of Lemma 1

We prove this result by induction. Fix k such that $k+1 \le n$ and consider pairs $(\tilde{a}, \tilde{b})$ such that $\tilde{a} + \tilde{b} + 1 = k + 1$. The induction hypothesis is

$$V^A(\tilde{a}+1, \tilde{b}, p) > V^A(\tilde{a}, \tilde{b}+1, p). \tag{A.18}$$

Now consider pairs (a, b) such that $a + b + 1 = k$. Using (4) yields

$$\begin{aligned} V^A(a+1,b,p) - V^A(a,b+1,p) = (1-g_{m-1})\,[\alpha+1] \\ + g_{m-1}\Big[[p\,V(a+2,b,p) + (1-p)\,V(a+1,b+1,p)] \\ - [p\,V(a+1,b+1,p) + (1-p)\,V(a,b+2,p)] \Big] \end{aligned}$$

Since $a+b+2 = k+1$, the induction hypothesis (A.18) implies that $V(a+2,b,p) \ge V(a+1,b+1,p)$ so that

$$p\,V(a+2,b,p) + (1-p)\,V(a+1,b+1,p) \ge V(a+1,b+1,p);$$

the induction hypothesis also implies that $V(a,b+2,p) \le V(a+1,b+1,p)$ so that

$$p\,V(a+1,b+1,p) + (1-p)\,V(a,b+2,p) \le V(a+1,b+1,p).$$

It follows that

$$V^A(a+1,b,p) - V^A(a,b+1,p) \ge (1-g_{m-1})\,[\alpha+1] > 0, \tag{A.19}$$

completing the induction step.

Finally, consider pairs $(\tilde{a}, \tilde{b})$ such that $\tilde{a} + \tilde{b} + 1 = n$. By (3) we have that $V(\tilde{a}+1, \tilde{b}, p) = (\alpha(\tilde{a}+1) - \tilde{b}) \vee 0$ and $V(\tilde{a}, \tilde{b}+1, p) = (\alpha\tilde{a} - (\tilde{b}+1)) \vee 0$. Since $\alpha(\tilde{a}+1) - \tilde{b} - (\alpha\tilde{a} - (\tilde{b}+1)) = \alpha + 1 > 0$, it follows that $V(\tilde{a}+1, \tilde{b}, p) \ge V(\tilde{a}, \tilde{b}+1, p)$ for every $p \in [0,1]$, establishing the induction hypothesis (A.18) for $k+1 = n$.

A.3 Proof of Lemma 2

Proof. We prove the lemma by induction. Fix k such that $k+1 < n$ and consider pairs $(\tilde{a}, \tilde{b})$ such that $\tilde{a} + \tilde{b} = k+1$. The induction hypothesis is that $V^A(\tilde{a}, \tilde{b}, p)$ is continuous in p and differentiable almost everywhere on $[0,1]$, with

$$\frac{d}{dp} V^A(\tilde{a}, \tilde{b}, p) > 0. \tag{A.20}$$

Now consider pairs (a, b) such that $a + b = k$. Observe that $\frac{d}{dp} u(a,b,p) = m(\alpha+1) > 0$. We distinguish three cases, according to whether continuation values are positive. Recall that, by Lemma 1, $V^A(a+1,b,p) > V^A(a,b+1,p)$ for every $p \in [0,1]$.

Case 1: p is such that $V^A(a+1,b,p) \le 0$. Lemma 1 implies that $V^A(a,b+1,p) < 0$. By (4), we then have that $V^A(a,b,p) = (1-g_m)\,u(a,b,p)$, which is continuous, differentiable and strictly increasing at every $p \in [0,1]$, completing the induction step.

Case 2: p is such that $V^A(a+1,b,p) > 0$ but such that $V^A(a,b+1,p) \le 0$. Then (4) becomes

$$V^A(a,b,p) = (1-g_m)\,u(a,b,p) + g_m\,p\,V^A(a+1,b,p). \tag{A.21}$$

Differentiating the second term with respect to p, we find that

$$\frac{d}{dp}\Big(p\,V^A(a+1,b,p)\Big) = V^A(a+1,b,p) + p\,\frac{d}{dp} V^A(a+1,b,p).$$

On the right-hand side, the first term is positive by assumption while the second term is positive under the induction hypothesis. Equation (A.21) therefore implies that $V^A(a,b,p)$ is a convex combination of two functions that are continuous and increasing at every $p \in [0,1]$, and differentiable at almost every $p \in [0,1]$, completing the induction step.

Case 3: p is such that $V^A(a,b+1,p) > 0$. Lemma 1 implies that $V^A(a+1,b,p) > 0$. Then (4) becomes

$$V^A(a,b,p) = (1-g_m)\,u(a,b,p) + g_m\left[p\,V^A(a+1,b,p) + (1-p)\,V^A(a,b+1,p)\right]. \tag{A.22}$$

Differentiating the second term with respect to p, we find that

$$\begin{aligned}\frac{d}{dp}\Big(p\,V^A(a+1,b,p) + (1-p)\,V^A(a,b+1,p)\Big) &= V^A(a+1,b,p) - V^A(a,b+1,p)\\ &\quad + p\,\frac{d}{dp}V^A(a+1,b,p) + (1-p)\,\frac{d}{dp}V^A(a,b+1,p).\end{aligned}$$

On the right-hand side, the first line is positive by Lemma 1 while the second line is positive under the induction hypothesis. Equation (A.23) therefore implies that $V^A(a,b,p)$ is a convex combination of two functions that are continuous and increasing at every $p \in [0,1]$, and differentiable at almost every $p \in [0,1]$, completing the induction step.

Finally, consider pairs $(\tilde{a},\tilde{b})$ such that $\tilde{a}+\tilde{b} = n-1$. By (3) we have that $V(\tilde{a}+1,\tilde{b},p) = (\alpha(\tilde{a}+1)-\tilde{b}) \vee 0$ and $V(\tilde{a},\tilde{b}+1,p) = (\alpha\tilde{a}-(\tilde{b}+1)) \vee 0$, both constant and continuous in p. From (4) it is immediate that $V^A(\tilde{a},\tilde{b},p)$ is continuous and differentiable almost everywhere in $[0,1]$. Differentiating gives

$$\frac{d}{dp}V^A(\tilde{a},\tilde{b},p) = (1-g_m)\,\frac{d}{dp}u(\tilde{a},\tilde{b},p) + g_m\left[V(\tilde{a}+1,\tilde{b},p) - V(\tilde{a},\tilde{b}+1,p)\right] > 0, \tag{A.23}$$

where the inequality follows from $V(\tilde{a}+1,\tilde{b},p) \geq V(\tilde{a},\tilde{b}+1,p)$ and $\frac{d}{dp}u(a,b,p) = m(\alpha+1) > 0$. This establishes the induction hypothesis for $k+1 = n-1$. □

A.4 Proof of Proposition 1

First, observe that, for every pair (a,b) such that $a+b<n$, by Lemma 2, $V^A(a,b,p)$ crosses zero at most once on $(0,1)$. Consequently, there exists a unique $p^*(a,b) \in [0,1]$ satisfying $V^A(a,b,p) \geq 0$ if and only if $p \geq p^*(a,b)$. We now show that the statements (i) and (ii) hold.

(i) Assume by way of contradiction that $p^*(a,b) \leq p^*(a+1,b)$. Then $V^A(a+1,b,p^*(a,b)) \leq 0$. Moreover, Lemma (1) gives $V^A(a,b+1,p^*(a,b)) < V^A(a+1,b,p^*(a,b))$, implying that $V^A(a,b+1,p^*(a,b)) < 0$. Consequently, (4) gives $V^A(a,b,p^*(a,b)) = (1-g_m)\,u(a,b,p^*(a,b))$, so that $V^A(a,b,p^*(a,b)) = 0$ if and only if $p^*(a,b) = \frac{b+m-a\alpha}{m(1+\alpha)} = \frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)}$, where the last expression is obtained by substituting $m = n-a-b$.

Suppose first that $\frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)} \leq 0$ or, equivalently, that $\alpha a - (n-a) \geq 0$. Observe that the left hand-side in the last inequality is $u(a,b,0)$. Since $u(a,b,p)$ is strictly increasing in p, we therefore have that $V^A(a,b,p) > 0$ for every $p \in (0,1)$. A contradiction.

Suppose next that $\frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)} \geq 1$ or, equivalently, $(n-b)\alpha - b \leq 0$. Observe that the left hand-side in the last inequality is $u(a,b,1)$. Since $u(a,b,p)$ is strictly increasing in p, we therefore have that $V^A(a,b,p) < 0$ for every $p \in (0,1)$. A contradiction.

Finally, suppose that $\frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)} \in (0,1)$ or, equivalently, that $n-a(1+\alpha) > 0$ and $n\alpha - b(1+\alpha) > 0$. Then

$$p^*(a,b) - p^*(a+1,b) = \frac{n\alpha - b(1+\alpha)}{(n-a-b)(n-(a+1)-b)(1+\alpha)} > 0,$$

a contradiction.

(ii) Assume by way of contradiction that $p^*(a,b) \geq p^*(a,b+1)$. Together with (i), this implies that if A is optimal in state (a,b,p), it remains optimal in all following states. Consequently, the payoff from A in state (a,b,p) is $u(a,b,p)$, and we have that $V^A(a,b,p^*(a,b)) = 0$ if and only if $p^*(a,b) = \frac{b+m-a\alpha}{m(1+\alpha)} = \frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)}$. For parameters such that $\frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)} \notin (0,1)$, the arguments from (i) establish the contradiction.

Now suppose that $\frac{n-a(1+\alpha)}{(n-a-b)(1+\alpha)} \in (0,1)$ or, equivalently, that $n - a(1+\alpha) > 0$ and $n\alpha - b(1+\alpha) > 0$. Then

$$p^*(a,b) - p^*(a,b+1) = \frac{a(1+\alpha) - n}{(n-a-b)(n-(a+1)-b)(1+\alpha)} < 0,$$

a contradiction.

A.5 Proof of Lemma 3

(i) Switching to B immediately and never switching to B are both feasible policies for the planner. Therefore, $V(a,b,p) \geq u(a,b,p) \vee 0$. For $a > n/(1+\alpha)$ we have $u(a,b,p) > 0$, establishing the claim.

(ii) We prove this claim by induction. Fix k such that $k+1 \leq n$ and consider pairs $(\tilde{a},\tilde{b})$ such that $\tilde{a}+\tilde{b} = k+1$. The induction hypothesis is that $V^A(\tilde{a},\tilde{b},p) < 0$ for every $p \in [0,1]$. Now consider pairs (a,b) such that $a+b = k$. Under the induction hypothesis, $V(a+1,b,p) = V(a,b+1,p) = 0$. Thus, (4) gives that $V^A(a,b,p) = (1-g_m)\,u(a,b,p)$. For $b > \alpha n/(1+\alpha)$ we have $u(a,b,p) < 0$, completing the induction step. Finally, consider pairs $(\tilde{a},\tilde{b})$ such that $\tilde{a}+\tilde{b} = n$. By (4), we have that $V^A(a,b,p) = u(a,b,p) < 0$, where the inequality follows from $b > \alpha n/(1+\alpha)$. This establishes the induction hypothesis for $k+1 = n$.

A.6 Claim in Section 3.3

The next lemma verifies that, indeed, S_n^{***} corresponds to the diagonal $\{(a,b)|b = a-1 \leq \bar{a}-1\}$. We verify that, under our assumptions, we indeed have that $V^A(a,b,m,p^\star) > 0$ whenever $(a,b) \in S_n^{***}$ or, equivalently, whenever $b = a-1$, $a \in \{1,\dots,\bar{a}\}$, and that $V^A(a,a,m,p^\star) \leq 0$, $a \in \{0,\dots,\bar{a}\}$.

Lemma A.10. *Fix λ, r, and let $n \leq \bar{n}(\lambda,r)$. Assume that $S_n^{***} = \{(a,b)|b = a-1 \leq \bar{a}-1\}$. Then, (i) $V^A(a,a-1,n-2a+1,p^\star) > 0$ for every $a \in \{1,\dots,\bar{a}\}$, and (ii) $V^A(a,a,n-2a,p^\star) \leq 0$ for every $a \in \{0,\dots,\bar{a}\}$.*

Proof. (i) Fix λ, r, and let $n \leq \bar{n}(\lambda,r)$. Observe that,

$$\begin{aligned}
& n \leq \bar{n}(\lambda,r) \\
\Leftrightarrow\quad & n \leq 1 - \frac{1}{v^M(p^\star)} \\
\Leftrightarrow\quad & 1 + (n-1)v^M(p^\star) \geq 0 \\
\Leftrightarrow\quad & u(1,0,n-1,p^\star) \geq 0
\end{aligned}$$

where the first equivalence follows from (5), the second uses $v^M(p^\star) < 0$, and the third just applies the definition of $u(1,0,n-1,p^\star)$. Moreover, observe that for every $a \in \{1,\dots,\bar{a}-1\}$,

$$u(a,a-1,n-2a+1,p^\star) = 1 + (n-2a+1)v^M(p^\star) \geq 1 + (n-1)v^M(p^\star) = u(1,0,n-1,p^\star),$$

where the inequality follows from $v^M(p^\star) < 0$. Consequently we have that

$$n \leq \bar{n}(\lambda,r) \quad \Rightarrow \quad u(a,a-1,n-2a+1,p^\star) \geq 0, \forall a \in \{1,\dots,\bar{a}-1\}. \tag{A.24}$$

We now verify that $V^A(a, a-1, n-2a+1, p^\star) > 0$, $a \in \{1, \dots, \bar{a}-1\}$. By (6) we have that

$$\begin{aligned} V^A(a, a-1, n-2a+1, p) &= u(a, a-1, n-2a+1, p) \\ &\quad - \sum_{k=a}^{\bar{a}} \binom{2k-2a}{k-a} \frac{1}{k-a+1}\, p^{k-a}(1-p)^{k-a+1} \left(\prod_{j=n-2k+1}^{n-2a+1} g_j \right) u(k, k, n-2k, p). \end{aligned}$$

For $p < p^M$ we have that $u(k, k, n-2k, p) = (n-2k)v^M(p) < 0$. Consequently, $V^A(a, a-1, n-2a+1, p) > u(a, a-1, n-2a+1, p)$ for every $p < p^M$. In particular,

$$V^A(a, a-1, n-2a+1, p^\star) > u(a, a-1, n-2a+1, p^\star) \geq 0,$$

where the second inequality follows from $n \leq \bar{n}(\lambda, r)$, as shown in (A.24). This establishes the claim.

(ii) Next, we verify that $V^A(a, a, n-2a, p^\star) \leq 0$ for every $a \in \{0, \dots, \bar{a}\}$. Given that $S_n^{***} = \{(a, b) | b = a-1 \leq \bar{a}-1\}$, we have that $(a, a+1) \notin S^{**}$ and that, consequently, $V(a, a+1, n-2a-1, p) < 0$. Hence, the recursion (4) for V^A under the planner solution gives that

$$V^A(a, a, n-2a, p) = (1-g_{n-2a})\, u(a, a, n-2a, p) + g_{n-2a}\, p\, V^A(a+1, a, n-2a-1, p).$$

Consequently, we have that $V^A(a, a, n-2a, p) \leq 0$ if and only if

$$\begin{aligned} & V^A(a+1, a, n-2a-1, p) \leq -\frac{1-g_{n-2a}}{g_{n-2a}p} u(a, a, n-2a, p) \\ \Leftrightarrow\quad & V^A(a+1, a, n-2a-1, p) \leq -\frac{1-g_{n-2a}}{g_{n-2a}} \frac{(n-2a)\, v^M(p)}{p} \\ \Leftrightarrow\quad & V^A(a+1, a, n-2a-1, p) \leq -\frac{r}{(n-2a)\lambda} \frac{(n-2a)\,(2p-1)}{p} \\ \Leftrightarrow\quad & V^A(a+1, a, n-2a-1, p) \leq \frac{r}{\lambda} \frac{1-2p}{p} \end{aligned} \tag{A.25}$$

By Lemma 2, $V^A(a, b, m, p)$ is a weakly increasing function of p for every $(a, b) \in S_n^{**}$. Using (6), we have that

$$V^A(a+1, a, n-2a-1, 0) = u(a+1, a, n-2a-1, 0) = 1 + (n-2a-1)(-1) = -(n-2a) < 0,$$

where the inequality holds for every $a \in \{0, \dots, \bar{a}\}$, and that

$$\begin{aligned} V^A(a+1, a, n-2a-1, p^M) &= 1 + (n-2a-1)v^M(p^M) \\ &\quad - \sum_{k=a+1}^{\bar{a}} \binom{2k-2a-2}{k-a-1} \frac{1}{k-a}\, p^{k-a-1}(1-p)^{k-a} \left(\prod_{j=n-2k+1}^{n-2a-1} g_j \right) (n-2k)v^M(p^M) \\ &= 1. \end{aligned}$$

By the intermediate value theorem, we therefore have that $V^A(a+1, a, n-2a-1, p^\star) \leq 1$. Conversely, the right hand side of (A.25) equals 1 when $p = p^\star$. Consequently, (A.25) holds when $p = p^\star$, establishing the claim. $\square$

A.7 Proof of Lemma 4

Proof. In this proof, we use the shorthand (a, b) for state (a, b, m). We verify that the expression in (6) satisfies the recursion (4) for the value function. We first consider states $(a, b) \in S_n^{***}$. Here, we have $V(a, b+1) = 0$ and $V(a+1, b) = V^A(a+1, b)$. Therefore, (4) gives

$$V^A(a, b) = (1-g_m)\, u(a, b) + g_m\, p\, V^A(a+1, b).$$

Substituting (6) for $V^A(a+1,b)$,

$$V^A(a,b) = (1-g_m)\,u(a,b) + g_m\,p\,u(a+1,b) - g_m\,p \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\,p^{k-a-1}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{m-1} g_j\right) u(k,k).$$

Next, we use $u(a,b) = pu(a+1,b) + (1-p)u(a,b+1)$ and the fact that $a = b+1$ for every $(a,b) \in S_n^{***}$ to substitute for $u(a+1,b)$. We then distribute $g_m\,p$ into the sum. Thus,

$$V^A(a,b) = u(a,b) - g_m\,(1-p)\,u(a,a) - \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\,p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{m} g_j\right) u(k,k).$$

Next, observe that for $(a,b) \in S_n^{***}$ we have $a = b+1$. Therefore,

$$\frac{(a+1-b)(k-a)}{(k-b)(2k-a-b-1)} = \frac{2(k-a)}{(k-b)(2k-2a)} = \frac{1}{k-b} = \frac{a-b}{k-b},$$

where we substituted $a = b+1$ at the first and last equality. Consequently,

$$\begin{aligned} C(a+1,b,k) &= \frac{(2k-a-b-2)!}{(k-a-1)!(k-b-1)!}\,\frac{a+1-b}{k-b} \\ &= \frac{(2k-a-b-1)!}{(k-a)!(k-b-1)!}\,\frac{(a+1-b)(k-a)}{(k-b)(2k-a-b-1)} \\ &= \binom{2k-a-b-1}{k-a}\frac{a-b}{k-b} \\ &= C(a,b,k). \end{aligned} \tag{A.26}$$

The intuition for this result is that, for each state $(a,b) \in S_n^{***}$, the first signal being B results in the planner switching to policy B. Consequently, every path from (a,b) to (k,k), $k \in \{a,\ldots,\bar{a}\}$, that only visits states in the continuation region S_n^{**} must start with an A signal, and therefore necessarily visits the state $(a+1,b)$. It follows that the number of paths from (a,b) to (k,k) equals the number of paths from $(a+1,b)$ to (k,k).

Thus,

$$V^A(a,b) = u(a,b) - g_m\,(1-p)\,u(a,a) - \sum_{k=a+1}^{\bar{a}} C(a,b,k)\,p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{m} g_j\right) u(k,k).$$

Finally, evaluating the summand at the second line for $k=a$ gives

$$(1-p)^{k-b}\left(\prod_{j=n-2a+1}^{m} g_j\right) u(a,a) = g_m\,(1-p)\,u(a,a),$$

where we substituted $a = b+1$ to get $m = n-a-b = n-2a+1$. We therefore obtain the expression in (6) for $V^A(a,b)$. This verifies that the expression in (6) for $V^A(a,b)$ satisfies the recursion in (4) in states $(a,b) \in S_n^{***}$.

Next, we consider states $(a,b) \in S_n^{**} \backslash S_n^{***}$. Here, we have $V(a,b+1) = V^A(a,b+1)$ and $V(a+1,b) = V^A(a+1,b)$. Therefore, (4) gives

$$V^A(a,b) = (1-g_m)\,u(a,b) + g_m\left[p\,V^A(a+1,b) + (1-p)\,V^A(a,b+1)\right].$$

Substituting (6) for $V^A(a+1,b)$ and $V^A(a,b+1)$,

$$\begin{aligned} V^A(a,b) = {} & (1-g_m)\,u(a,b) + g_m\,\left(p\,u(a+1,b) + (1-p)u(a,b+1)\right) \\ & - g_m\,p \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\,p^{k-a-1}(1-p)^{k-b}\left(\prod_{j=n-2k+1}^{m-1} g_j\right) u(k,k) \\ & - g_m\,(1-p) \sum_{k=a}^{\bar{a}} C(a,b+1,k)\,p^{k-a}(1-p)^{k-b-1}\left(\prod_{j=n-2k+1}^{m-1} g_j\right) u(k,k). \end{aligned}$$

We use $u(a,b) = pu(a+1,b) + (1-p)u(a,b+1)$ at the first line, and distribute the terms $g_m\, p$ and $g_m(1-p)$ into the sums to get

$$V^A(a,b) = u(a,b) - \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\, p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{n-a-b} g_j\right) u(k,k)$$
$$- \sum_{k=a}^{\bar{a}} C(a,b+1,k)\, p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{n-a-b} g_j\right) u(k,k).$$

Observe that

$$\begin{aligned}
C(a+1,b,k) + C(a,b+1,k) &= \binom{2k-a-b-2}{k-a-1}\frac{a+1-b}{k-b} + \binom{2k-a-b-2}{k-a}\frac{a-b-1}{k-b-1} \\
&= \frac{(2k-a-b-2)!}{(k-a-1)!(k-b-1)!}\frac{a+1-b}{k-b} + \frac{(2k-a-b-2)!}{(k-a)!(k-b-2)!}\frac{a-b-1}{k-b-1} \\
&= \frac{(2k-a-b-2)!}{(k-a-1)!(k-b-1)!}\left(\frac{a+1-b}{k-b} + \frac{a-b-1}{k-a}\right) \\
&= \frac{(2k-a-b-2)!}{(k-a-1)!(k-b-1)!}\frac{(a-b)(2k-a-b-1)}{(k-a)(k-b)} \\
&= \binom{2k-a-b-1}{k-a}\frac{a-b}{k-b} \\
&= C(a,b,k). \qquad \text{(A.27)}
\end{aligned}$$

Therefore,

$$V^A(a,b) = u(a,b) \; - \sum_{k=a+1}^{\bar{a}} C(a,b,k)\, p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{n-a-b} g_j\right) u(k,k)$$
$$- \left[C(a,b+1,k)\, p^{k-a}(1-p)^{k-b} \left(\prod_{j=n-2k+1}^{n-a-b} g_j\right) u(k,k)\right]_{k=a},$$

where (7) gives $C(a,b+1,a) = 1 = C(a,b,a)$. Hence, the expression above simplifies to the expression in (6) for $V^A(a,b)$. This verifies that the expression in (6) for $V^A(a,b)$ satisfies the recursion in (4) in states $(a,b) \in S_n^{**} \setminus S_n^{***}$. $\square$

A.8 Proof of Lemma 6

By Lemma (5), $p^*(\bar{a},\bar{b}) < p^\star$ if and only if $\alpha\bar{a} - \bar{b} > 0$. First, consider $\alpha < \frac{1}{n-1}$. Here, α is so small that policy B is optimal unless all agents prefer A ($\bar{a} = n-1$) or, equivalently, policy B is optimal if there is at least one agent who prefers B ($\bar{b} = 0$). Since $\alpha\bar{a} - \bar{b} = \alpha(n-1) > 0$, we have that $p^*(\bar{a},\bar{b}) < p^\star$, establishing the "if" part of the claim for $k = n-1$. Conversely, suppose that $\alpha > n-1$. Here, α is so large that policy A is optimal if at least one agent prefers A ($\bar{a} = 0$) or, equivalently, policy A is optimal unless all agents prefer B ($\bar{b} = n-1$) Since $\alpha\bar{a} - \bar{b} = -(n-1) < 0$, we have that $p^*(\bar{a},\bar{b}) > p^\star$, establishing the "only if" claim for $\alpha > n-1$. Finally, fix an integer $k \in \{1,2,\ldots,n-1\}$ and let $\alpha \in \left(\frac{n-(k+1)}{k+1}, \frac{n-k}{k}\right)$. Then $\bar{a} = k$ and $\bar{b} = n-(k-1)$. Thus, $\alpha\bar{a} - \bar{b} > 0$ if and only if $\alpha > \frac{n-(k+1)}{k}$.

A.9 Proof of Lemma 7

Proof. To prove this result, we first derive the recursion satisfied by $\mathrm{u}^{\theta^i}(a,b,m,p)$. We then verify that the expression in (9) satisfies this recursion. The first part of the proof follows the same steps as the derivation of the recursion (4) for V^A in the planner problem. The second part of the proof follows the same steps as the proof of Lemma 6.

Suppose that k of the novices are in fact A-types, and the remaining $m-k$ are B-types. The probability of this event is $\binom{m}{k}p^k(1-p)^{m-k}$. In this case, agent i's expected payoff is

$$\int_0^\infty e^{-m\lambda t}\Big[k\lambda\Big[(1-e^{-rt})v(\theta^i)+e^{-rt}\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)\Big] \\ +(m-k)\lambda\Big[(1-e^{-rt})v(\theta^i)+e^{-rt}\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\Big]\Big]dt,$$

which can be understood as follows. If the first piece of news comes from one of the k uninformed A-types at date t, then the number of declared A-types increases by 1 and the number of novices decreases by one. The policy A is implemented until t, resulting in a flow payoff of $v(\theta^i)$. The payoff in case the first piece of news comes from on of the $m-k$ uninformed B-types is derived in the same way. Rearranging gives

$$\int_0^\infty e^{-m\lambda t}\Big[m\lambda\,(1-e^{-rt})v(\theta^i)+e^{-rt}\lambda\Big[k\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)+(m-k)\,\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\Big]\Big]dt,$$

which yields

$$\frac{r}{r+m\lambda}v(\theta^i)+\frac{\lambda}{r+m\lambda}\Big[k\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)+(m-k)\,\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\Big]$$

Thus, we have that

$$\begin{aligned}
\mathfrak{u}^{\theta^i}(a,b,m,p) &= \sum_{k=0}^{m}\Bigg\{\binom{m}{k}p^k(1-p)^{m-k}\Bigg[\frac{r}{r+m\lambda}v(\theta^i) \\
&\quad +\frac{\lambda}{r+m\lambda}\Big[k\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)+(m-k)\,\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\Big]\Bigg]\Bigg\} \\
&= \frac{r}{r+m\lambda}\sum_{k=0}^{m}\binom{m}{k}p^k(1-p)^{m-k}v(\theta^i) \\
&\quad +\frac{\lambda}{r+m\lambda}\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)\sum_{k=0}^{m}\binom{m}{k}p^k(1-p)^{m-k}k \\
&\quad +\frac{\lambda}{r+m\lambda}\,\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\sum_{k=0}^{m}\binom{m}{k}p^k(1-p)^{m-k}(m-k) \\
&= \frac{r}{r+m\lambda}\,v(\theta^i) \\
&\quad +\frac{\lambda}{r+m\lambda}\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)\,mp \\
&\quad +\frac{\lambda}{r+m\lambda}\,\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\,m(1-p).
\end{aligned}$$

Simplifying the latter yields the following recursion for $\mathfrak{u}^{\theta^i}$:

(A.28) $$\mathfrak{u}^{\theta^i}(a,b,m,p)=(1-g_m)\,v(\theta^i)+g_m\Big[p\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)+(1-p)\,\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)\Big].$$

The next step consists in verifying that the expression in (9) satisfies the recursion above. We first consider states $(a,b)\in S_n^{***}$. Here, we have $\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)=0$. Therefore, (A.28) gives

$$\mathfrak{u}^{\theta^i}(a,b,m,p)=(1-g_m)\,v(\theta^i)+g_m\,p\,\mathfrak{u}^{\theta^i}(a+1,b,m-1,p).$$

Substituting (9) for $\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)$ gives

$$\mathfrak{u}^{\theta^i}(a,b,m,p)=(1-g_m)\,v(\theta^i)+g_m\,p\,v(\theta^i)-g_m\,p\,v(\theta^i)\sum_{k=a+1}^{\bar a}C(a+1,b,k)\,p^{k-a-1}(1-p)^{k-b}\prod_{j=n-2k+1}^{m-1}g_j.$$

Factoring $v(\theta^i)$ and distributing $g_m\, p$ into the sum gives

$$\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\left[1 - g_m(1-p) - \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\, p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j\right].$$

Since $(a,b) \in S_n^{***}$, we use Using (A.26) to substitute for C(a+1,b,k) and obtain

$$\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\left[1 - g_m(1-p) - \sum_{k=a+1}^{\bar{a}} C(a,b,k)\, p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j\right] \tag{A.29}$$

Finally, evaluating the summand for $k = a$ gives

$$C(a,b,a)\,(1-p)^{a-b} \prod_{j=n-2a+1}^{m} g_j = g_m\,(1-p)$$

where (7) gives $C(a,b,a) = 1$ and we substitute $a = b+1$ to get $m = n-a-b = n-2a+1$. We conclude that the expression in (A.29) is equivalent to the expression in (9). This verifies that the expression in (9) for $\mathfrak{u}^{\theta^i}(a,b,n,p)$ satisfies the recursion in (A.28) in states $(a,b) \in S_n^{***}$.

Next, we consider states $(a,b) \in S_n^{**} \setminus S_n^{***}$. Here, we have that $\mathfrak{u}^{\theta^i}(a+1,b,m-1,p)$ and $\mathfrak{u}^{\theta^i}(a,b+1,m-1,p)$ are given by (9). Substituting in (A.28) gives

$$\begin{aligned}\mathfrak{u}^{\theta^i}(a,b,m,p) &= (1-g_m)\,v(\theta^i) + g_m\left[p\,v(\theta^i) + (1-p)\,v(\theta^i)\right]\\ &\quad - g_m\,p\,v(\theta^i) \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\, p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m-1} g_j\\ &\quad - g_m\,(1-p)\,v(\theta^i) \sum_{k=a}^{\bar{a}} C(a,b+1,k)\, p^{k-a}(1-p)^{k-b-1} \prod_{j=n-2k+1}^{m-1} g_j.\end{aligned}$$

Factoring $v(\theta^i)$ and distributing $g_m\, p$ and $g_m\,(1-p)$ into the sum gives

$$\begin{aligned}\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\Bigg[1 &- \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\, p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j\\ &- \sum_{k=a}^{\bar{a}} C(a,b+1,k)\, p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j\Bigg].\end{aligned}$$

Using (A.27) gives that

$$\begin{aligned}\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\Bigg[1 &- \sum_{k=a+1}^{\bar{a}} C(a,b,k)\, p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j\\ &- \left[C(a,b+1,k)\, p^{k-a}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j\right]_{k=a}\Bigg].\end{aligned}$$

Finally, (7) gives $C(a,b+1,a) = 1 = C(a,b,a)$, so the expression above simplifies to the expression in (9). This verifies that the expression in (9) for $\mathfrak{u}^{\theta^i}(a,b,m,p)$ satisfies the recursion in (A.28) in states $(a,b) \in S_n^{**} \setminus S_n^{***}$. $\square$

A.10 Proof of Lemma 8

Proof. To prove equation (11), we first derive the recursion satisfied by $\mathfrak{u}_i(a,b,m,p)$. We then verify that the expression in (11) satisfies this recursion. The first part of the proof follows the same steps as the derivation of the recursion for V^A in the planner problem, and we do not repeat the detail here. The second part of the proof is in the spirit of the proof of Lemma 6.

First, we have that the payoff $\mathfrak{u}_i(a,b,m,p)$ of an uninformed agent i under the planner policy, assuming all m uninformed agents, including agent i, report truthfully, satisfies the recursion

$$\text{(A.30)}\quad \mathfrak{u}_i(a,b,m,p) = (1-g_m)\, v^M(p) + \frac{\lambda}{r+m\lambda}\Big[(m-1)\Big(p\,\mathfrak{u}_i(a+1,b,m-1,p) + (1-p)\,\mathfrak{u}_i(a,b+1,m-1,p)\Big) + p\,\mathfrak{u}^A(a+1,b,m-1,p) + (1-p)\,\mathfrak{u}^B(a,b+1,m-1,p)\Big],$$

where $v^M(p) = p-(1-p)$ is the single agent myopic payoff, and $\mathfrak{u}^A$ and $\mathfrak{u}^B$ are defined in (9). The recursion can be understood as follows. In state (a,b,m), the next piece of news arrives at rate g_m. With probability $(m-1)/m$ it comes from on of the $m-1$ uninformed agents other than agent i, so agent i remains uninformed and the state changes to account for the new pieces of news. With probability $1/m$, the next piece of news comes from agent i, so her continuation payoff u^{θ^i} now accounts for the fact that agent i has learnt (and truthfully reported) her type.

Next, we show that the expression in (11) satisfies this recursion. Our approach is separately to derive the coefficients in $\mathfrak{u}_i$, which we rewrite as

$$\text{(A.31)}\qquad \mathfrak{u}_i(a,b,m,p) = p\,\mathfrak{a}(a,b,m,p) + (1-p)\,(-1)\,\mathfrak{b}(a,b,m,p),$$

and we let

$$\text{(A.32)}\qquad \mathfrak{c}(a,b,m,p) := 1 - \sum_{k=a}^{\bar{a}} C(a,b,k)\,p^{k-a}(1-p)^{k-b}\left(\prod_{j=n-2k+1}^{n-a-b} g_j\right),$$

so that (9) gives $\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\,\mathfrak{c}(a,b,m,p)$. Substituting (A.31) and (A.32) into (A.30) yields the following recursions for the coefficients $\mathfrak{a}$ and $\mathfrak{b}$:

$$\text{(A.33)}\quad \mathfrak{a}(a,b,m,p) = \frac{1}{m}\Big[m(1-g_m) + g_m\,(m-1)\,[p\,\mathfrak{a}(a+1,b,m-1,p) + (1-p)\,\mathfrak{a}(a,b+1,m-1,p)] + g_m\,\mathfrak{c}(a+1,b,m-1,p)\Big],$$

$$\text{(A.34)}\quad \mathfrak{b}(a,b,m,p) = \frac{1}{m}\Big[m(1-g_m) + g_m\,(m-1)\,[p\,\mathfrak{b}(a+1,b,m-1,p) + (1-p)\,\mathfrak{b}(a,b+1,m-1,p)] + g_m\,\mathfrak{c}(a,b+1,m-1,p)\Big].$$

Observe that

$$\text{(A.35)}\qquad p\,\mathfrak{a}(a,b,m,p) + (1-p)\,\mathfrak{b}(a,b,m,p) = \mathfrak{c}(a,b,m,p).$$

Indeed, substituting (A.33) and (A.34) into the left hand side of (A.35) gives

$$
\begin{aligned}
[p\,\mathfrak{a} + (1-p)\,\mathfrak{b}](a,b,m,p) &= p\,\frac{1}{m}\Big[m(1-g_m) + g_m\,(m-1)\,[p\,\mathfrak{a}(a+1,b,m-1,p) + \,(1-p)\,\mathfrak{a}(a,b+1,m-1,p)] \\
&\qquad\qquad + g_m\,\mathfrak{c}(a+1,b,m-1,p)\Big] \\
&\quad + (1-p)\,\frac{1}{m}\Big[m(1-g_m) + g_m\,(m-1)\,[p\,\mathfrak{b}(a+1,b,m-1,p) + \,(1-p)\,\mathfrak{b}(a,b+1,m-1,p)] \\
&\qquad\qquad + g_m\,\mathfrak{c}(a,b+1,m-1,p)\Big] \\
&= (1-g_m) \\
&\quad + p\,\frac{1}{m} g_m\,(m-1)\,[p\,\mathfrak{a}(a+1,b,m-1,p) + (1-p)\,\mathfrak{b}(a+1,b,m-1,p)] \\
&\quad + (1-p)\,\frac{1}{m} g_m\,(m-1)\,[p\,\mathfrak{a}(a,b+1,m-1,p) + (1-p)\,\mathfrak{b}(a,b+1,m-1,p)] \\
&\quad + p\,\frac{1}{m} g_m\,\mathfrak{c}(a+1,b,m-1,p) + (1-p)\,\frac{1}{m} g_m\,\mathfrak{c}(a,b+1,m-1,p) \\
&= (1-g_m) \\
&\quad + p\,\frac{1}{m} g_m\,(m-1)\mathfrak{c}(a+1,b,m-1,p) \\
&\quad + (1-p)\,\frac{1}{m} g_m\,(m-1)\mathfrak{c}(a,b+1,m-1,p) \\
&\quad + p\,\frac{1}{m} g_m\,\mathfrak{c}(a+1,b,m-1,p) + (1-p)\,\frac{1}{m} g_m\,\mathfrak{c}(a,b+1,m-1,p) \\
&= (1-g_m) + \,g_m\,[p\,\mathfrak{c}(a+1,b,m-1,p) + (1-p)\,\mathfrak{c}(a,b+1,m-1,p)] \\
&= \mathfrak{c}(a,b,m,p),
\end{aligned}
$$

where the last equality is obtained by substituting $\mathfrak{u}^{\theta^i}(a,b,m,p) = v(\theta^i)\,\mathfrak{c}(a,b,m,p)$ into (A.28), yielding

$$
\begin{aligned}
v(\theta^i)\,\mathfrak{c}(a,b,m,p) &= (1-g_m)\,v(\theta^i) + g_m\,\Big[p\,v(\theta^i)\,\mathfrak{c}(a+1,b,m-1,p) + (1-p)\,v(\theta^i)\,\mathfrak{c}(a,b+1,m-1,p)\Big] \\
&= v(\theta^i)\Big[(1-g_m) + g_m\,[p\,\mathfrak{c}(a+1,b,m-1,p) + (1-p)\,\mathfrak{c}(a,b+1,m-1,p)]\Big].
\end{aligned}
$$

We therefore proceed by showing that $1 - \mathbb{E}[e^{-r\tau(a,b,m,p)}|\theta^i = A]$, where the expression for the conditional expectation is given in (12), satisfies the recursion (A.33). We then use (A.35) to obtain the expression in (13) for $\mathbb{E}[e^{-r\tau(a,b,m,p)}|\theta^i = B]$.

Substituting $1 - \mathbb{E}[e^{-r\tau(a,b,m,p)}|\theta^i = A]$ for $\mathfrak{a}(a,b,m,p)$ and $1 - \mathbb{E}[e^{-r\tau(a,b,m,p)}]$ for $\mathfrak{c}(a,b,m,p)$ in (A.33) gives

$$
\begin{aligned}
\mathfrak{a}(a,b,m,p) = \frac{1}{m}\Big[m(1-g_m) &+ g_m\,(m-1)\,p\,\Big(1 - \mathbb{E}[e^{-r\tau(a+1,b,m-1,p)}|\theta^i = A]\Big) \\
&+ g_m\,(m-1)\,(1-p)\,\Big(1 - \mathbb{E}[e^{-r\tau(a,b+1,m-1,p)}|\theta^i = A]\Big) \\
&+ g_m\,\Big(1 - \mathbb{E}[e^{-r\tau(a+1,b,m-1,p)}]\Big)\Big]
\end{aligned}
$$

which simplifies to

$$
\begin{aligned}
\mathfrak{a}(a,b,m,p) = \frac{1}{m}\Big[m - g_m\,(m-1)\,\Big(p\,\mathbb{E}[e^{-r\tau(a+1,b,m-1,p)}|\theta^i = A] &+ (1-p)\,\mathbb{E}[e^{-r\tau(a,b+1,m-1,p)}|\theta^i = A]\Big) \\
&- g_m\,\mathbb{E}[e^{-r\tau(a+1,b,m-1,p)}]\Big]
\end{aligned}
$$

Substituting the expression in (12) for the conditional expectations and the experssion in (10) for the unconditional expectation gives

$$
\begin{aligned}
\mathfrak{a}(a,b,m,p) = \frac{1}{m}\Big[m \\
&- g_m\,(m-1)p \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\,\frac{1}{m-1}\,[k-a-1+(n-2k)p]\,p^{k-a-2}(1-p)^{k-b} \prod_{j=n-2k+1}^{m-1} g_j \\
&- g_m\,(m-1)(1-p) \sum_{k=a}^{\bar{a}} C(a,b+1,k)\,\frac{1}{m-1}\,[k-a+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b-1} \prod_{j=n-2k+1}^{m-1} g_j \\
&- g_m \sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\; p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m-1} g_j\Big].
\end{aligned}
$$

Distributing the factors in g_m into the sums and isolating the summand for $k=a$ in the second sum, while observing that, by (7), $C(a,b+1,a)=1$, gives

$$
\begin{aligned}
\mathfrak{a}(a,b,m,p) = 1 &- \frac{1}{m}\sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\;[k-a-1+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j \\
&- \frac{1}{m}(n-2a)(1-p)^{a-b} \prod_{j=n-2a+1}^{m} g_j \\
&- \frac{1}{m}\sum_{k=a+1}^{\bar{a}} C(a,b+1,k)\;[k-a+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j \\
&- \frac{1}{m}\sum_{k=a+1}^{\bar{a}} C(a+1,b,k)\; p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j.
\end{aligned}
$$

Assembling the three sums and simplifying the resulting summand gives

$$
\begin{aligned}
\mathfrak{a}(a,b,m,p) = 1 &- \frac{1}{m}(n-2a)(1-p)^{a-b} \prod_{j=n-2a+1}^{m} g_j \\
&- \frac{1}{m}\sum_{k=a+1}^{\bar{a}} \Big\{C(a+1,b,k)+C(a,b+1,k)\Big\}\;[k-a+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j.
\end{aligned}
$$

Using (A.27) to substitute for the term in braces gives

$$
\begin{aligned}
\mathfrak{a}(a,b,m,p) = 1 &- \frac{1}{m}(n-2a)(1-p)^{a-b} \prod_{j=n-2a+1}^{m} g_j \\
&- \frac{1}{m}\sum_{k=a+1}^{\bar{a}} C(a,b,k)\;[k-a+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j.
\end{aligned}
$$

Finally, observe that the summand at the second line evaluated at $k=a$ equals $(n-2a)(1-p)^{a-b}\prod_{j=n-2a+1}^{m} g_j$, so that we indeed have that

$$
\mathfrak{a}(a,b,m,p) = 1 - \frac{1}{m}\sum_{k=a}^{\bar{a}} C(a,b,k)\;[k-a+(n-2k)p]\,p^{k-a-1}(1-p)^{k-b} \prod_{j=n-2k+1}^{m} g_j, \tag{A.36}
$$

which confirms that $\mathfrak{a}(a,b,m,p) = 1-\mathbb{E}[e^{-r\tau(a,b,m,p)}|\theta^i = A]$ with the conditional expectation given in (12).

Finally, we obtain that $\mathfrak{b}(a,b,m,p) = 1-\mathbb{E}[e^{-r\tau(a,b,m,p)}|\theta^i = B]$ by substituting (A.36) for $\mathfrak{a}$ and (A.32) for $\mathfrak{c}$ into (A.35), and obtaining the expression in (13) for the conditional expectation.

□

A.11 Proof of Lemma 9

Proof. The proof explores the relationship between the single agent's payoff and the planner (joint) payoff. First, we verify that the planner payoff in (6), the joint payoff of agents in this society, indeed aggregates the single-agent payoffs under the planner solution, (9) and (11). That is,

$$V^A(a,b,m,p) = a\,\mathfrak{u}^A(a,b,m,p) + b\,\mathfrak{u}^B(a,b,m,p) + m\,\mathfrak{u}_i(a,b,m,p). \tag{A.37}$$

Indeed, substituting (9) and (11) in the right-hand side above gives

$$\begin{aligned}
[a\,\mathfrak{u}^A + b\,\mathfrak{u}^B + m\,\mathfrak{u}_i](a,b,m,p) &= (a-b)\,\Big(1-\mathbb{E}\big[e^{-r\tau(a,b,m,p)}\big]\Big) \\
&\quad + m\,\Big[p\,\Big(1-\mathbb{E}\big[e^{-r\tau(a,b,m,p)}|\theta^i=A\big]\Big) - (1-p)\,\Big(1-\mathbb{E}\big[e^{-r\tau(a,b,m,p)}|\theta^i=B\big]\Big)\Big] \\
&= a-b+m(p-(1-p)) - (a-b)\,\mathbb{E}\big[e^{-r\tau(a,b,m,p)}\big] \\
&\quad - m\,p\,\mathbb{E}\big[e^{-r\tau(a,b,m,p)}|\theta^i=A\big] + m\,(1-p)\,\mathbb{E}\big[e^{-r\tau(a,b,m,p)}|\theta^i=B\big] \\
&= u(a,b,m) - (a-b)\sum_{k=a}^{\bar{a}} C(a,b,k)\;p^{k-a}(1-p)^{k-b}\prod_{j=n-2k+1}^{m} g_j \\
&\quad - \sum_{k=a}^{\bar{a}} C(a,b,k)\;[k-a+(n-2k)p]\,p^{k-a}(1-p)^{k-b}\prod_{j=n-2k+1}^{m} g_j \\
&\quad + \sum_{k=a}^{\bar{a}} C(a,b,k)\;[k-b+(n-2k)(1-p)]\,p^{k-a}(1-p)^{k-b}\prod_{j=n-2k+1}^{m} g_j \\
&= u(a,b,m) - \sum_{k=a}^{\bar{a}} C(a,b,k)\,[(n-2k)(p-(1-p))]\,p^{k-a}(1-p)^{k-b}\prod_{j=n-2k+1}^{m} g_j
\end{aligned}$$

where, at the third equality, we substitute (10), (12) and (13) and simplify, and where using $u(k,k,n-2k) = (n-2k)(p-(1-p))$ establishes that the last expression equals (6). $\square$

A.12 Proof of Proposition 2

We use a "guess and verify" approach. Consider the Lagrangian where the IC for $t=0$ has multiplier $M(\geq 0)$ and the IC for $t>0$ has multiplier $\mu(t)(\geq 0)$:

$$L(F;M,\mu) = \mathrm{Obj}(F) + M\,\mathrm{IC}_0(F) + \int_{\tau>0} \mu(\tau)\,\mathrm{IC}_\tau(F).$$

Observe that:

$$\mathrm{Obj}(F) \leq L(F;M,\mu) \leq \max_F L(F;M,\mu)$$

given any feasible plan F, and thus, if (under certain parametric condition) we could find multipliers (M^*,μ^*) such that $\max_F L(F;M,\mu) = Obj(F^*)$ for some feasible plan F^*, then it must be that F^* is the optimal policy (under that parametric condition). Moreover,

$$F^* = \arg\max_F L(F;M,\mu)$$
$$M^*\,\mathrm{IC}_0(F^*) + \int_{\tau>0} \mu^*(\tau)\,\mathrm{IC}_\tau(F^*) = 0.$$

We consider the following guess of the optimal plan F^*:

- F_A is such that $A(t) = 1$ for all t;

- F_B is such that $B(t)$ is interior for $t < t^*$, and $B(t) = 0$ for $t > t^*$;
- $F_N(t) = 0$ for $t < t^*$, and $F_N(t) = 1$ for $t > t^*$.

The associated multipliers (M^*, μ^*) are such that the slope for each $B(t)$ is zero:

$$M = \frac{q(1-p)(1-\alpha)}{1-p-p\alpha}$$
$$\mu(t)e^{\lambda t} = \frac{\lambda(1-p)}{1-p-p\alpha}((1-\alpha)(1-q) + M + \int_{\tau=0}^{t} \mu(\tau)e^{\lambda\tau}).$$

Seeing the system of the second equations for $t > 0$ as a differential equation for $G(t) = \int_{\tau=0}^{t} \mu(\tau)e^{\lambda\tau}d\tau$ and solving it (with initial condition $G(0) = 0$), we obtain:

$$G(t) = ((1-\alpha)(1-q) + M)\exp(\frac{\lambda(1-p)}{1-p-p\alpha}t),$$

and thus,

$$\mu(t)e^{\lambda t} = ((1-\alpha)(1-q) + M)\frac{\lambda(1-p)}{1-p-p\alpha}\exp(\frac{\lambda(1-p)}{1-p-p\alpha}t).$$

We now verify that our guess is correct under certain parametric condition. First, the optimality of setting $A(t) = 1$ for all t is immediate from the Lagrangian. The optimality of $B(t)$ is also fine, because the multipliers are such that any choice of $B(t)$ is optimal. Later, $B(t)$ is to be chosen so that all the incentive compatibility conditions are binding with equality.

Finally, we verify that there exists t^* such that it is optimal to set $F_N(t) = 0$ for $t < t^*$ and $F_N(t) = 1$ for $t > t^*$. The Lagrangian is linear in each $F_N(t)$, whose coefficient is proportional to:

$$\begin{aligned}
\phi(t) \quad := \quad & (2\alpha p - (1-p)(1-\alpha))(1-q)[\lambda(1-e^{-t}) - e^{-t}] - 2\alpha p(1-q)\lambda \\
& -M[\lambda((1-p-p\alpha)(1-e^{-t}) - p\alpha) - (1-p-p\alpha)e^{-t}] \\
& -(1-p-p\alpha)\int_{\tau=0}^{t} \mu(\tau)e^{\lambda\tau}[\lambda(1-e^{-(t-\tau)}) - e^{-(t-\tau)}].
\end{aligned}$$

Observe:

$$\begin{aligned}
& \int_{\tau=0}^{t} \mu(\tau)e^{\lambda\tau}[\lambda(1-e^{-(t-\tau)}) - e^{-(t-\tau)}] \\
= \quad & \frac{\lambda}{\lambda(1-p) + 1-p-p\alpha}[(\lambda+1)(1-p)e^{-t} - p\alpha\exp(\frac{\lambda(1-p)}{1-p-p\alpha}t)],
\end{aligned}$$

and thus, the coefficient of $F_N(t)$ further simplifies to:

$$\begin{aligned}
\phi(t) \quad = \quad & -\lambda((1-p)(1-\alpha) - M\alpha p) \\
& +\frac{\lambda p\alpha(1-p-p\alpha)}{\lambda(1-p) + 1-p-p\alpha}\exp(\frac{\lambda(1-p)}{1-p-p\alpha}t) \\
& -(\lambda+1)[2\alpha p(1-q) - (1-p)(1-\alpha) + \frac{\lambda(1-p)(1-p-p\alpha)}{1-p-p\alpha+\lambda(1-p)}]e^{-t}.
\end{aligned}$$

It is obvious that $\phi(t) > 0$ for large enough t. If $\phi(t)$ crosses with the horizontal axis only once ("single-crossing"), then there must exist t^* such that $F_N(t) = 0$ for $t < t^*$ and $F_N(t) = 1$ for $t > t^*$ (possibly $t^* = 0$).[14]

Finally, we set $B(t)$ to satisfy the incentive compatibility constraints with equality. First, we can set $B(t) = 0$ for all $t > t^*$. For $t < t^*$, by IC:

$$\begin{aligned}
(1-p-p\alpha)B(t) \quad = \quad & (1-p)\int_{\tau=t}^{t^*} B(\tau)d\tau + (1-p)e^{-\lambda(t^*-t)}(1-e^{-(t^*-t)}) - p\alpha(1-e^{-(t^*-t)(\lambda+1)}) \\
= \quad & (1-p)\int_{\tau=t}^{t^*} B(\tau)d\tau + (1-p)e^{-\lambda(t^*-t)} - p\alpha - (1-p-p\alpha)e^{-(\lambda+1)(t^*-t)}.
\end{aligned}$$

[14] In case it crosses with the horizontal axis multiple times, a possible idea may be to use the fact that F_N must be non-decreasing.

Solving this differential equation, we can obtain a closed-form expression for $B(t)$.

Finally, we verify that truthful reporting is optimal for the novice upon observing either signal, A or B. A novice getting a B signal at $t < t^*$ strictly prefers reporting B at t, thereby receiving a continuation payoff of $-B(t)$, to

- reporting A at $s \in [t, t^*]$, which yields a continuation payoff of $-(1 - e^{-(s-t)}) - e^{-(s-t)}A(s) = -1$, but we have $B(t) \leq 1$ by definition;
- reporting N on $[t, t^*]$, which results in a switch to B at date t^* yielding a continuation payoff of $-(1 - e^{-(t^*-t)})$, which is no greater than $-B(t) = -\int_t^{t^*}(1 - e^{-(s-t)})\mathrm{d}F_B(s,t)$, since F_B induces a switch prior to t^* with positive probability;
- reporting B at $s \in (t, t^*]$, which yields a continuation payoff of $-(1 - e^{-(s-t)}) - e^{-(s-t)}B(s)$, which, again, is no greater than $-B(t)$, since F_B induces a switch on (t, s) with positive probability

A novice getting an A signal at $t < t^*$ strictly prefers reporting A at t, thereby receiving a continuation payoff of α, to

- reporting B at $s \in [t, t^*]$, which yields a continuation payoff of $\alpha(1 - e^{-(s-t)}) + \alpha e^{-(s-t)}B(s) \leq \alpha$, where the inequality follows from $B(t) \leq 1$;
- reporting N on $[t, t^*]$, yielding a continuation payoff of $\alpha(1 - e^{-(t^*-t)}) < \alpha$.
- reporting A at $s \in (t, t^*]$, yielding a continuation payoff of $\alpha(1 - e^{-(s-t)}) + \alpha e^{-(s-t)}A(s) = \alpha$.

A.13 Proof of Proposition 3

We first show that there exists an optimal solution $\mathcal{D}^t$ to the planner's relaxed problem with only constraint (16), in which $\mathcal{D}^t(a, b+1, m) = \mathcal{D}^t(a, b, m+1)$ for every $(a, b, m) \in S_{n-1}$. Then we verify that other constraints.

The relaxed problem, i.e., maximising (15) with the constraint (16) and the feasibility constraint $\mathcal{D}^t(a, b, m) \in [0, 1]$ for any $(a, b, m) \in S_n$, is a linear programming problem, in which the objective function is bounded and a feasible solution exists. Thus, an optimal solution exists.

Let us pick one optimal solution $\mathcal{D}^t$ and denote the number of states in S_{n-1} in which $\mathcal{D}^t(a, b+1, m) \neq \mathcal{D}^t(a, b, m+1)$ to be $d \geq 0$, where $(a, b, m) \in S_{n-1}$. We call such a state a *disagreement* state. Note that d is finite, since the number of total states is finite. If $d = 0$, we have our result.

If $d \geq 1$, we now show we can find another optimal solution in which the number l' of the disagreement states is at most $l - 1$, i.e., $l' \leq l - 1$. Let $s^1 = (a^1, b^1, m^1) \in S_{n-1}$ be a disagreement state. IC constraint (16) implies that it cannot be the only disagreement state, since they must balance out in expectation. Let $s^2 = (a^2, b^2, m^2) \in S_{n-1}$ be the paired disagreement state. Without loss of generality, assume $0 \leq \mathcal{D}^t(a^1, b^1, m^1+1) < \mathcal{D}^t(a^1, b^1+1, m^1) \leq 1$ and $0 \leq \mathcal{D}^t(a^2, b^2+1, m^2) < \mathcal{D}^t(a^2, b^2, m^2+1) \leq 1$. Let $\gamma \neq 0$ be the Lagrange multiplier for the constraint (16).

Since the optimal policy $\mathcal{D}^t$ satisfies $0 \leq \mathcal{D}^t(a^1, b^1, m^1+1) < \mathcal{D}^t(a^1, b^1+1, m^1) \leq 1$, the first order conditions on the state (a^1, b^1, m^1+1) and (a^1, b^1+1, m^1) of the Lagrange function satisfy

$$\pi_{n-1}^t(a^1, b^1, m^1)\left[q^t(m)u(a^1, b^1, p) - \gamma\right] \leq 0, \tag{A.38}$$

and

$$\pi_{n-1}^t(a^1, b^1, m^1)\left[q^t(b)u(a^1, b^1+1, p) + \gamma\right] \geq 0, \tag{A.39}$$

where $q^t(m) = \pi_n^t(a^1, b^1, m^1+1)/\pi_{n-1}^t(a^1, b^1, m^1)$ and $q^t(b) = \pi_n^t(a^1, b^1+1, m^1)/\pi_{n-1}^t(a^1, b^1, m^1)$ are the probabilities that an agent has observed no news and B-news by time t, respectively. Because the news process is independent across agents, $q^t(m)$ and $q^t(b)$ do not depend on the state $(a^1, b^1, m^1) \in S_{n-1}$. Thus,

$$q^t(b)u(a^1, b^1+1, p) + \gamma \geq 0 \geq q^t(m)u(a^1, b^1, p) - \gamma. \tag{A.40}$$

Similarly, since the optimal policy $\mathcal{D}^t$ satisfies $0 \le \mathcal{D}^t(a^2, b^2+1, m^2) < \mathcal{D}^t(a^2, b^2, m^2+1) \le 1$, the first order conditions on the state (a^2, b^2+1, m^2) and (a^2, b^2, m^2+1) of the Lagrange function satisfy

$$\pi^t_{n-1}(a^2, b^2, m^2)\left[q^t(b)u(a^2, b^2+1, p) + \gamma\right] \le 0, \tag{A.41}$$

$$\pi^t_{n-1}(a^2, b^2, m^2)\left[q^t(m)u(a^2, b^2, p) - \gamma\right] \ge 0, \tag{A.42}$$

which implies

$$q^t(m)u(a^2, b^2, p) - \gamma \ge 0 \ge q^t(b)u(a^2, b^2+1, p) + \gamma. \tag{A.43}$$

Inequalities (A.40) and (A.43) together, we have

$$q^t(b)u(a^1, b^1+1, p) + \gamma \ge 0 \ge q^t(b)u(a^2, b^2+1, p) + \gamma, \tag{A.44}$$

$$q^t(m)u(a^2, b^2, p) - \gamma \ge 0 \ge q^t(m)u(a^1, b^1, p) - \gamma. \tag{A.45}$$

Equivalently, we have

$$u(a^1, b^1+1, p) \ge u(a^2, b^2+1, p), \tag{A.46}$$

$$u(a^2, b^2, p) \ge u(a^1, b^1, p). \tag{A.47}$$

Note that $u(a, b+1, p) = u(a, b, p) - (1+\alpha)p$, thus (A.46) implies that $u(a^1, b^1, p) \ge u(a^2, b^2, p)$. With (A.47), this means that $u(a^1, b^1, p) = u(a^2, b^2, p)$ and $u(a^1, b^1+1, p) = u(a^2, b^2+1, p)$. Thus, they imply (A.44) and (A.45) hold with equalities, which in turn imply that (A.38), (A.39), (A.41) and (A.42) hold with equalities.

The first order conditions all hold with equalities in this linear programming problem. This means that the complementary slackness conditions behind the inequalities are slack and do not restrict the optimal solution on those relevant states to the boundary of $[0, 1]$. Thus, we are free to modify the policy on those states within the feasible region, without affecting its optimality.

Let $\Delta^1 = \mathcal{D}^t(a^1, b^1+1, m^1) - \mathcal{D}^t(a^1, b^1, m^1+1) \in (0, 1]$ and $\Delta^2 = \mathcal{D}^t(a^2, b^2+1, m^2) - \mathcal{D}^t(a^2, b^2, m^2+1) \in [-1, 0)$. Without loss, assume that $\pi^t_{n-1}(a^1, b^1, m^1)\Delta^1 + \pi^t_{n-1}(a^2, b^2, m^2)\Delta^2 \ge 0$. Consider a new policy $\hat{\mathcal{D}}^t$, in which the policy on all other states coincides with $\mathcal{D}^t$, expect on states $(a^1, b^1+1, m^1), (a^1, b^1, m^1+1), (a^2, b^2+1, m^2), (a^2, b^2, m^2+1) \in S_n$. If $\hat{\mathcal{D}}^t$ is feasible, it is also optimal. To be feasible, all we need is that $\hat{\mathcal{D}}^t \in [0, 1]$ on those relevant states. Note that (16) implies

$$\begin{aligned}
& \pi^t_{n-1}(a^1, b^1, m^1)\Delta^1 + \pi^t_{n-1}(a^2, b^2, m^2)\Delta^2 \\
= \;& \pi^t_{n-1}(a^1, b^1, m^1)\left[\hat{\mathcal{D}}^t(a^1, b^1+1, m^1) - \hat{\mathcal{D}}^t(a^1, b^1, m^1+1)\right] \\
+ \;& \pi^t_{n-1}(a^2, b^2, m^2)\left[\hat{\mathcal{D}}^t(a^2, b^2+1, m^2) - \hat{\mathcal{D}}^t(a^2, b^2, m^2+1)\right].
\end{aligned}$$

If $\pi^t_{n-1}(a^1, b^1, m^1)\Delta^1 + \pi^t_{n-1}(a^2, b^2, m^2)\Delta^2 = 0$, we make a feasible modification $\hat{\mathcal{D}}^t = 1$ on $(a^1, b^1+1, m^1), (a^1, b^1, m^1+1), (a^2, b^2+1, m^2), (a^2, b^2, m^2+1) \in S_n$ so that the number of the disagreement states on S_{n-1} in the new policy is $l' = l - 2$. If $\pi^t_{n-1}(a^1, b^1, m^1)\Delta^1 + \pi^t_{n-1}(a^2, b^2, m^2)\Delta^2 > 0$, then $\Delta^1 + \frac{\pi^t_{n-1}(a^2, b^2, m^2)}{\pi^t_{n-1}(a^1, b^1, m^1)}\Delta^2 \in (0, 1)$. It is easy to see that the following modification is feasible: $\hat{\mathcal{D}}^t(a^1, b^1+1, m^1) = \hat{\mathcal{D}}^t(a^2, b^2+1, m^2) = \hat{\mathcal{D}}^t(a^2, b^2, m^2+1) = 1$ and

$\hat{\mathcal{D}}^t(a^1, b^1, m^1+1) = 1 - \Delta^1 - \frac{\pi^t_{n-1}(a^2,b^2,m^2)}{\pi^t_{n-1}(a^1,b^1,m^1)}\Delta^2 \in (0,1)$. The the number of the disagreement states on S_{n-1} in the new policy is $l' = l-1$.

Continuing this process, we can find an optimal policy in which the number of the disagreement states on S_{n-1} is zero. Thus, there exists an optimal solution to the relaxed problem in which for every $(a,b,m) \in S_{n-1}$, $\mathcal{D}^t(a, b+1, m) = \mathcal{D}^t(a, b, m+1)$. Thus, for any fixed $a \in \{0, 1, \cdots, n\}$, $\mathcal{D}^t(a,b,m) = \mathcal{D}^t(a, b', m')$, where $(a,b,m), (a,b',m') \in S_n$. That is, the optimal policy $\mathcal{D}^t$ depends only on the number of reports on A-news. Let us write it as $\mathcal{D}^t(a)$, for $a \in \{0, 1, \cdots, n\}$. Thus, the planner's objective can be written as

$$\sum_{a=0}^{n} \mathcal{D}^t(a) \sum_{\{(b,m):(a,b,m)\in S_n\}} \pi^t_n(a,b,m)\, u(a,b,p).$$

Obviously, it is optimal that $\mathcal{D}^t(a) = 1$ ($\mathcal{D}^t(a) = 0$) if the expected payoff conditional on a is positive (negative), i.e., $\sum_{\{(b,m):(a,b,m)\in S_n\}} \pi^t_n(a,b,m)\, u(a,b,p) > 0$ (< 0), for any given a. Let $p_t = \frac{pe^{-\lambda(t+\underline{T})}}{pe^{-\lambda(t+\underline{T})}+1-p}$ denote the posterior belief on the agent being type A but having not observed A-news by time t. Note that $p_t < p$ and it is strictly decreasing in t.We can calculate the expected payoff conditional on a:

$$\frac{\sum_{\{(b,m):(a,b,m)\in S_n\}} \pi^t_n(a,b,m)\, u(a,b,p)}{\sum_{\{(b,m):(a,b,m)\in S_n\}} \pi^t_n(a,b,m)} = \alpha a + [\alpha p_t - (1-p_t)]\,(n-a),$$

which is strictly increasing in both a and p_t. This policy means that there exists a threshold $\bar{a}^t$, which is increasing in t, such that $\mathcal{D}^t(a) = 1$ if and only if $a \geq \bar{a}^t$. Moreover, $\bar{a}^t \geq 1$ as $p_t < p < p^M$, and $\bar{a}^t \leq \bar{a}+1$ as the expected payoff is positive when $a = \bar{a}+1$ even if $p_t = 0$.

Finally, the above solution satisfies all other IC constraints that are excluded from the relaxed problem, as reporting A-news strictly increases the probability that the total number of A-reports exceeds $\bar{a}^t$.

A.14 Proof of Proposition 4

In both scenarios, note that for $\iota \in \{F, B\}$,

$$\text{(A.48)} \qquad \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\right] < \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(0)}\right] < \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(1)}\right].$$

The first inequality comes from that

$$\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(0)}\right] = \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{B^\iota(0)}\right] < 0,$$

by the definition of the event $B^\iota(0)$.

The second inequality comes from that

$$\begin{aligned}
&\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(0)}\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(1)}\right] \\
=&\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, (\mathbb{1}_{A^\iota(0)\cap A^\iota(1)} + \mathbb{1}_{A^\iota(0)\cap B^\iota(1)})\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, (\mathbb{1}_{A^\iota(1)\cap A^\iota(0)} + \mathbb{1}_{A^\iota(1)\cap B^\iota(0)})\right] \\
=&\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(0)\cap B^\iota(1)}\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(1)\cap B^\iota(0)}\right] \\
=&\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, |B^\iota(1)\right] \mathbb{P}\left[A^\iota(0)\cap B^\iota(1)\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, |A^\iota(1)\right] \mathbb{P}\left[A^\iota(1)\cap B^\iota(0)\right] < 0.
\end{aligned}$$

The last inequality is due to the definition of the events $B^\iota(1)$ and $A^\iota(1)$. The last equality is due to the fact that the conditional expected payoff conditional on information on both date 0 and 1 is the same to the conditional expected payoff conditional on information only on date 1.

The difference of the expected payoffs between holding the election on date 0 and that on date 1 in scenario $\iota \in \{F, B\}$ is

$$\begin{aligned}
&\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(0)}\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\right] \\
&-\delta\left[\mathbb{E}\left[(\alpha\, x_A - (n-x_A))\, \mathbb{1}_{A^\iota(1)}\right] - \mathbb{E}\left[(\alpha\, x_A - (n-x_A))\right]\right].
\end{aligned}$$

Define

$$\delta^{\iota} = \frac{\mathbb{E}\left[(\alpha\, x_A - (n - x_A))\, \mathbb{1}_{A^{\iota}(0)}\right] - \mathbb{E}\left[(\alpha\, x_A - (n - x_A))\right]}{\mathbb{E}\left[(\alpha\, x_A - (n - x_A))\, \mathbb{1}_{A^{\iota}(1)}\right] - \mathbb{E}\left[(\alpha\, x_A - (n - x_A))\right]}.$$

It is immediate that $\delta^{\iota} \in (0,1)$ due to A.48 and our results follow.

B Second-best mechanism in the discrete-time framework

Although the continuous-time framework considered in the main text has the advantage of simplicity and clarity in presentation, its formal treatment could be much trickier. In fact, we do not show that the "second-best mechanism" considered in the main text is optimal within the continuous-time framework. What we show here is that the same kind of mechanism is indeed optimal in the discrete-time framework, and as the time length shrinks to zero, it becomes closer and closer to the one described in the main text. That is the sense that we (informally) argue that that is the second-best mechanism "in continuous time".

B.1 Problem and Lagrangian

The discrete-time version of the problem shares the same notation as in the main text, except that the time $t \in \{0, 1, \ldots\}$. The discount factor is $\delta \in (0,1)$, and the probability that an uninformed voter learns that he is A (resp. B) type is Λp (resp. $\Lambda(1-p)$) at any given t, where $\Lambda \in (0,1)$. In order to make the discrete-time model closer to the continuous-time counterpart, we may let $\delta = e^{-r\Delta}$, $\Lambda = e^{-\lambda\Delta}$ and let $\Delta \to 0$, as standard. For the rest of the material, we keep the notation δ and Λ to simplify the notation.

Assumption 2. *Assume p is less than the single-agent experimentation threshold (i.e., $p < \frac{1}{1+\alpha+\frac{\Lambda\delta\alpha}{1-\delta}}$).*

Also, the parameters are such that the first-best mechanism looks as follows, as a function of (a_t, b_t).

	$b_t = 0$	$b_t = 1$	$b_t = 2$
$a_t = 0$	A	B	B
$a_t = 1$	A	A	
$a_t = 2$	A		

The principal's second-best problem is given as follows:

$$\begin{aligned} \max_{q} \quad & \sum_{t=0}^{\infty} \delta^t \mathbb{E}_{(a_{\le t}, b_{\le t})}[q_t(a_{\le t}, b_{\le t}) w(a_t, b_t)] \\ \text{sub. to} \quad & q_{t-1}(a_{\le t-1}, b_{\le t-1}) \ge q_t(a_{\le t-1}, a_t, b_{\le t-1}, b_t), \ \forall t \\ & IC_t^N, \ \forall t, \end{aligned}$$

where $q_{-1}(\cdot) = 1$ by convention.

The uninformed voter's incentive compatibility at time t, IC_t^N, is the same as in the main text:

$$\mathbb{E}^{n-1}_{(a_{\le t}, b_{\le t})}[V_t^N(a_{\le t}, b_{\le t}) - V_t^B(a_{\le t}, (b_{<t}, b_t + 1)) \frac{p\alpha - (1-p)}{(-1)}] \ge 0,$$

but it is rewritten in the discrete-time framework without "dt":

Lemma B.11.

$$\begin{aligned} & \mathbb{E}^{n-1}_{(a_{\le t}, b_{\le t})}[V_t^N(a_{\le t}, b_{\le t}) - V_t^B(a_{\le t}, (b_{<t}, b_t + 1)) \frac{p\alpha - (1-p)}{(-1)}] \ge 0 \\ \Leftrightarrow \quad & \mathbb{E}_{(a_{\le t}, b_{\le t})}[V_t^N(a_{\le t}, b_{\le t})(2 - a_t - b_t) - V_t^B(a_{\le t}, b_{\le t}) \frac{b_t - b_{t-1}}{\Lambda p} \frac{p\alpha - (1-p)}{-1}] \ge 0. \end{aligned}$$

Proof. For the first term, we use:

$$\begin{aligned}
&\mathbb{E}^{n-1}_{(a_{\le t},b_{\le t})}[V_t^N(a_{\le t},b_{\le t})] - \mathbb{E}_{(a_{\le t},b_{\le t})}[V_t^N(a_{\le t},b_{\le t})\frac{2-a_t-b_t}{n(1-\Lambda)}] \\
= \; & \frac{(n-1)!(\Lambda p)^{\sum_{s\le t} a_s}(\Lambda(1-p))^{\sum_{s\le t} b_s}}{a_0!b_0!\ldots a_t!b_t!(n-1-\sum_{s\le t}(a_s+b_s))!}(1-\Lambda)^{n-1-\sum_{s\le t}(t+1-s)(a_s+b_s)} \\
& -\frac{n!(\Lambda p)^{\sum_{s\le t} a_s}(\Lambda(1-p))^{\sum_{s\le t} b_s}}{a_0!b_0!\ldots a_t!b_t!(n-\sum_{s\le t}(a_s+b_s))!}(1-\Lambda)^{n-1-\sum_{s\le t}(t+1-s)(a_s+b_s)}\frac{2-a_t-b_t}{n} = 0,
\end{aligned}$$

(recalling $n=2$); and similarly for the second term. □

Now we are ready to provide the Lagrangian:

$$\begin{aligned}
L(\eta_{(\cdot)};q_{(\cdot)}(\cdot)) \;=\; & \sum_{t=0}^{\infty}\delta^t\mathbb{E}_{(a_{\le t},b_{\le t})}\Big[q_t(a_{\le t},b_{\le t})w(a_t,b_t) \\
& +\eta_t\left\{V_t^N(a_{\le t},b_{\le t})(2-a_t-b_t) - V_t^B(a_{\le t},b_{\le t})\frac{b_t-b_{t-1}}{\Lambda p}\frac{p\alpha-(1-p)}{-1}\right\}\Big],
\end{aligned}$$

where $\eta_{(\cdot)}$ and $q_{(\cdot)}(\cdot)$ are any non-negative vectors (including those which are infeasible in the primal problem). As in the main text, we ignore the monotonicity constraint. We have the following result, which corresponds to the (weak) duality result for Lagrangian.

Lemma B.12. *If there exist non-negative multiplier vector $\eta^*_{(\cdot)}$, and an incentive compatible mechanism $q^*_{(\cdot)}(\cdot)$ such that (i)*

$$L(q^*_{(\cdot)}(\cdot);\eta^*_{(\cdot)}) \ge L(q_{(\cdot)}(\cdot);\eta^*_{(\cdot)})$$

for any non-negative $q_{(\cdot)}(\cdot)$, and (ii) the value of the primal problem given mechanism $q^*_{(\cdot)}(\cdot)$ coincides with $L(q^*_{(\cdot)}(\cdot),\eta^*_{(\cdot)})$, then $q^*_{(\cdot)}(\cdot)$ solves the primal problem, that is, is a second-best mechanism.

Proof. Suppose that the hypothesis in the statement is satisfied.

Let $W(q_{(\cdot)}(\cdot))$ denote the value of the primal objective (i.e., the principal's ex ante expected payoff) given any incentive compatible mechanism $q_{(\cdot)}(\cdot)$. It is immediate that, for any non-negative multiplier vector $\eta_{(\cdot)}$, we have:

$$L(\eta_{(\cdot)};q_{(\cdot)}(\cdot)) \ge W(q_{(\cdot)}(\cdot)).$$

Therefore, for any incentive compatible mechanism $q_{(\cdot)}(\cdot)$, we have:

$$\begin{aligned}
W(q^*_{(\cdot)}(\cdot)) &= L(\eta^*_{(\cdot)};q^*_{(\cdot)}(\cdot)) \\
&\ge L(\eta^*_{(\cdot)};q_{(\cdot)}(\cdot)) \\
&\ge W(q_{(\cdot)}(\cdot)),
\end{aligned}$$

meaning that $q^*_{(\cdot)}(\cdot)$ is optimal in the primal problem. □

B.2 Candidate solution

The form of the candidate solution is similar as in the main text: There exists t^* such that

1. For any $t<t^*$: (i) unless $(a_t,b_t)=(0,2)$, stay in A; (ii) if $(a_t,b_t)=(0,2)$, then either delay or probabilistic switch to B.
2. At $t=t^*$: (i) if $a_t=0$, then immediately switch to B; (ii) otherwise, stay in A.

As before, this may be seen as a class of mechanisms. As in Observation 1 and 2 of the main text, the extreme versions of the mechanism in this class either makes all the IC constraints (for $t < t^*$) slack or all violated. Observation 3 claims (informally in the main text) that, as a consequence, we should be able to find a mechanism in this class that satisfies all the IC constraints with equality (for $t < t^*$). The first part of the following theorem formalizes those observations, and the second part is about its optimality.

Theorem 1. *There exists a mechanism in the above class such that all the IC constraints for $t < t^*$ hold with equality. Moreover, it is a second-best mechanism.*

The rest of this section provides the proof. First, we show the existence part. To show this, first, we make two observations that correspond to Observations 1 and 2 in the main text.

Lemma B.13. *The version of the mechanism in the above class such that $q_t((\phi, t_1 t_2)_t) = 1$ for all $t_1, t_2 < t^*$ satisfies all the IC constraints (possibly with strict inequalities).*

We omit the proof as it is straightforward.

Lemma B.14. *The version of the mechanism in the above class such that $q_t((\phi, t_1 t_2)_t) = 0$ for all $t_1, t_2 < t^*$ violates all the IC constraints.*

Proof. By direct computation of

$$\mathbb{E}^{n-1}_{a_{\le t}, b_{\le t}}[V^N_t(a_{\le t}, b_{\le t}) - V^B_t(a_{\le t}, b_{<t}, b_t + 1)\frac{p\alpha - (1-p)}{-1}].$$

□

To conclude the existence part, we now obtain an analog of Observation 3: Consider the extreme version of the mechanism considered in Observation 1 (i.e., the version with $q_t((\phi, t_1 t_2)_t) = 1$ for all $t_1, t_2 < t^*$), where all the IC constraints are satisfied, possibly with strict inequalities.

Consider the period $t = t^* - 1$. We decrease $q_{t^*-1}((\phi, t_1 t_2)_{t^*-1})$ gradually, uniformly across all $t_1, t_2 \le t^* - 1$, until $IC^N_{t^*-1}$ holds with equality. This operation also affects IC^N_t for $t < t^* - 1$, but never violates IC^N_t: Indeed, the first term of the left-hand side of IC^N_t (i.e., the expected continuation value of truth-telling N) comprises (i) the flow payoff of time t ($p\alpha - 1 - p$) and (ii) the (discounted) expected continuation payoff from $t + 1$ on, which comprises (ii-A) the case where this uninformed voter becomes A type at $t + 1$, (ii-B) the case where he becomes B type at $t + 1$, and (ii-N) the other case. If misreporting B at t, his payoff would comprise (i') the flow payoff of time t, $p\alpha - 1 - p$ (recall $q_t(\cdot) = 1$ for $t < t^* - 1$ at this point) and (ii') the (discounted) expected continuation payoff from $t + 1$ on, which comprises (ii'-A) the case where this uninformed voter becomes A type at $t + 1$, (ii'-B) the case where he becomes B type at $t + 1$, and (ii'-N) the other case. (i) and (i') attain the same payoff, so do (ii-B) and (ii'-B) (because $q_{t^*-1}((\phi, t_1 t_2)_{t^*-1})$ does not depend on t_1, t_2). (ii-A) clearly attains higher payoff than (ii'-A). Finally, (ii-N) attains higher payoff than (ii'-N) as long as IC^N_{t+1} holds (weakly or strictly).

Apply the same procedure inductively: For any given $t < t^* - 1$, suppose that all $IC^N_{t'}$ hold with equality for all $t' = t + 1, \dots t^* - 1$, with $q_{t'}((\phi, t_1 t_2)_{t'})$ invariant in t_1, t_2. By the same logic as above, we can gradually decrease $q_t((\phi, t_1 t_2)_t)$ uniformly across all $t_1, t_2 \le t$ until IC^N_t holds with equality. This operation does not violate $IC^N_{t''}$ for $t'' < t$ by the same logic above, and it does not affect $IC^N_{t'}$ for $t' > t$ at all because $q_t((\phi, t_1 t_2)_t)$ does not matter for $IC^N_{t'}$.

Next, we verify that this mechanism is indeed second-best. Based on the previous weak-duality lemma, our task is to identify the right Lagrange multiplier and an incentive compatible mechanism that is optimal in the Lagrangian problem, with the same value as in the previous expression.

Because our candidate optimal mechanism requires interior $q_t((\phi, t_1t_2)_t)$ for $t_1, t_2 < t^*$. Because the Lagrangian is linear in $q_t((\phi, t_1t_2)_t)$, that would be only possible if its slope is zero. This leads to our guess of the multipliers:

- For $s < t^*$: $\tilde{\eta}_s = \frac{2-\alpha_P}{2}\gamma(1+\gamma)^t$ where $\gamma = \frac{\Lambda(1-p)}{(1-\Lambda)(1-p-p\alpha)} (> 0)$.
- For $s > t^*$: $\tilde{\eta}_s = 0$.
- t^* is such that:

$$(1+\gamma)^{t^*} = \frac{\alpha_P p(1+\alpha)}{(2-\alpha_P)(1-p-p\alpha)} \tag{B.49}$$

Hence, $\int_{s\le t} \tilde{\eta}_s ds = \frac{2-\alpha_P}{2}((1+\gamma)^t - 1)$ for $t < t^*$, and $\int_{s\le t} \tilde{\eta}_s ds = \frac{2-\alpha_P}{2}((1+\gamma)^{t^*} - 1)$ for $t > t^*$.

Note that, then: If $t_1 < t^*$,

$$-\int_{s\le t_1} \tilde{\eta}_s ds + \tilde{\eta}_{t_1} \frac{1-p-p\alpha}{\lambda(1-p)} = \frac{2-\alpha_P}{2}.$$

Therefore, the principal is indifferent for any $q_t((\phi, t_1t_2)_t)$ if $t_1, t_2 < t^*$.

Once the multipliers are given as above, the rest of the first-order conditions are obtained accordingly. This part is basically the same as in the text, so we omit it for brevity, but it confirms that our guess is indeed correct in that the above candidate policy is a second-best mechanism.

Finally, we verify the other ignored incentive constraints. First, in this mechanism, a voter with A type does not have any incentive to misreport his type, because doing so can only (weakly) decrease the probability of staying in policy A. A voter with B type does not have any incentive to misreport his type either, because doing so can only (weakly) increase the probability of staying in policy A. Therefore, this mechanism is incentive compatible for any type, and at any point in time.

C Revelation Principle in Continuous-time Mechanism Design

In this section, we show a revelation-principle result in a continuous-time dynamic mechanism design environment.

The key assumption is that each agent never learns a deviation of other agents, and in this sense, there is no "off-path" information set.[15] Hence, we consider (ex ante) Nash equilibrium as a natural solution concept, and show that, for any Nash equilibrium outcome given any (indirect) mechanism, there exists a direct mechanism where everyone's truth-telling is a Nash equilibrium.

Although the result is based on a relatively standard argument, it seems that there has not been any formal revelation-principle result in continuous-time environments. In what follows, we consider a general continuous-time mechanism design environment (which includes our main environment as a special case), hoping that this section can be read independently by readers who are interested in the revelation principle in other continuous-time environments.

C.1 Model and Result

There exist I agents. The time is continuous $t \in [0, \infty)$. It is useful to imagine that each agent i's private information is drawn all at time 0, $\theta_i = (\theta_{it})_t \in \Theta_i = \prod_t \Theta_{it}$, although his choice at each t can only depend on

[15]Except for trivial cases where he himself deviates, which is well-treated in our analysis.

$\theta_i^t = (\theta_{is})_{s\le t}$. Each Θ_{it} is a Polish space with Borel σ-algebra and $\theta = (\theta_i)_i$ is a stochastic process following a probability measure F.[16]

The principal makes a sequence of social choices, $a = (a_t)_t \in A = \prod_t A_t$, where A is a measurable space. Each agent i's (ex post) payoff in the game is given by a measurable function $u_i(\theta_i, a)$.

Formally, at the beginning of the game, the principal commits to a mechanism, comprising (i) a collection of message spaces, M_{it} for each i, t, each of which is a measurable space, and (ii) a collection of social choice rules, $\alpha_t((m_i^t)_i) \in A_t$ for each $t, (m_i^t)_i$, where $m_i^t = (m_{is})_{s\le t}$, and each α_t is a measurable mapping. We discuss its interpretation in the next paragraph. After observing the mechanism, each agent chooses a plan of the play, which is denoted by a collection of measurable mappings $\sigma_{it}(\theta_i^t) \in M_{it}$ for each t, θ_i^t. Let σ_i denote this collection, i.e., $\sigma_i = (\sigma_{it}(\theta_i^t))_{t,\theta_i^t}$. [17]

Informally, one can interpret that the above notation reflects the following timing of the game: At each time t, first, agent i observes a new piece of his private information $\theta_{it} \in \Theta_{it}$, and then he privately sends a message $m_{it} \in M_{it}$ to the principal. Next, the principal makes a social decision of that period as a function of the collected messages up to t (and of the past social choices, which is irrelevant here because α_s, $s < t$, are all deterministic): $\alpha_t((m_i^t)_i) \in A_t$. Then the game proceeds to the "next instance of time".

Given α and $\sigma = (\sigma_i)_i$, together with F, a joint probability measure for (θ, a) is induced, and hence i's expected payoff:

$$\mathbb{E}[u_i(\theta, a)|\alpha, \sigma].$$

Definition 1. *A profile of the player's plans of the play, $\sigma = (\sigma_i)$, is a Nash equilibrium given α, if for all i and any plan of the play of i, σ_i', we have:*

$$\mathbb{E}[u_i(\theta, a)|\alpha, \sigma] \ge \mathbb{E}[u_i(\theta, a)|\alpha, \sigma_i', \sigma_{-i}].$$

Definition 2.
1. *A mechanism is direct if each $M_{it} = \Theta_{it}$.*
2. *A plan of the play of agent i, σ_i, is truth-telling if $\sigma_{it}(\theta_i^t) = \theta_{it}$ for all $t, \theta_i^t = (\theta_{is})_{s\le t}$.*
3. *A direct mechanism, represented by α, is incentive compatible if every agent's truth-telling is a Nash equilibrium given α.*

Theorem 2. *Fix α and σ so that σ is a Nash equilibrium given α. Then, there exists a direct mechanism α^* such that (i) truth-telling σ^* is a Nash equilibrium (i.e., α^* is incentive compatible) and (ii) $(\alpha, \sigma; F)$ and $(\alpha^*, \sigma^*; F)$ induce the same joint distribution for (θ, a).*

Proof. Define α^* so that:

$$\alpha_t^*((\theta_i^t)_i) = \alpha_t((\sigma_{it}(\theta_i^t))_i),$$

for each t and $(\theta_i^t)_i$. Then (ii) is immediate.

To show (i), consider a deviation of agent i in the direct mechanism α^*. Any possible deviation is identified by a measurable mapping $\eta_i : \Theta_i \to \Theta_i$, where its i-th component is denoted by $\eta_{it}(\theta_i^t) \in \Theta_{it}$, interpreted as i's (mis)reported type at t given his true type realization up to t (i.e., θ_i^t), reflecting the restriction that i's choice

[16] Thus, σ-algebra over Θ_i is given by any finite cylinders. For a countable Cartesian product of measurable spaces, we always consider product σ-algebra.

[17] Implicitly, I assume the least informed setting for the agents: he does not observe anything except for his type realization. To show the revelation principle, this seems the easiest situation.

at time t can only depend on θ_i^t (not entire θ_i). Let $\eta_i^t(\theta_i^t) = (\eta_{is}(\theta_i^s))_{s\leq t}$. The joint distribution for (θ, a) given $(\alpha^*, \tau_i, \sigma^*_{-i}; F)$ is the same as that given $(\alpha, \sigma'_i, \sigma_{-i}; F)$, where $\sigma'_i = (\sigma'_{it})_t$ with:

$$\sigma'_{it}(\theta_i^t) = \sigma_{it}(\eta_i^t(\theta_i^t)),$$

for each t. However, because σ is a Nash equilibrium given α, it is not profitable for i to deviate from σ_i to σ'_i. Therefore, it is not profitable for i to deviate from truth-telling to τ_i. □

References

Albrecht, J., A. Anderson, and S. Vroman (2010, July). Search by committee. *Search Theory and Applications 145*(4), 1386–1407.

Ali, S. N., M. Mihm, and L. Siga (2017). The Perverse Politics of Polarization. SSRN Scholarly Paper ID 2907489, Social Science Research Network, Rochester, NY.

Azrieli, Y. and S. Kim (2014). Pareto Efficiency and Weighted Majority Rules. *International Economic Review 55*(4), 1067–1088.

Balseiro, S. R., H. Gurkan, and P. Sun (2019). Multiagent Mechanism Design Without Money. *Operations Research 67*(5), 1417–1436.

Barelli, P., S. Bhattacharya, and L. Siga (2020). Full information equivalence in large elections. *Available at SSRN 3183959*.

Battaglini, M. (2016, 11). Public Protests and Policy Making. *The Quarterly Journal of Economics 132*(1), 485–549.

Bergemann, D. and J. Välimäki (2010). The dynamic pivot mechanism. *Econometrica 78*(2), 771–789.

Bhattacharya, S. (2013). Preference monotonicity and information aggregation in elections. *Econometrica 81*(3), 1229–1247.

Black, D. (1958). *The Theory of Committees and Elections.* Springer.

Börgers, T. (2004). Costly Voting. *The American Economic Review 94*(1), 57–66.

Callander, S. (2011, October). Searching and Learning by Trial and Error. *American Economic Review 101*(6), 2277–2308.

Casella, A. (2005, May). Storable votes. *Special Issue in Honor of Richard D. McKelvey 51*(2), 391–419.

Chakravarty, S., T. R. Kaplan, and G. Myles (2018). When costly voting is beneficial. *Journal of Public Economics 167*, 33–42.

Chan, J., A. Lizzeri, W. Suen, and L. Yariv (2018, April). Deliberating Collective Decisions. *The Review of Economic Studies 85*(2), 929–963.

Compte, O. and P. Jehiel (2010). Bargaining and Majority Rules: A Collective Search Perspective. *Journal of Political Economy 118*(2), 189–221.

Damiano, E., H. Li, and W. Suen (2012). Optimal deadlines for agreements. *Theoretical Economics 7*(2), 357–393.

Damiano, E., H. Li, and W. Suen (2021). Optimal delay in committees. Technical report.

Dekel, E. and M. Piccione (2000). Sequential voting procedures in symmetric binary elections. *Journal of Political Economy 108*(1), 34–55.

Ehlers, L., D. Majumdar, D. Mishra, and A. Sen (2020, May). Continuity and incentive compatibility in cardinal mechanisms. *Journal of Mathematical Economics 88*, 31–41.

Ekmekci, M. and S. Lauermann (2019). Informal elections with dispersed information. Technical report, University of Bonn and University of Mannheim, Germany.

Escobar, J. F. and Q. Zhang (2020). Delegating learning. *Theoretical Economics, Forthcoming*.

Feddersen, T. and W. Pesendorfer (1997). Voting behavior and information aggregation in elections with private information. *Econometrica 65*(5), 1029–1058.

Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, 1146–1155.

Forges, F. (1986). An Approach to Communication Equilibria. *Econometrica 54*(6), 1375–1385.

Gersbach, H. (1993). Politics and the Choice of Durability: Comment. *The American Economic Review 83*(3), 670–673.

Ghosal, S. and B. Lockwood (2009). Costly voting when both information and preferences differ: Is turnout too high or too low? *Social Choice and Welfare 33*(1), 25–50.

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica*, 587–601.

Grenadier, S. R., A. Malenko, and N. Malenko (2015). Timing decisions in organizations: Communication and authority in a dynamic environment. *Available at SSRN*.

Grüner, H. P. and T. Tröger (2019). Linear Voting Rules. *Econometrica 87*(6), 2037–2077.

Guo, Y. (2016). Dynamic delegation of experimentation. *American Economic Review 106*(8), 1969–2008.

Guo, Y. and J. Hörner (2020). Dynamic allocation without money.

Jackson, M. O. and H. F. Sonnenschein (2007, January). Overcoming Incentive Constraints by Linking Decisions. *Econometrica 75*(1), 241–257.

Keller, G., S. Rady, and M. Cripps (2005). Strategic experimentation with exponential bandits. *Econometrica*, 39–68.

Khan, U. and M. B. Stinchcombe (2015). The virtues of hesitation: Optimal timing in a non-stationary world. *The American Economic Review 105*(3), 1147–1176.

Kim, S. (2017, July). Ordinal versus cardinal voting rules: A mechanism design approach. *Games and Economic Behavior 104*, 350–371.

Krasa, S. and M. K. Polborn (2009). Is mandatory voting better than voluntary voting? *Games and Economic Behavior 66*(1), 275–291.

Krishna, V. and J. Morgan (2015). Majority rule and utilitarian welfare. *American Economic Journal: Microeconomics 7*(4), 339–75.

Majumdar, D. and A. Sen (2004). Ordinally Bayesian Incentive Compatible Voting Rules. *Econometrica 72*(2), 523–540.

Messner, M. and M. K. Polborn (2012, June). The option to wait in collective decisions and optimal majority rules. *Journal of Public Economics 96*(5), 524–540.

Moldovanu, B. and F. Rosar (2021, November). Brexit: A comparison of dynamic voting games with irreversible options. *Games and Economic Behavior 130*, 85–108.

Murto, P. and J. Välimäki (2013). Delay and information aggregation in stopping games with private information. *Journal of Economic Theory 148*(6), 2404–2435.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica: Journal of the Econometric Society*, 323–358.

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory 10*(2), 187–217.

Stanley, R. P. (2015). *Catalan numbers.* Cambridge University Press.

Stinchcombe, M. B. (1992). Maximal strategy sets for continuous-time game theory. *Journal of Economic Theory 56*(2), 235–265.

Strulovici, B. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica 78*(3), 933–971.

Sugaya, T. and A. Wolitzky (2020, 08). The Revelation Principle in Multistage Games. *The Review of Economic Studies*.