Conditional Retrospective Voting
in Large Elections*

Ignacio Esponda       Demian Pouzo
(NYU Stern)            (UC Berkeley)

January 2, 2012

Abstract

We propose a new model of elections where a continuum of voters have private information about the uncertain performance of electoral alternatives. A (retrospective) voting equilibrium formalizes the idea that a large number of voters choose between alternatives (e.g., political parties) based on their previously observed performance. This equilibrium notion provides a tractable framework to study several implications of retrospective voting regarding information aggregation, optimal electoral rules, the value of information, endogenous preferences, and party polarization. We also provide a game-theoretic foundation for our voting equilibrium by studying elections where the (finite) number of voters goes to infinity.

* A previous version of this paper was circulated under the title “Information aggregation under non-strategic voting”. We thank Heski Bar-Isaac, Ernesto Dal Bó, Erik Eyster, Frederico Finan, Gregory Huber, Navin Kartik, Alessandro Lizzeri, Santiago Oliveros, Andy Skrzypacz, several seminar participants, and especially Wolfgang Pesendorfer for their comments. Esponda: 44 W Fourth Street KMC 7-76, New York, NY 10012, iesponda@stern.nyu.edu; Pouzo: 508-1 Evans Hall #3880, Berkeley, CA 94704, dpouzo@econ.berkeley.edu.
1 Introduction

The extent to which markets aggregate private information has received considerable
attention at least since Hayek (1945). In many contexts, however, allocations are
determined not by the price mechanism but rather by some voting mechanism. Feddersen and Pesendorfer (1997) provide the seminal treatment of information aggregation
in large elections under the assumption that voters play a Nash equilibrium, which
embodies the requirement that voters use not only their own private information but
also the information that they can infer from other voters from the hypothetical event
that their vote is pivotal. In this paper, we study information aggregation in elections
when voters are boundedly rational and do not account for others’ information.

The formal political economy literature usually portrays voters as sophisticated indi-
viduals who understand the ideological and policy positions of political parties, can
solve complicated signal-extraction problems, and have rational expectations about
future policies. While this view of voters has led to many insights about the role of
political elections in the economy, it does contrast sharply with a long-standing and
robust empirical finding in political science: electorates are often poorly informed and
have little understanding of ideology and policy.\footnote{E.g., Delli Carpini and Keeter (1997) and Converse (2000).}

How can elections provide accountability with such modest levels of sophistication
in the electorate? Political scientists have long emphasized that retrospective voting
may be sufficient for satisfactory democratic performance (Key, 1966). Under retro-
spective voting, voters reward or punish politicians and their parties based on their
past performance. In the words of Fiorina (1981, p. 5), voters “need not know the
precise economic or foreign policies of the incumbent administration in order to see
or feel the results of those policies.” While retrospective voting has received empiri-
cal support (e.g., Kramer (1971), Fiorina (1978), Lewis-Beck and Stegmaier (2000)),
there have been few attempts to incorporate it into voting models.\footnote{An early exception is Nordhaus’ (1975) theory of political business cycles. Our model follows
Down’s (1957) view of retrospective voting as a way to predict how parties will perform in the future rather than as a way to simply punish or reward the party for past performance (Fiorina (1981), Chapter 1). Barro (1973), Ferejohn (1986), and Maskin and Tirole (2004) follow a complementary,
principal-agent approach to accountability. Also, see Bendor et al. (2010) for a (non-equilibrium)
model of retrospective voting in political science, and Bendor et al. (2011) for a textbook treatment
of bounded rationality in elections.}

More generally, the idea that voters repeatedly decide between the same set of
issues or alternatives by learning from past outcomes seems natural in many different types of elections.\textsuperscript{3} Several important questions about retrospective voting are hard to answer without a theoretical framework. To what extent is information aggregated? Is there a mechanism inherent to elections that mitigates the kind of mistakes that a boundedly-rational electorate can be expected to make? Does better information result in better electoral outcomes? In which cases are majority and supermajority rules desirable? In a political context, do parties have an incentive to mitigate or exacerbate their ideological positions under retrospective voting?

We provide a new framework to answer these questions. The setup is based on the ‘jury model’ (Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997, 1998)), which captures the realistic feature that voters are uncertain about the payoffs obtained from electing each alternative. In each election, voters observe information correlated with a randomly drawn state of the world and then simultaneously cast a vote for one of two alternatives. The payoffs are increasing in the state of the world for one alternative and decreasing for the other. For example, consider an election between two parties. The state of the world can represent whether the economy is overheated or in a recession, and one of the parties may be better at dealing with an overheated economy and the other better at dealing with a recession (perhaps because of their different attitudes toward monetary and fiscal policy).

We propose a new solution concept, \textit{(retrospective) voting equilibrium}, to formalize retrospective voting in large elections. We also provide a game-theoretic foundation to justify this solution concept. To illustrate the foundation, suppose that an election takes place every period by drawing a new state from a fixed distribution. In every period, voters keep track of the performance of each alternative conditional on the information they observe in that period. Voters then vote for the alternative that has delivered the best past performance. A voting equilibrium represents the limit, as the size of the electorate grows to infinity, of the steady-states corresponding to this dynamic environment.\textsuperscript{4}

A voting equilibrium consists of a strategy profile and an election cutoff that satisfy

\textsuperscript{3}For example, voters repeatedly choose between the same political parties in Presidential, Congressional, and local elections; legislators choose whether to vote with or against their affiliated party; shareholders vote on takeover proposals; union members vote to accept or reject negotiated contracts; and residents vote whether to approve additional funding for school districts.

\textsuperscript{4}Esponda and Pouzo (2011) show that these steady-states are captured by the behavioral equilibrium introduced by Esponda (2008).
two conditions. First, a tie occurs at the election cutoff, so that one alternative is
chosen for states higher than the election cutoff and the other alternative is chosen for
lower states. Second, the strategy profile must be optimal given the election cutoff.
Optimality here is defined in terms of retrospective voting: Voters’ perceptions of the
benefits of each alternative derive from the observed performance of each alternative,
which depends on the states in which each alternative is elected, and, therefore, on
the election cutoff. Our definition of a voting equilibrium has some parallels with the
definition of a competitive equilibrium in a market economy. In our context, the role
of prices is played by the election cutoff. And the fact that the number of voters is
sufficiently large means that each voter has a negligible effect on the election cutoff,
which is therefore taken as given by the electorate but is endogenously determined by
the strategies of all voters. The fact that we can characterize retrospective behavior
in large elections in this parsimonious manner is a major advantage of our framework.

Retrospective voting has been criticized for its simplicity. Stigler (1973, p. 164)
argues that it is not sensible for the electorate to punish the party in power for “de-
velopments (e.g., a foreign recession) beyond the powers or responsibilities of the
party”\textsuperscript{5}. Motivated by such critiques, our theory of retrospective voting incorporates
more sophisticated behavior by letting voters condition their behavior on informa-
tion correlated with the state of the world—thus we refer to our theory as conditional
retrospective voting. In a political context, information may include campaign plat-
forms, policy announcements, information about the state of the national or foreign
economy, or any other signal correlated with party performance.\textsuperscript{6}

By allowing voters to condition on their information, our model emphasizes an
important, but often overlooked, aspect of retrospective voting: Voters learn to make
decisions from past experience but counterfactuals are not observed. Consequently,
learning suffers from a selection problem. For example, while voters can judge the
elected party’s performance, they do not observe how the losing party would have
performed if elected. But, since the electorate uses valuable information to make
decisions, then parties are not randomly elected into office, and, therefore, voters’
perceptions about performance are likely to be biased. This bias casts doubts on the

\textsuperscript{5}Achen and Bartels (2004), Leigh (2009), and Wolfers (2009) find that voters punish politicians
for events that are outside of their control, though Healy and Malhotra (2010) find that punishment
is related to the politician’s response to these events.

\textsuperscript{6}For example, Campbell et al. (2010) find that retrospective evaluations depend on whether the
candidate is an incumbent or a successor for the in-party.
capacity of retrospective voting to produce good outcomes.

We use the framework to study some implications of retrospective voting. Efficient aggregation of information usually fails in the sense that mistakes (i.e., choosing the wrong alternative) happen in equilibrium. To understand why mistakes happen, suppose that party $A$ is best if the underlying (unobservable) state of the economy is strong, while party $B$ is best if the economy is weak. If equilibrium were efficient, so that party $A$ were elected in a strong economy and party $B$ in a weak economy, then voters would always observe party $A$ performing better than party $B$ (since it is easier to govern in a strong economy). Hence, all voters would prefer to vote for party $A$, thus contradicting the hypothesis that the right party is chosen in its corresponding state of the world. In equilibrium, party $A$ will have to be occasionally elected into office in a weak economy; this mistake will then reduce party $A$’s popularity and provide incentives for voters to choose both parties in equilibrium.

In richer environments, with multiple states of the world, information may be efficiently aggregated (we provide necessary and sufficient conditions). But the general point remains: If the alternatives are asymmetric, in the sense that one alternative performs much better than the other when compared in the states of the world in which these alternatives perform best, then equilibrium must be characterized by mistakes. In particular, retrospective voting yields certain convergence, not necessarily in policy positions or in potential performance, but rather in observed performance.

One concern is that, if mistakes are costly, then retrospective voting may perform poorly. We show that this concern is overstated: In equilibrium, the probability of mistakes adjusts in order to mitigate the effect of more costly mistakes. In the previous example, suppose that party $A$’s performance is disastrous in a weak economy (say, because $A$’s typical policies are particularly damaging in that state). Then, very few mistakes are sufficient to decrease $A$’s popularity and induce voters to also vote for $B$. The fact that equilibrium retrospective voting inherently limits the expected damage that results from choosing the wrong party may explain why elections perform relatively well despite the lack of sophistication of the electorate.

The quality of information and electoral rules play an important role in determining the amount of mistakes in equilibrium. In particular, a more informed electorate need not guarantee better electoral outcomes under retrospective voting. The intuition is that better information makes players more confident in voting for an alternative when they receive a stronger signal in favor of that alternative. The problem
is that, as long as information is imperfect, a large enough fraction of voters may still observe those stronger signals by mistake, and their higher level of confidence may result in a higher amount of mistakes. The model can also explain why societies are conservative and require supermajorities to adopt risky alternatives. In the previous example, suppose that party $B$’s performance does not depend on the state of the world but that party $A$’s performance is risky. Since the risky alternative tends to be chosen when the economy is strong, then voters will overestimate its true performance and a supermajority rule prevents the risky party $A$ from being chosen too often. Thus, the biases that arise under retrospective voting provide an alternative justification for the common use of supermajority rules for changes in the status quo (e.g., raising taxes and other legislative proposals, shareholder voting on takeovers, and constitutional changes).

The model can also be used to endogenize people’s preferences over the alternatives. For example, in the political context, we can study the relationship between information and endogenous party affiliation. We find that conditioning on better information decreases the amount of partisanship in the electorate. In addition, given that partisanship is endogenously determined in equilibrium, a change in the information of some types of voters may affect the voting patterns of other types. An implication is that one should control for the preferences of other voters when empirically investigating the relationship between information and voting behavior.

Finally, we study whether parties have incentives to mitigate or exacerbate their ideological differences by analyzing a game where parties strategically choose policies to maximize their chance of being elected. Unlike the standard environment in the median voter theorem (Downs, 1957), voters do not consider a party or policy always better than another but, rather, preferences are state-dependent. We show that, due to retrospective voting, each party has an incentive to choose relatively extreme policies that work well in those states in which it is elected into office, since those are the states that retrospective voters use to evaluate their performance. This is true even if all voters would prefer a neutral policy that delivers a state-independent payoff. In particular, our model may partly explain the well-documented political

---


8 In the spirit of Downs (1957) and Fiorina (1981), partisanship is endogenously determined by parties’ past performance. Gerber and Green (1998) provide a model of endogenous party affiliation.
polarization in the United States.9

Most of the literature studies voting under uncertainty under the assumption that voters play a Nash equilibrium (Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997, 1998)).10 In the context that we consider, Feddersen and Pesendorfer (1997) show that information is efficiently aggregated in a Nash equilibrium of large elections for any non-unanimous electoral rule.11 In particular, the quality of information and the electoral rule play essentially no role in the final outcome. Also, the large electorate asymptotics under a Nash equilibrium do not produce the type of parsimonious and tractable framework that we obtain under retrospective voting.12 Finally, the strategic (i.e., Nash) literature originated as a response to the earlier assumption that people vote sincerely. Our retrospective voting equilibrium can be viewed as a notion of sincere voting where beliefs are determined endogenously in equilibrium.

More generally, this paper follows a recent literature that studies game-theoretic equilibrium concepts for boundedly rational players (e.g., Osborne and Rubinstein (1998), Jehiel (2005), Eyster and Rabin (2005), Jehiel and Samet (2007), Jehiel and Koessler (2008), and Esponda (2008))).13 Also, our foundation of retrospective voting equilibrium is in the spirit of the literature that seeks foundations for rational expectations equilibrium (e.g., Milgrom (1981b), Pesendorfer and Swinkels (1997), Reny and Perry (2006), and Vives (2011)).

In Section 2, we illustrate our notion of retrospective voting, the selection problem faced by voters, and how our solution concept differs from Nash equilibrium. In Section 3, we present the framework and the solution concept of a voting equilibrium. In Section 4, we study some implications of retrospective voting. In Section 5, we present a game-theoretic foundation for the notion of a voting equilibrium. We conclude in Section 6 by mentioning some limitations of our work and possible extensions.

---


10Guarnaschelli et al. (2000) provide early experimental evidence on Nash equilibrium in small committees, while Esponda and Vespa (2011) recently find that most subjects fail to engage in pivotal voting in the lab. See Huber et al. (2011) for an experiment on retrospective voting.

11Feddersen and Pesendorfer (1997) show that information may fail to aggregate with multidimensional state variables.

12Myerson’s (1998) extended Poisson model makes these large games more tractable.


## 2 Illustration of retrospective voting

In this section, we illustrate the behavioral assumptions underlying our solution concept and some implications that are important for understanding the intuition behind our results. Suppose that a particular voter has participated in 8 past elections between parties $A$ and $B$. Figure 1 depicts her history, including her signal ($a$ or $b$), her vote ($A$ or $B$), the outcome of the election ($A$ or $B$), and the performance of the party in power. The signals represent information that is correlated with each party’s performance, such as campaign platforms, economic indicators, the possibility of international conflict, or even information about variables that are outside the control of politicians. For simplicity, suppose that the voter knows that party $B$ always delivers a constant payoff of 0, while she is uncertain about the performance of party $A$.

We postulate the following retrospective behavior. Suppose that in the 9th election, the voter observes signal $a$. First, she forms beliefs about the expected performance of party $A$ conditional on information $a$. These beliefs are given by the average observed payoff obtained from $A$ when her signal was $a$, which in this case is \((-1 + 1 + 1)/3 = 1/3\). Second, the player votes for the party that she believes has the highest expected payoffs: in this case $1/3 > 0$ and, therefore, she votes for $A$.

Esponda and Pouzo (2011) show that the steady-states of this dynamic environment, when all voters behave in the way described above, correspond to the notion of behavioral equilibrium (Esponda, 2008), which differs from Nash equilibrium. In fact, retrospective voting does not take into account two sources of sample selection. The first source is exogenous: Since counterfactuals are not observed (e.g., periods 3 and 7 in Figure 1), estimates are likely to be biased upwards since, given that people use their private information to make decisions, alternatives then tend to be chosen when they are most likely to perform well. The second source is endogenous: The vote of a player, whenever pivotal, affects the sample that she will observe and, therefore, her beliefs. In both the exogenous and endogenous cases, the underlying source of the bias is that other players use their private information to make decisions. Failing to account for selection in a learning environment is, then, analogous to failing to account for the informational content of other players’ actions.

In this paper, we characterize the steady-states of this dynamic environment as the size of the electorate grows to infinity, and we refer to the resulting solution con-
cept as a (retrospective) voting equilibrium. We show that, as the electorate grows, the probability of being pivotal becomes negligible. Therefore, the endogenous selection problem disappears. However, the exogenous selection problem remains, thus explaining why our solution concept differs from Nash equilibrium even when the electorate becomes large. In particular, while in a Nash equilibrium voters learn from “marginal” events in which a party gets elected by the narrowest margin, in a voting equilibrium voters also learn from the “inframarginal” events in which a party gets elected even if the election was not close. While the entire sequence (as the electorate size increases) of equilibria affects inferences drawn from the marginal event, only the limit of the sequence matters for inframarginal events. Therefore, retrospective voting has a straightforward characterization in large elections.

3 Voting framework

A continuum of voters participate in an election to choose between alternatives (e.g., parties) $A$ and $B$. A state $\omega \in \Omega$ is first drawn according to a probability distribution $G$ and, conditional on the state, each player observes an independently-drawn private signal. Players then simultaneously submit a vote for either $A$ or $B$. Votes are aggregated according to an electoral rule $\rho \in (0, 1)$: Alternative $A$ wins the election if
the proportion of votes in favor of $A$ is greater than $\rho$; otherwise, $B$ wins the election.

We model heterogeneity (in preferences and information) by assuming that each voter is of a particular type $\theta \in \Theta$, where $\Theta$ is a finite set of types and $\phi(\theta) \in (0,1)$ is the proportion of voters of type $\theta$. Conditional on a state $W = \omega$, players of type $\theta$ independently draw a signal $S_\theta = s$ from a finite, nonempty set $S_\theta \subset \mathbb{R}$ with probability $q_\theta(s | \omega)$; they also obtain utility $u_\theta(o, \omega)$, where $o \in \{A, B\}$ is the elected alternative. Let $s^L_\theta \equiv \min\{s : s \in S_\theta\}$ and $s^H_\theta \equiv \max\{s : s \in S_\theta\}$ denote the lowest and highest signals of type $\theta$. We maintain the following assumptions throughout the paper, for all types $\theta \in \Theta$:

**A1.** $\Omega = [-1,1]$ and $G$ is an absolutely continuous probability distribution over $\Omega$ with density $g$.

**A2.** (i) $u_\theta(A, \cdot) : \Omega \to \mathbb{R}$ is nondecreasing and $u_\theta(B, \cdot) : \Omega \to \mathbb{R}$ is nonincreasing, and one of them is strictly monotone; (ii) $u_\theta(A, \cdot)$ and $u_\theta(B, \cdot)$ are both continuously differentiable and bounded.

**A3.** MLRP: For all $\omega' > \omega$, and $s' > s$:

$$\frac{q_\theta(s'|\omega')}{q_\theta(s'|\omega)} - \frac{q_\theta(s|\omega')}{q_\theta(s|\omega)} > 0.$$

**A4.** (i) $\inf_\Omega g(\omega) > 0$; (ii) there exists $d > 0$ such that $q_\theta(s|\omega) > d$ for all $s \in S_\theta$ and $\omega \in \Omega$; (iii) $q_\theta(s | \cdot)$ is continuous for all $s \in S_\theta$.

Assumptions A1-A3 provide an ordering between states, information, and players’ preferences.\footnote{It is important that $u_\theta(A, \cdot)$ and $u_\theta(B, \cdot)$ are separately increasing; it does not suffice that their difference is increasing. In addition, the interval state space yields a convenient cutoff characterization of equilibrium; with a finite state space, we would also have to indicate the probability that each alternative is elected at the cutoff state.} Note that A3 is trivially satisfied for types with a unique signal (i.e., no private information). Assumption A4 requires densities to be uniformly bounded—in particular, “strong signals” (Milgrom, 1979) are ruled out—and, for simplicity in the statement of results, continuity of $q_\theta(s | \cdot)$.

Let $\sigma_\theta : S_\theta \to [0,1]$ denote the strategy of type $\theta$, where $\sigma_\theta(s)$ is the probability of voting for alternative $A$ after observing signal $s$. A strategy $\sigma_\theta$ is **nondecreasing** if $\sigma_\theta(s') \geq \sigma_\theta(s)$ for all $s' > s$. A strategy profile $\sigma = (\sigma_\theta)_{\theta \in \Theta}$ is nondecreasing if $\sigma_\theta$ is nondecreasing for each $\theta$. 
Let
\[ \kappa(\omega; \sigma) = \sum_{\theta \in \Theta} \phi(\theta) \sum_{s \in \mathcal{S}_\theta} q_\theta(s \mid \omega) \sigma_\theta(s) \]
denote the proportion of votes in favor of A. Assumption A3 implies that \( \kappa(\cdot; \sigma) \) is nondecreasing if \( \sigma \) is nondecreasing. In the case where the strategy depends on private information, so that \( \sigma \) is not flat, then \( \kappa(\cdot; \sigma) \) is increasing and the outcome of the election can be characterized by a cutoff: A is elected if and only if \( \kappa(\omega; \sigma) \geq \rho \), or, equivalently, for all sufficiently high states. This observation motivates the following definition.\(^{15}\)

**Definition 1.** A state \( c \in \Omega \) is an *election cutoff* given a strategy profile \( \sigma \) if \( \kappa(\omega; \sigma) \geq \rho \) for all \( \omega > c \) and \( \kappa(\omega; \sigma) \leq \rho \) for all \( \omega < c \).

When making her decision, each voter takes the cutoff as given. A cutoff determines the set of states for which each alternative is chosen, and, consequently, each voter’s evaluation of the benefits of electing each alternative. For a given cutoff \( c \in \Omega \), the conditional difference in benefits from electing A over B that is perceived by a voter of type \( \theta \) who observes signal \( s \) is

\[ v_\theta(s; c) \equiv E(u_\theta(A, W) \mid W \geq c, S_\theta = s) - E(u_\theta(B, W) \mid W \leq c, S_\theta = s). \quad (1) \]

To interpret the above expression, note that alternative A is elected whenever \( W \geq c \), so that a voter’s retrospective evaluation of A is given by the first term in the right hand side of (1). A similar interpretation holds for the second term.

The following definition captures the idea that each voter votes for the alternative with the highest perceived benefits.

**Definition 2.** A strategy profile \( \sigma \) is *optimal* given an election cutoff \( c \) if

\[ \sigma_\theta(s) \in \arg \max_{a \in [0,1]} a \cdot v_\theta(s; c) \]

for all \( s \in \mathcal{S}_\theta \) and \( \theta \in \Theta \).

\(^{15}\)When \( \sigma \), and, therefore, \( \kappa(\cdot; \sigma) \) are constant, this definition is motivated by the limiting case where signals satisfy MLRP but become uninformative; see Section 5.
By assumptions A1-A3, \( v_\theta(\cdot; c) \) is increasing. Therefore, any strategy that is optimal given some cutoff must be nondecreasing.

A voting equilibrium requires players to optimize given an election cutoff that is endogenously determined by players’ strategies. We provide a game-theoretic foundation for this solution concept in Section 5.

**Definition 3.** A (retrospective) voting equilibrium is a strategy profile \( \sigma \) and an election cutoff \( c \) such that: (i) \( \sigma \) is optimal given \( c \), and (ii) \( c \) is an election cutoff given \( \sigma \).

It is straightforward to characterize a voting equilibrium in terms of the primitives of the voting environment. For each type and signal, define the personal cutoffs

\[
c_\theta(s) \equiv \arg \min_{c \in \Omega} |v_\theta(s; c)|. \tag{2}
\]

Since \( \Omega \) is compact and \( v_\theta(s; \cdot) \) is continuous and increasing (by A1-A4), there exists a unique solution \( c_\theta(s) \) that is nonincreasing in \( s \). Let \( \zeta \equiv \min_\theta c_\theta(s^H_\theta) \) and \( \bar{c} = \max_\theta c_\theta(s^L_\theta) \) denote the lowest and highest personal cutoffs across all types.

If we knew the equilibrium election cutoff \( c^* \), then it would be straightforward to characterize the equilibrium strategy: a type \( \theta \) with signal \( s \) such that \( c_\theta(S_\theta) < c^* \) must have \( v_\theta(s; c^*) > 0 \) and, therefore, she will optimally vote for \( A \); similarly, if \( c_\theta(s) > c^* \), then she will optimally vote for \( B \).\footnote{Mixed strategies are characterized by \( \sum_{(\theta,s) : c_\theta(s) = c^*} \phi(\theta) q_\theta(s | c^*) \sigma_\theta(s) = \rho - \overline{\kappa}(c^*) \), where \( \overline{\kappa} \) is defined next.} Consequently, we now characterize the set of equilibrium cutoffs. For a possible election cutoff \( \omega \in \Omega \),

\[
\pi(\omega) \equiv \sum_{\theta \in \Theta} \phi(\theta) q_\theta(c_\theta(S_\theta) < \omega | \omega), \tag{3}
\]

may be interpreted as the proportion of players that vote for \( A \) conditional on the state being the election cutoff \( \omega \).\footnote{The interpretation is correct unless \( \omega \) is one of the personal cutoffs \( c_\theta(s) \) for some \( \theta, s \).}

**Lemma 1.** \( \overline{\kappa} : \Omega \to [0, 1] \) is nondecreasing, left-continuous, and satisfies: \( \overline{\kappa}(\omega) = 0 \) if \( \omega \leq \zeta \), \( \overline{\kappa}(\omega) \in (0, 1) \) if \( \zeta < \omega \leq \bar{c} \), and \( \overline{\kappa}(\omega) = 1 \) if \( \omega > \bar{c} \).
Proposition 1. The set of equilibrium election cutoffs under electoral rule $\rho \in (0, 1)$ is given by

$$C^{\text{eqm}}(\rho) \equiv \left[ \inf_{\omega \in \Omega} \{ \bar{\kappa}(\omega) \geq \rho \}, \sup_{\omega \in \Omega} \{ \bar{\kappa}(\omega) \leq \rho \} \right].$$

Moreover, the set is nonempty, so a voting equilibrium always exists.

Proof. See the Appendix.

Figure 2 depicts the function $\bar{\kappa}$ for two different environments. In the left panel, there is a unique voter type with two signals, $\bar{\kappa}$ is increasing, and the unique equilibrium cutoff is the state $c^*$ where $\bar{\kappa}$ intersects the electoral rule $\rho$. In the right panel, there are two types and any state in $[c^\theta(s^L_\theta), c^\theta(s^H_\theta)]$ is an equilibrium cutoff.

Corollary 1. For any $c \in [\underline{c}, \bar{c}]$, there exists $\rho \in (0, 1)$ such that $c$ is a voting equilibrium cutoff under electoral rule $\rho$.

Proof. Fix any $c \in (\underline{c}, \bar{c}]$ and let $\rho = \bar{\kappa}(c)$. By Lemma 1, $\rho \in (0, 1)$. By Proposition 1, $c \in C^{\text{eqm}}(\rho)$ under electoral rule $\rho$. For $c = \underline{c}$, Proposition 1 implies that $c \in C^{\text{eqm}}(\rho)$ under electoral rule $\rho = \lim_{c \searrow \underline{c}} \bar{\kappa}(c) \in (0, 1)$.

\[\]
4 Implications of Retrospective Voting

In this section, we apply the framework introduced above to study several issues discussed in the Introduction: mistakes and information aggregation (Section 4.1), optimal electoral rules and rationales for supermajority and (simple) majority rules (Section 4.2), the value of information and implications for welfare and partisanship (Section 4.3), endogenous mitigation of mistakes (Section 4.4), and endogenous party polarization (Section 4.5). We maintain assumptions A1-A4 from Section 3. In addition,

\[
W(c) = \sum_{\theta \in \Theta} \phi(\theta) \left[ E(u_\theta(A, W) \mid W > c) \Pr(W > c) + E(u_\theta(B, W) \mid W < c) \Pr(W < c) \right]
\]

denotes welfare as a function of the election cutoff \(c\). Let \(c^{FB}\) denote the first-best cutoff that maximizes welfare; it exists because \(\Omega\) is compact (A1) and \(W(\cdot)\) is continuous (by A2). For simplicity we make the additional assumption that \(c^{FB}\) is unique and interior, i.e., \(c^{FB} \in (-1, 1)\). We refer to the special case where every type has the same utility functions as a case with homogenous preferences. An implication of homogenous preferences is that \(W\) is single-peaked, i.e. \(W\) is increasing for all \(c < c^{FB}\) and decreasing for all \(c > c^{FB}\). The proofs of all results in Section 4 follow from the characterization results in Section 3 and are relegated to the Online Appendix.\(^{18}\)

4.1 Information aggregation

We now define information aggregation.

**Definition 4.** Information is aggregated if there exists an electoral rule \(\rho\) such that \(c^{FB}\) is an equilibrium cutoff, i.e., \(c^{FB} \in C^{eqm}(\rho)\).

Feddersen and Pesendorfer (1997) show that, if the planner’s preference coincides with the preference of the median (or any other percentile) voter, then the first-best outcome can be achieved with majority rule (or the corresponding percentile electoral rule) in a Nash equilibrium.\(^{19}\) In contrast, information may or may not

\(^{18}\)Available at http://people.stern.nyu.edu/iesponda/Ignacio_Esponda/Research.html.

\(^{19}\)Feddersen and Pesendorfer (1997) establish full-information equivalence, meaning that for any \(\rho\), there exists a Nash equilibrium outcome that coincides with the outcome the \(\rho\)-median voter would choose if the state were known by all voters. Since full-information equivalence does not usually
be aggregated in our context, therefore providing a role for information, electoral rules, and preferences that would not exist under Nash equilibrium. Corollary 1 immediately implies the following necessary and sufficient condition for information aggregation in terms of the primitives.

**Proposition 2.** Information is aggregated if and only if $c \leq c^{FB} \leq \bar{c}$.

What makes information aggregation difficult in a voting equilibrium is that players’ beliefs do not depend on their equilibrium strategies once we assume that the outcome is the first-best outcome.\textsuperscript{20} To see intuitively why $c^{FB} \leq \bar{c}$ is necessary for information aggregation, suppose that $c^{FB} > \bar{c}$. If information were aggregated, then, even after observing their lowest signals, all types would prefer to vote for A. But the fact that no one would vote for B contradicts the assumption that information is aggregated in the first place. Roughly speaking, a situation such as $c^{FB} > \bar{c}$, where information is not aggregated, occurs whenever the alternatives are asymmetric in the sense that one alternative performs much better than the other when compared in the states of the world in which each alternative performs best. A simple example is a status quo model where one alternative yields a risky payoff and the other alternative yields a state-independent payoff.

### 4.2 Optimal electoral rules

Inspection of Figure 2 reveals that different electoral rules can result in different voting equilibrium outcomes. We first characterize optimal electoral rules and then present rationales for majority and supermajority rules.

**Definition 5.** An electoral rule $\rho$ is *optimal* if there exists $c \in C^{eqm}(\rho)$ such that $W(c) \geq W(\tilde{c})$ for all $\tilde{c} \in \bigcup_{0<\rho<1}C^{eqm}(\rho)$.

Proposition 1 immediately implies the following characterization of optimal electoral rules in terms of the primitives of the voting environment.

\textsuperscript{20}In contrast, in a Nash equilibrium, even conditional on the first-best outcome, the information conveyed by the pivotal event depends on equilibrium strategies.
Proposition 3. Suppose that information is not aggregated and that \( W \) is single-peaked. If \( c < c_{FB} \), then \( \rho \) is optimal if and only if
\[
\rho \geq 1 - \sum_{\theta} \phi(\theta) q_{\theta}(c_{\theta}(S_{\theta}) = c \mid \bar{c})
\]
if \( c > c_{FB} \), then \( \rho \) is optimal if and only if
\[
\rho \leq \sum_{\theta} \phi(\theta) q_{\theta}(c_{\theta}(S_{\theta}) = c \mid \bar{c})
\]
Suppose that information is aggregated. Then, \( \rho \) is optimal if and only if
\[
\rho \in \left[ \bar{\pi}(c_{FB}), \lim_{\omega \downarrow c_{FB}} \overline{\pi}(\omega) \right]
\]

To understand Proposition 3, suppose that \( c_{FB} > c = c_{\theta'}(s_{\theta'}^L) \) in the right panel of Figure 2. In particular, information is not aggregated because, if it were, everyone would prefer to vote for \( A \), irrespective of their signal. For some type to be willing to vote for \( B \), the electorate must mistakenly elect \( A \) in states of the world where \( B \) would have been best; such mistakes make \( B \) more attractive. But mistakes carry a welfare cost. The lowest level of this mistake that still provides incentives for some type to play \( B \) is the mistake that makes the type with the highest personal cutoff, \( \theta' \), indifferent between \( A \) and \( B \) when observing its lowest signal. Given such indifference, there is at least a proportion \( 1 - \phi(\theta') q_{\theta'}(s_{\theta'}^L \mid \bar{c}) \) of players who would vote for \( A \) conditional on \( c \) being an equilibrium cutoff. But then, the electoral rule must be higher than the previous proportion if \( B \) is to be the outcome with positive probability. In addition, electoral rules that require a lower proportion to choose \( A \) also require a larger mistake in order to induce more people to vote for \( B \), so that \( B \) remains an outcome with positive probability. Since larger mistakes are associated with lower welfare, such electoral rules are suboptimal.

4.2.1 Simple majority and supermajority rules

Majority rule usually plays a prominent role in large elections. Condorcet (1785) provided an early informational rationale for (simple) majority rule by arguing that, if each individual makes the right choice with probability greater than \( 1/2 \), then the best alternative is elected under majority rule with probability that goes to one as the size of the electorate increases. The strategic-voting literature (Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997, 1998)) argues that, even if each individual could make the right decision with probability greater than \( 1/2 \), her strategic behavior may differ from her behavior as a decision maker. In fact, when voters are strategic (i.e., in a Nash equilibrium), the electoral rule is essentially irrelevant and Condorcet’s specific rationale for majority rule disappears.
In our setting, it is also the case that equilibrium behavior (which is endogenous) may differ from individual decision-making and, therefore, Condorcet’s rationale no longer holds. However, majority rule still has a special role provided that each type makes the right decision with probability greater than 1/2 in every state. We say that \( \sigma_\theta \) is an optimal decision policy for type \( \theta \) if it maximizes expected utility in the case where type \( \theta \) makes dictatorial decisions, i.e., for every \( s \in S_\theta \), \( \sigma_\theta(s) \in \arg \max_{x \in [0,1]} x \cdot E(u_\theta(A,W) - u_\theta(B,W) \mid S_\theta = s) \).

**Definition 6.** The voting environment satisfies Condorcet informativeness if the probability of voting for \( A \), \( \sum_{s \in S_\theta} \sigma_\theta(s) q_\theta(s \mid \omega) \), is strictly greater than 1/2 for all \( \omega > c^{FB} \) and strictly lower than 1/2 for all \( \omega < c^{FB} \), for every optimal decision policy \( \sigma_\theta \) and every type \( \theta \in \Theta \).

In our context, majority rule need not aggregate information and need not even be the optimal electoral rule. However, Proposition 4 shows that majority rule is the only rule that avoids worst-case outcomes. Intuitively, Condorcet informativeness implies that the proportion of votes for \( A \) is below \( 1/2 \) at \( c \) and above \( 1/2 \) at \( \overline{c} \), so that majority rule avoids extreme outcomes (at least one of which must be a worst-case outcome when \( W \) is single-peaked).

**Proposition 4.** Suppose that \( W \) is single-peaked. Then simple majority rule is the unique electoral rule that yields an equilibrium payoff strictly higher than the worst equilibrium payoff for every environment that satisfies Condorcet informativeness.

In many settings, the electorate chooses between an alternative that is safe (e.g., status quo) and an alternative that is risky. If there is sufficient heterogeneity in information types, then it is optimal to require a supermajority to choose the risky option, therefore providing a new rationale for conservatism.

**Proposition 5.** Consider a voting environment with homogeneous preferences where (i) \( B \) is a safe alternative, i.e., \( u(B, \omega) = u(B) \) for all \( \omega \in \Omega \), (ii) \( u(B) > E(u(A,W) \mid s^{B}_{\theta''}) \) for some \( \theta'' \in \Theta \), (iii) \( \phi(\theta) < 1/2 \) for all \( \theta \in \Theta \), and (iv) \( c_\theta(s^{B}_{\theta}) \neq c_\theta(s^{B}_{\theta'}) \) for all \( \theta \neq \theta'. \) Then any optimal electoral rule requires a supermajority (i.e., \( \rho > 1/2 \)) to elect the risky alternative \( A \), but falls short of aggregating information.
Intuitively, retrospective voters are biased in favor of the risky alternative. The reason is that they learn the value of the risky alternative by conditioning on all cases where the risky alternative is chosen, and these cases happen to be the ones where the risky alternative performs better than average. Electoral rules that are biased against the risky alternative mitigate this problem.

### 4.3 The value of information

We fix primitives \( \{(\Omega, g), u\} \) and perform comparative statics with respect to the information structure \( \mathcal{I} = (\mathcal{S}, q) \). We explicitly index previously-defined objects by their corresponding information structure. For simplicity, we assume that there is a unique type and drop the type subscripts. We follow Blackwell (1953) in saying that an information structure \( \mathcal{I}' = (\mathcal{S}', q') \) is more informative than \( \mathcal{I} = (\mathcal{S}, q) \) if \( \mathcal{I} \) is obtained by a garbling of \( \mathcal{I}' \).

**Definition 7.** \( \mathcal{I}' = (\mathcal{S}', q') \) is more informative than \( \mathcal{I} = (\mathcal{S}, q) \) if there exists an \( |\mathcal{S}'| \times |\mathcal{S}| \) matrix \( M \) with entries \( m_{s's} \) that satisfy \( \sum_{s' \in \mathcal{S}} m_{s's} = 1 \) for all \( s' \in \mathcal{S}' \) and \( q(s' | \cdot) = \sum_{s' \in \mathcal{S}} m_{s's} q'(s' | \cdot) \) for all \( s \in \mathcal{S} \). If, in addition, \( |\mathcal{S}'| \geq 2 \) and \( m_{s's} > 0 \) for all \( s, s' \), then we say that \( \mathcal{I}' \) is strictly more informative than \( \mathcal{I} \).

By Blackwell (1953), \( \mathcal{I}' \) is more informative than \( \mathcal{I} \) if and only if, in any environment \( \{(\Omega, g), u\} \), the expected utility of an individual making dictatorial decisions is (weakly) higher under \( \mathcal{I}' \) compared to \( \mathcal{I} \). Our objective is to understand the effect of more informative signals on the outcome of a voting equilibrium. The following result is key.

**Lemma 2.** If \( \mathcal{I}' \) is more informative than \( \mathcal{I} \), then the personal cutoffs under \( \mathcal{I}' \) are more extreme than under \( \mathcal{I} \), i.e., \( c(\mathcal{I}') \leq c(\mathcal{I}) \) and \( \overline{c}(\mathcal{I}') \geq \overline{c}(\mathcal{I}) \). These inequalities are strict if \( \mathcal{I}' \) is strictly more informative than \( \mathcal{I} \) and the personal cutoffs satisfy \( c(s; \mathcal{I}) \in (-1, 1) \) for all \( s \in \mathcal{S} \).

The idea behind Lemma 2 is that more informative information structures widen the difference in perceived utility between the lowest and highest signals, and, consequently, the range of personal cutoffs.
4.3.1 Information and welfare

The view that better information must help ignores the aggregate effect of information in an environment where people learn retrospectively. In fact, the next proposition shows that more information may decrease welfare.

**Proposition 6.** Fix any environment with information structures \( I', I \) such that \( I' \) is strictly more informative than \( I \) and \( c(s; I) \in (-1, 1) \) for all \( s \in S \). Then there exists an electoral rule \( \rho \) such that the highest equilibrium welfare under \( I' \) is lower than the lowest equilibrium welfare under \( I \).

Figure 3 illustrates this result. Originally, type \( \theta_3 \) has no information and \( c < c^{FB} \) is the equilibrium cutoff—the other types are there to illustrate a future result. When type \( \theta_3 \) becomes informed, those observing \( s^{IH} \) are more confident that \( A \) is best. Therefore, they are willing to vote for \( A \) even if \( A \) is elected for low states of the world, and their personal cutoff decreases to \( c_{\theta_3}(s^{IH}) \). The problem is that, since information is not perfect, a (small) fraction of voters will observe \( s^{IH} \) in low states of the world and incorrectly vote for \( A \). The new equilibrium cutoff moves further away from the first-best cutoff to \( c' < c \), and welfare decreases.

**Proposition 7.** Fix any environment with information structures \( I', I \) such that \( I' \) is strictly more informative than \( I \) and \( c(s; I) \in (0, 1) \) for all \( s \in S \). Then there exists an electoral rule \( \rho^* \) such that the highest equilibrium welfare under \( I' \) is weakly higher than the highest equilibrium welfare under \( I \) with any electoral rule. Moreover, welfare is strictly higher if the first-best outcome cannot be achieved under \( I \).

Proposition 7 shows that the intuition that more information increases welfare is correct as long as the electoral rule is chosen optimally. But there are two reasons to be cautious about the value of more information in practice. First, the electoral rule is often fixed. Second, implementing the optimal rule requires precise knowledge of the primitives (see Section 4.2).
4.3.2 Information and partisanship

We show that a voter becomes less partisan (in the sense of always voting for the same party) when her signal structure is more informative. Nevertheless, a change in one voter’s information structure may also affect the partisanship of other voters.

**Definition 8.** Given a strategy profile $\sigma$, type $\theta$ is a partisan for party A (B) if $\sigma_\theta(s) = 1 (0)$ for all $s \in S_\theta$.

**Proposition 8.** Let $\mathcal{I} = (I_\theta)_{\theta \in \Theta}$ and $\mathcal{I}' = (I'_\theta)_{\theta \in \Theta}$ be information structures such that $I'_\theta$ is more informative than $I_\theta$ for some $\theta^* \in \Theta$ and $I_\theta = I'_\theta$ for all $\theta \neq \theta^*$. Suppose, in addition, that $c_\theta(s; \mathcal{I}) \neq c_\theta(s^H_\theta; \mathcal{I})$ for all $s \in S_\theta$, $\theta \neq \theta^*$. Then, if $(\sigma, c)$ with $c \in (-1, 1)$ is an equilibrium under $\mathcal{I}$ such that type $\theta^*$ is a partisan for A (B), there cannot be an equilibrium under $\mathcal{I}'$ such that type $\theta^*$ is a partisan for B (A).

Figure 3 illustrates this result. Originally, none of the three types have private information and $c$ is the equilibrium cutoff. In particular, type $\theta_3$’s personal cutoff is above $c$ and, therefore, she is a partisan for $B$ in this equilibrium. When type $\theta_3$ becomes informed, her personal cutoffs become more extreme, as predicted by Lemma 2. Then the equilibrium cutoff becomes $c'$ and type $\theta_3$ becomes an “independent”, i.e., she votes for $A$ after observing $s^H$ and for $B$ after $s^L$. In general, the new equilibrium cutoff must lie to the left of the original cutoff and, therefore, type $\theta_3$ cannot become
a partisan for the other party. Figure 3 illustrates an additional point. In the original
equilibrium, type $\theta_1$ is a partisan for $A$, but in the new equilibrium she becomes
a partisan for $B$. Thus, partisanship is an equilibrium phenomenon and, therefore,
changes in the information structure of one type may affect behavior of another type.

4.4 Mitigation of equilibrium mistakes

Since mistakes often happen in equilibrium with retrospective voters, one concern is
that elections may perform poorly in cases where the performance of an alternative
is disastrous in some states of the world (say, in the case of a political party, because
of its ideology). To study this issue, we consider a family of environments, indexed
by $\beta \geq 0$, with homogeneous preferences $u(B, \omega)$ and

$$u^\beta(A, \omega) = u(A, \omega) - \beta \max \{ (h(c^{FB}) - h(\omega)), 0 \},$$

where $h(\cdot)$ is increasing and bounded and $c^{FB}$ solves $u(A, c^{FB}) = u(B, c^{FB})$ and
satisfies $c^{FB} \in (-1, 1)$. By construction, the first-best cutoff is $c^{FB}$ for every $\beta$. The
role of a higher $\beta$ is to make it more costly to elect $A$ in states of the world where $B$
is best. We conduct comparative statics with respect to $\beta$.

**Proposition 9.** Suppose that information is not aggregated for $\beta = 0$. Then either
$C^{eqm,\beta}(\rho) = C^{eqm,0}(\rho)$ for all $\beta \geq 0$ and $\rho \in (0, 1)$ (which happens if $c^0 > c^{FB}$) or

$$\lim_{\beta \to \infty} W^{\beta}(c^\beta) \geq E \left( u(B, W) \mid W < c^{FB} \right)$$

for any sequence $(c^\beta)_\beta$ where $c^\beta$ is the equilibrium cutoff under an optimal electoral
rule $\rho^\beta$ (which happens if $c^0 < c^{FB}$).

Proposition 9 provides a lower bound on limiting equilibrium welfare as the cost
of wrongly choosing alternative $A$ goes to infinity. Intuitively, we argued after Propo-
sition 2 that, when $c^0 < c^{FB}$, the mistake of wrongly choosing $A$ is necessary to
provide incentives for voters to also be willing to vote for $B$. As the exogenous cost
of wrongly choosing $A$ goes to infinity, then incentives can be provided by having the
probability of making the mistake go to zero. Moreover, the benefits from choosing
$B$ in the corresponding states of the world constitute a lower bound to the amount
of equilibrium damage that must be done to $A$. 21
4.5 Endogenous party polarization

We endogenize the primitives of the voting environment by focusing on the case where
A and B are political parties and allowing the parties to choose policies. To motivate
the analysis, suppose that the economy is either in a recessionary or overheated
state. Party A ideologically favors expansionary fiscal policy while party B favors
contractionary policy. Expansionary policy does best in a recession but hurts an
overheated economy, while contractionary policy does best in an overheated economy
but hurts during a recession. There is also a neutral, hands-off policy that neither
helps nor hurts the economy. Parties can choose to mitigate their ideological positions
by choosing any combination between their ideological position and the neutral policy.
Suppose that, in expectation, the neutral policy does better compared to the extreme
policies. Does electoral competition mitigate or exacerbate ideological positions?
Does it increase welfare?

The formal game is as follows. For simplicity, we assume that there is a unique type
and that neither the electorate nor the parties have private information. Each party
j ∈ {A, B} simultaneously chooses a policy pj ∈ [0, 1] that represents a weighted av-
erage between its extreme ideological policy, uj(ω), and a neutral policy that delivers
a constant payoff (normalized to zero). Given policies (pA, pB), the electorate’s payoff
from electing party j in state of the world ω is u(j, ω) = pjuj(ω), where we assume
that uA(·) is increasing and uB(·) is decreasing, so that assumption A2 is satisfied
provided that pA + pB > 0. In that case, the outcome of the election is given by the
voting equilibrium defined in Section 3. For the case where both parties choose the
neutral policy (pA = pB = 0), we naturally assume that, in each state of the world,
the parties have an equal chance of being elected. We model electoral competition
by solving for the Nash equilibrium in policies under the assumption that parties
maximize their probability of being elected. We add the following assumptions.

B1. Euj(W) < 0 for j = 1, 2; uA(1) > 0 and uB(−1) > 0.

B2. Let cA and cB be defined by E(uA(W) | W > cA) = 0 and E(uB(W) | W < cB) =
0.21 We assume that cA < cB.

Assumption B1 says that, in the case without electoral competition where either
one of these parties are in charge, then the neutral policy is preferred in expectation

21Existence and uniqueness of (cA, cB) follows from assumption B1.
by the electorate. However, there are some states of the world in which the extreme policies are preferred. Assumption B2 requires the parties’ extreme policies to be, in expectation, not too much worse than the neutral policy.

**Proposition 10.** Suppose that B1-B2 hold and let \( \rho \in (0,1) \) be any electoral rule. Then, in the unique Nash equilibrium, parties choose to exacerbate their ideological positions, \( p_A = p_B = 1 \). Moreover, electoral competition leads to strictly higher welfare compared to the case where either party is put in charge and chooses the neutral policy.

While the tendency of both parties to polarize and go against the preference of the (median) voter seems unintuitive, the result is a natural consequence of an environment where policy outcomes are uncertain and people vote retrospectively based on observed performance. Retrospective voting implies that party A will be elected in high states and party B in low states.\(^{22}\) Therefore, each party has an incentive to choose an extreme policy that works well in those extreme states in which it is usually elected into office, since those are the payoffs that voters use to evaluate their performance. In addition, since each policy tends to be appropriately matched with the right state of the world, the coexistence of extreme policies increases welfare.

## 5 Foundation for voting equilibrium

We provide a game-theoretic foundation for (retrospective) voting equilibrium. We begin by describing the voting game played by \( n \) players and the solution concept of non-strategic equilibrium, which captures the steady-state of a learning environment where players vote retrospectively. Players’ payoffs are perturbed to guarantee that each alternative is elected with positive probability in equilibrium. The main result is that the definition of voting equilibrium in Section 3 corresponds to the limit of non-strategic equilibria as, first, the number of players goes to infinity and, then, the payoff perturbations vanish.

\(^{22}\)As explained in Section 5, our setup with no private information captures the limiting case of a setup where people have private information that is very uninformative, so that the outcome of the election is still characterized by a cutoff.
5.1 Voting game

The rules of the game are as described in Section 3. The difference is that there are now a finite number of players indexed by $i = 1, \ldots, n$, and their utility from electing alternative $o \in \{A, B\}$ is given by

$$u_\theta(o, \omega) + 1 \{o = B\} v_i,$$

where $v_i \in \mathbb{R}$ is a privately-observed payoff perturbation drawn from a probability distribution $F_\theta$ and $\theta_i$ is the type of player $i$. By assumption A2, let $K < \infty$ denote a uniform bound on utility: $\sup_{o=\{A,B\}, \omega \in \Omega} |u_\theta(o, \omega)| < K$. In addition to A1-A4, we maintain the following assumptions, for all $\theta \in \Theta$:

**A5.** $F_\theta$ is absolutely continuous and satisfies $F_\theta(-2K) > 0$ and $F_\theta(2K) < 1$; its density $f_\theta$ satisfies $\inf_{x \in [-2K, 2K]} f_\theta(x) > 0$.

**A6.** $S_\theta$ has at least two elements and there exists $z > 0$ such that for $\omega' > \omega$, and $s' > s$:

$$\frac{q_\theta(s'|\omega')}{q_\theta(s'|\omega)} - \frac{q_\theta(s|\omega')}{q_\theta(s|\omega)} \geq z(\omega' - \omega).$$

Assumption A5 guarantees that each alternative is voted with positive probability. This property implies that the probability that players are pivotal (i.e., that their vote decides the election) becomes negligible as $n \to \infty$. Assumption A6 is a strengthening of MLRP that establishes a uniform bound on the rate at which the likelihood ratio changes.

Let $x = (x_1, \ldots, x_n)$, $x_i \in \{A, B\}$, denote the votes of every player. We integrate out the payoff perturbations and denote the resulting mixed strategy of player $i$ by $\alpha_i \in A_{\theta_i}$, where

$$A_{\theta_i} = \{\alpha_i : F_{\theta_i}(-2K) \leq \alpha_i(s_i) \leq F_{\theta_i}(2K) \forall s_i \in S_{\theta_i}\}$$

is the set of player $i$’s strategies, and $\alpha_i(s_i)$ is the probability that player $i$ votes

\footnote{A5 also yields a refinement, which is standard in the literature, that rules out equilibria where everyone votes for the same alternative because a unilateral deviation cannot change the outcome. Espinosa and Pouzo (2011) show that the perturbations are also important for providing a learning foundation for non-strategic equilibrium.}
for $A$ after observing signal $s_i$ (see Esponda and Pouzo (2011) for more details). Each strategy profile $\alpha = (\alpha_1, ..., \alpha_n) \in \times_{i=1}^n A_{\theta_i}$ induces a distribution over outcomes $P_\alpha^n \in \Delta (Z^n)$, where $Z^n \equiv X^n \times S^n \times \Omega$, $S^n \equiv \times_{i=1}^n S_{\theta_i}$, and

$$P(\alpha)(X', S', \Omega') = \sum_{x \in X'} \sum_{s \in S'} \int_{\Omega'} \left( \times_{i=1}^n \alpha_i(s_i) 1\{x_i = A\} (1 - \alpha_i(s_i)) 1\{x_i = B\} q_{\theta_i}(s_i \mid W) G(dW) \right),$$

for any $X' \subset X^n$, $S' \subset S^n$, $\Omega' \in \mathcal{B}$, where $\mathcal{B}$ is the Borel $\sigma$-algebra over $[-1, 1]$.

To gain intuition for our notion of equilibrium, suppose that player $i$ repeatedly faces a sequence of stage games where players use strategies $\alpha$ every period. Then, under the assumption that the payoff to alternative $A$ is observed only whenever $A$ is chosen, player $i$ will come to observe that, conditional on observing signal $s_i$, alternative $A$ yields in expectation $E_{P^n(\alpha)} (u_{\theta_i}(A, W) \mid o = A, s_i)$. A similar expression holds for alternative $B$.

A non-strategic (or naive) player who observes $v_i$ and $s_i$ believes that expected utility is maximized by voting for $A$ whenever $\Delta_i(P^n(\alpha), s_i) - v_i > 0$ and voting for $B$ otherwise, where

$$\Delta_i(P^n, s_i) \equiv E_{P^n(\alpha)} (u_{\theta_i}(A, W) \mid o = A, S_{\theta_i} = s_i) - E_{P^n(\alpha)} (u_{\theta_i}(B, W) \mid o = B, S_{\theta_i} = s_i)$$

is well-defined because of the payoff perturbations.

**Definition 9.** A strategy profile $\alpha \in \times_{i=1}^n A_{\theta_i}$ is a (non-strategic or naive) equilibrium of the voting game if for every player $i = 1, ..., n$ and for every $s_i \in S_i$,

$$\alpha_i(s_i) = F_i(\Delta_i(P^n(\alpha), s_i)).$$

In equilibrium, players best respond to beliefs that are endogenously determined by both their own strategy and those of other players and that are consistent with observed equilibrium outcomes. Non-strategic players, however, do not account for the correlation between others’ votes and the state of the world (conditional on their own private information). Esponda and Pouzo (2011) show existence of equilibrium and provide a learning foundation for this solution concept.\(^{25}\)

\(^{24}\)Whenever an expectation $E_P$ has a subscript $P$, this means that the probabilities are taken with respect to the distribution $P$.

\(^{25}\)The definition of equilibrium is a special case of Esponda’s (2008) behavioral equilibrium, which
5.2 Large number of players

We analyze games in which the number of players goes to infinity by studying sequences of voting games. We build such sequences by independently drawing infinite sequences of types \( \xi = (\theta_1, \theta_2, \ldots, \theta_n, \ldots) \in \Xi \) according to the probability distribution \( \phi \in \Delta(\Theta) \); we denote the distribution over \( \Xi \) by \( \Phi \). We interpret each sequence of types as describing an infinite number of \( n \)-player games by letting the first \( n \) elements of \( \xi \) represent the types of the \( n \) players.

Let \( \alpha \) denote a strategy mapping from sequences of types \( \Xi \) to sequences of strategy profiles—i.e., for all \( \xi \in \Xi \), let \( \alpha(\xi) = (\alpha^1(\xi), \ldots, \alpha^n(\xi), \ldots) \), where

\[
\alpha^n(\xi) = (\alpha^1_1(\xi), \ldots, \alpha^1_n(\xi), \ldots) \in \times_{i=1}^n A_{\theta_i},
\]

is the strategy profile that is played in the \( n \)-player game with types \( \theta_1, \ldots, \theta_n \). Let \( P^n(\alpha(\xi)) \) be the probability distribution over \( X^n \times S^n \times \Omega \) induced by the strategy profile \( \alpha^n(\xi) \) in the \( n \)-player game. We define three properties of strategy mappings.

**Definition 10.** A strategy mapping \( \alpha \) is an \( \varepsilon \)-equilibrium mapping if there exists \( n_{\varepsilon} \) such that for all \( n \geq n_{\varepsilon} \), \( i = 1, \ldots, n \), and \( s_i \in S_i \),

\[
|\alpha^n_i(\xi)(s_i) - F_i(\Delta_i(P^n(\alpha(\xi)), s_i))| \leq \varepsilon
\]

for all \( \xi \in \Xi \). For \( \Xi' \subset \Xi \), a strategy mapping \( \alpha \) is \( \Xi' \)-asymptotically interior if

\[
\liminf_{n \to \infty} P^n(\alpha(\xi))(o = A) > 0 \quad \text{and} \quad \limsup_{n \to \infty} P^n(\alpha(\xi))(o = A) < 1
\]

a.s. \(-\Xi'\). A strategy mapping \( \alpha \) is \( \Xi' \)-asymptotically \( c \)-cutoff if there exists \( c \in (-1, 1) \) such that

\[
\lim_{n \to \infty} P^n(\alpha(\xi))(o = A | \omega) = \begin{cases} 1 & \text{for } \omega > c \\ 0 & \text{for } \omega < c \end{cases} \quad \text{a.s. } -\Xi'.
\]

The first property in Definition 10 requires that, for large enough \( n \), players play strategies that constitute an \( \varepsilon \) equilibrium. Our notion of limit equilibrium will require combines the idea of a self-confirming equilibrium (Battigalli (1987), Fudenberg and Levine (1993), Dekel et al. (2004)) with an information-processing bias. See de Figueiredo Jr et al. (2006) for an application of self-confirming equilibrium to the American Revolution.
this property to hold for all $\varepsilon > 0$; while being slightly weaker than requiring strategies to constitute an equilibrium, this condition yields a full characterization of limit equilibrium.\textsuperscript{26} The second property requires that the probabilities of choosing $A$ and $B$ remain bounded away from zero as the number of players increases. We will provide a full characterization of equilibria that have such a property. The final property specifies that, as the number of players increases, the probability that $A$ is elected goes to 1 for states above a cutoff and to 0 for states below it. We will show that equilibrium can be characterized by this convenient property.

In addition to characterizing the equilibrium $c$-cutoff, our objective is to characterize the profile of equilibrium strategies. A complete characterization of equilibrium strategies is cumbersome due to the nature of the equilibrium object: As the number of players increases, the dimension of $\alpha^n$ also increases. We overcome this inconvenience by characterizing the limit, as the number of players increases, of the average strategy chosen by each type of player.\textsuperscript{27}

For a given strategy mapping $\alpha$ and a sequence of types $\xi \in \Xi$, let $\sigma^n(\xi; \alpha) = (\sigma^n_\theta(\xi; \alpha))_{\theta \in \Theta} \in \mathcal{A}^* \equiv \times_{\theta \in \Theta} \mathcal{A}_\theta$ denote the average strategy played by each type in the $n$-player game with types $(\theta_1, \ldots, \theta_n)$ and strategy profile $\alpha^n(\xi)$. Formally,

$$\sigma^n_\theta(\xi; \alpha)(s_i) = \frac{\sum_{i=1}^n 1\{\theta_i(\xi) = \theta\} \alpha^n_i(\xi)(s_i)}{\sum_{i=1}^n 1\{\theta_i(\xi) = \theta\}} \in \mathcal{A}_\theta \tag{6}$$

whenever $\sum_{i=1}^n 1\{\theta_i(\xi) = \theta\} > 0$, and arbitrary otherwise. We call any element $\sigma \in \mathcal{A}^*$ an average strategy profile and say that $\sigma$ is increasing if for each type $\theta \in \Theta$, $s'_\theta > s_\theta$ implies $\sigma_\theta(s'_\theta) > \sigma_\theta(s_\theta)$.

**Definition 11.** An average strategy profile $\sigma \in \mathcal{A}^*$ is a limit $\varepsilon$-equilibrium if there exists $\alpha$ and $\Xi'$ with $\Phi(\Xi') > 0$ such that: (1) $\alpha$ is an $\varepsilon$-equilibrium mapping; (2) $\alpha$ is $\Xi'$-asymptotically interior; and (3) $\lim_{n \to \infty} \|\sigma^n(\xi; \alpha) - \sigma\| = 0$ for all $\xi \in \Xi'$. If, in addition, $\alpha$ is $\Xi'$-asymptotically $c$-cutoff, then $\sigma$ is a $c$-cutoff limit $\varepsilon$-equilibrium.

An average strategy profile $\sigma \in \mathcal{A}^*$ is a limit equilibrium if it is a limit $\varepsilon$-equilibrium for all $\varepsilon > 0$.

\textsuperscript{26}Our result that a limit equilibrium is a fixed point of a particular correspondence remains true under the stronger requirement that strategies constitute an equilibrium. But the converse result, that any fixed point is also a limit equilibrium, relies on the notion of $\varepsilon$-equilibrium.

\textsuperscript{27}Unlike most of the related literature, we do not restrict players of the same type to following the same strategy.
The next result shows that limit equilibria have a cutoff characterization.

**Lemma 3.** There exists $\varepsilon > 0$ and $\gamma(\varepsilon)$ with $\lim_{\varepsilon \to 0} \gamma(\varepsilon) = 0$ such that for all $\varepsilon < \varepsilon$: if $\sigma$ is a limit $\varepsilon$-equilibrium, then (i) $\sigma$ is increasing and is a $c$-cutoff limit $\varepsilon$-equilibrium, where

$$\kappa(c; \sigma) = \rho,$$

(7)

and (ii) for all $\theta \in \Theta$ and $s_\theta \in S_\theta$,

$$|\sigma_\theta(s_\theta) - F_\theta(v_\theta(s_\theta; c))| \leq \gamma(\varepsilon).$$

**Proof.** See the Online Appendix.

The intuition of Lemma 3 is as follows. Since $\sigma^n(\xi; \alpha)$ converges to $\sigma$, then the probability that a randomly chosen player votes for A, conditional on $\omega$, converges to $\kappa(\omega; \sigma)$. By standard asymptotic arguments, the proportion of votes for A becomes concentrated around $\kappa(\omega; \sigma)$. So, for states where $\kappa(\omega; \sigma) > \rho$, the probability that the outcome is A converges to 1. Similarly, for states where $\kappa(\omega; \sigma) < \rho$, the probability that the outcome is A converges to 0. Finally, the key is to show that $\sigma$ is increasing, which then implies that there is at most one (measure zero) state such that $\kappa(\omega; \sigma) = \rho$, so that the outcome can be characterized by a cutoff.

The proof that $\sigma$ is increasing is standard for Nash equilibrium, where it relies on the fact that, by MLRP, higher signals convey more favorable information about A. In our context, higher signals also have a second, indirect effect, because, to the extent that players can be pivotal and affect the outcome of the election, their beliefs about alternatives also depend on their own strategy. In fact, this indirect effect may go in the opposite direction of the standard effect. However, we establish that the probability of being pivotal goes to zero as the number of players increases and, therefore, the indirect effect eventually vanishes and becomes dominated by the direct

\[\text{To see this claim, fix a player and a signal and suppose that she votes for A with probability close to 1. Then, most often, A is the outcome of the election whenever at least } k-1 \text{ or more of the other players have voted for A. Compare this case to the case where she votes for B with probability close to 1. Then, most often, A is the outcome of the election whenever at least } k \text{ or more of the other players have voted for A. If strategies are increasing, then, by MLRP, the first event conveys less favorable information about A. Therefore, a higher signal leads this player to vote more for A, which then makes her less favorable about A.}\]
exists a \( a.s.-\xi \) of Online Appendix to obtain and, thus, theorem and the fact that trivially to \( \sigma \lim \) obtain that, for all \( \theta \in \Theta \) and \( s \in S_\theta \).

Therefore, we can follow the proof leading to equation (18) in the Online Appendix to play \( \sigma \lim \theta \)–i.e., \( \alpha \) implies that, for all \( \epsilon > 0 \), \( |\sigma_\theta(s) - F_\theta(v_\theta(s;c))| \leq \gamma(\epsilon) \) for all \( \theta, s \). Since \( \gamma(\epsilon) \to 0 \) as \( \epsilon \to 0 \), then \( \sigma_\theta(s) = F_\theta(v_\theta(s;c)) \)

If: Consider the strategy mapping \( \alpha \) defined by letting players of type \( \theta \) always play \( \sigma_\theta \)-i.e., \( \alpha_i(\xi)(s_i) = \sigma_\theta(s_{\theta_i}) \) for all \( \xi \), all \( i \). First, note that \( \sigma^n = \sigma \) converges trivially to \( \sigma \), and \( \sigma \) is increasing because \( F_\theta \) and \( v(\cdot;c) \) are increasing (by A1-A5).

Therefore, we can follow the proof leading to equation (18) in the Online Appendix to obtain that \( \lim_{n \to \infty} P^n(\xi)(o = A|W) = 1\{W < c\} \) a.s.-\( \Xi \). The dominated convergence theorem and the fact that \( c \in (-1,1) \) implies that \( \lim_{n \to \infty} P^n(\xi)(o = A) \in (0,1) \), and, thus, \( \alpha \) is \( \Xi \)-asymptotically interior. Therefore, we can apply Lemma 3.1 in the Online Appendix to obtain \( \lim_{n \to \infty} \Delta_i(P^n(\alpha(\xi)), s_i) = v_{\theta_i}(s_i;c) \) a.s.-\( \Xi \). By continuity of \( F_{\theta_i} \) (A5), \( \lim_{n \to \infty} F_{\theta_i}(\Delta_i(P^n(\alpha(\xi)), s_i)) = F_{\theta_i}(v_{\theta_i}(s_i;c)) \). Therefore, for \( \epsilon > 0 \), there exists a \( n_\epsilon \) such that for \( n \geq n_\epsilon \), all \( i, s_i \)

\[
|\alpha_i^n(\xi)(s_i) - F_{\theta_i}(\Delta_i(P^n(\alpha(\xi)), s_i))| = |\sigma_{\theta_i}(s_i) - F_{\theta_i}(\Delta_i(P^n(\alpha(\xi)), s_i))| \\
= |F_{\theta_i}(v_{\theta_i}(s_i;c^*)) - F_{\theta_i}(\Delta_i(P^n(\alpha(\xi)), s_i))| < \epsilon
\]
a.s.-\( \Xi \).

\[\text{The result that the probability of being pivotal vanishes would fail if the variance were zero. For example, suppose that} \ n \ \text{is even, voting is by majority rule, and half of the players vote for} \ A \ \text{and half vote for} \ B. \ \text{Then, each player is pivotal with probability} \ 1, \ \text{for all} \ n.\]

29
5.3 Vanishing perturbations

While the perturbations may have a natural interpretation in some contexts, we now consider sequences of equilibria where the perturbations vanish. We index games by a parameter \( \eta \) that affects the distribution \( F_\eta \) from which perturbations are drawn.

**Definition 12.** A family of perturbations \( \{F_\eta\}_{\eta \in \mathbb{N}} \), where \( F_\eta = \{F_\theta^\eta\}_{\theta \in \Theta} \), is vanishing if for all \( \theta \in \Theta \) and \( \eta \) assumption A5 is satisfied and

\[
\lim_{\eta \to 0} F_\theta^\eta(v) = \begin{cases} 
0 & \text{if } v < 0 \\
1 & \text{if } v > 0
\end{cases}
\]

The next two results provide a foundation for (retrospective) voting equilibrium.

**Theorem 2.** Suppose that there exists a vanishing family of perturbations \( \{F_\eta\}_{\eta} \) and a sequence \( (\sigma_\eta, c_\eta)_{\eta} \) such that \( \lim_{\eta \to 0} (\sigma_\eta, c_\eta) = (\sigma, c) \) and where \( \sigma_\eta \) is a limit equilibrium and \( c_\eta \) its corresponding cutoff for all \( \eta \). Then \( (\sigma, c) \) is a voting equilibrium.

**Proof.** Theorem 1 implies that \( \sigma_\theta(s) = \lim_\eta \sigma_\theta^\eta(s) = \lim_\eta F_\theta^\eta(v_\theta(s;c)) \). Since \( F_\eta \) is vanishing, then \( \sigma_\theta(s) = 1 \) if \( v_\theta(s;c) > 0 \) and \( \sigma_\theta(s) = 0 \) if \( v_\theta(s;c) < 0 \). Therefore, \( \sigma \) is optimal given \( c \). Next, fix any \( \omega' < c \). Since \( c_\eta \to c \), there exists \( \bar{\eta} \) such that, for all \( \eta < \bar{\eta} \), \( \omega' < c_\eta \), and, by Theorem 1, \( \kappa(\omega';\sigma_\eta) \leq \rho \). Since \( \sigma_\eta \to \sigma \), continuity of \( \kappa(\omega';\cdot) \) implies that \( \kappa(\omega';\sigma) \leq \rho \). Similarly, \( \kappa(\omega'';\sigma) \geq \rho \) for all \( \omega'' > c \). Therefore, \( c \) is an election cutoff given \( \sigma \).

**Theorem 3.** Suppose that \((\sigma, c)\) is a voting equilibrium with \( c \in (-1, 1) \). Then there exists a vanishing family of perturbations \( \{F_\eta\}_{\eta} \) and a sequence \( (\sigma_\eta, c_\eta)_{\eta} \) such that \( \lim_{\eta \to 0} (\sigma_\eta, c_\eta) = (\sigma, c) \) and where \( \sigma_\eta \) is a limit equilibrium and \( c_\eta \) its corresponding cutoff for all \( \eta \).

**Proof.** See the Appendix.

We conclude by making two observations. First, situations where one alternative is never chosen (i.e., \( c = -1 \) or \( c = 1 \)) are easily justified: if an alternative is never
chosen, then beliefs about its performance can be arbitrary. Our solution concept in Section 3 considers $c = -1$ or $c = 1$ to be an equilibrium only if it can be approached by a sequence of limit equilibria where both alternatives are chosen and, therefore, by a sequence of non-arbitrary beliefs.\footnote{Our solution concept implicitly assumes that people have the most optimistic beliefs for an alternative that is not chosen. For example, if $c = -1$ then $B$ is never chosen. However, players believe $u(B, -1)$, as if $B$ were indeed chosen (only) under state $\omega = -1$. This refinement can be justified formally provided that people use their private information when voting (so that $\kappa(\cdot; \sigma)$ is increasing and the probability that $B$ is chosen, while vanishing, is infinitely more likely for $\omega = -1$ than for higher states). However, the refinement is arbitrary when $\sigma$ is flat.}

Second, our game-theoretic foundation uses assumption A6, which is stronger than assumption A3 in Section 3. In particular, A3 allows for the case where people have no private information. We can provide a foundation for such a case by considering a sequence of voting games indexed by $r \in \mathbb{N}$, where $z^r > 0$ denotes the constant defined in assumption A6, and where $\lim_{r \to \infty} z^r = 0$. Therefore, we view the case of no information as the limiting case of an information structure that satisfies A6 but where informativeness vanishes.

6 Conclusion

We have provided a framework that formalizes the idea that voters vote retrospectively in large elections. The framework can be interpreted as the counterpart of the standard competitive equilibrium environment in a market economy but where actions are mediated by an electoral rule rather than by prices. In particular, the framework is easily applicable to obtain several insights regarding the extent of information aggregation, optimal electoral rules, the value of information, endogenous preferences or party affiliation, and party polarization under retrospective voting.

The model can be generalized in two directions. First, more structure can be added to players’ optimization problem in order to represent a priori information about the primitives that players can use to make inferences about counterfactuals.\footnote{In the extreme case where players can perfectly deduce counterfactuals, the model collapses to the standard notion of sincere voting.} Second, it is possible to allow for more sophisticated retrospective rules that condition also on past vote shares (see Esponda and Pouzo (2011)).

The framework requires non-trivial extensions to answer some other questions of interest. For example, given a theory of why people vote, we could extend the model...
to allow for abstention and study turnout. In addition, our simple characterization of voting outcomes in large elections draws heavily on the monotonicity assumptions on preferences and information. In settings where these assumptions fail, the model would need to be extended and the analysis would be more complicated.32 Finally, we have abstracted away from other interesting dynamics, such as the fact that the performance of the currently-elected alternative may be influenced by the performance of the previously-elected alternative. We believe that the current model can be a good starting point for these and other future extensions.

7 Appendix

Proof of Lemma 1. $\bar{\kappa}(\cdot)$ is nondecreasing: For any $\omega' > \omega$,

$$\sum_\theta \phi(\theta)q_\theta (c_\theta(S_\theta) < \omega' \mid \omega') \geq \sum_\theta \phi(\theta)q_\theta (c_\theta(S_\theta) < \omega \mid \omega') \geq \sum_\theta \phi(\theta)q_\theta (c_\theta(S_\theta) < \omega \mid \omega),$$

where the last inequality follows because, since $c_\theta(\cdot)$ is nondecreasing, the events $\{c_\theta(S_\theta) < \omega\}$ and $\{S_\theta \leq s_\theta(\omega)\}$ are equivalent for some threshold $s_\theta(\omega)$, and thus MLRP implies that $q_\theta(\cdot \mid \omega')$ first-order stochastically dominates $q_\theta(\cdot \mid \omega)$ (Milgrom, 1981a).

$\bar{\kappa}(\cdot)$ is left-continuous: Since there are a finite number of personal cutoffs (defined by equation (2)), then for each $c \in (-1, 1)$ there exists $\omega' < c$ such that all personal cutoffs are outside the interval $I_c = [\omega', c)$. Then, for all $\theta$ and all $\hat{\omega}$, $q_\theta (c_\theta(S_\theta) < \omega \mid \hat{\omega}) = q_\theta (c_\theta(S_\theta) < c \mid \hat{\omega})$ for all $\omega \in I_c$. In addition, $q_\theta(s \mid \cdot)$ is continuous by A4(iii). Therefore, $\lim_{\omega \downarrow c} q_\theta (c_\theta(S_\theta) < c \mid \omega) = q_\theta (c_\theta(S_\theta) < c \mid c)$ for all $\theta$.

Finally: If $\omega \leq c$, then $\{c_\theta(S_\theta) < \omega\} = \emptyset$, so that $\bar{\kappa}(\omega) = 0$. Similarly, if $\omega > \bar{c}$, then $\{c_\theta(S_\theta) < \omega\} = S_\theta$, so that $\bar{\kappa}(\omega) = 1$. Last, we establish that $\bar{\kappa}(\omega) \in (0, 1)$ for any $\omega \in (c, \bar{c}]$. First, we show that $\bar{\kappa}(\omega) < 1$ for any $\omega \in [c, \bar{c}]$. Suppose not, so that, because $\bar{\kappa}(\cdot)$ is nondecreasing, $\bar{\kappa}(\bar{c}) = 1$. Then $\sum_{\theta \in \Theta} \phi(\theta)q_\theta (c_\theta(S_\theta) = \bar{c} \mid \bar{c}) = 0$.

32Bhattacharya (2008) presents examples where monotonicity fails and, using the Nash equilibrium solution concept, shows that information aggregation may fail in those cases. In our model, the lack of monotonicity will change the conditioning events from an interval such as $W \geq c$ to a subset of $\Omega$, but complications will arise if the resulting $\kappa$ function is flat when intersecting the electoral rule.
but this contradicts A4(ii). We now show that \( \pi(\omega) > 0 \) for any \( \omega \in (c, \bar{c}] \). Since, \( \pi \) is nondecreasing, it suffices to show that \( \lim_{\omega \downarrow c} \pi(\omega) > 0 \). Suppose not; then \( \lim_{\omega \downarrow c} \sum_{\theta \in \Theta} \phi(\theta) q_\theta(c_\theta(S_\theta)) = c | \omega = 0 \). By continuity of \( q_\theta(s | .) \) (A4(iii)), the previous equation contradicts A4(ii). \( \square \)

**Proof of Proposition 1.** The proof relies on the following claim.

**Claim 1.1** Suppose that \( \sigma \) is optimal given election cutoff \( c \). Then

\[
\kappa(\omega; \sigma) = \sum_{\theta \in \Theta} \phi(\theta) \left( q_\theta(c_\theta(S_\theta) < c | \omega) + \sum_{s: c_\theta(s) = c} q(s | \omega) \sigma_\theta(s) \right). \tag{8}
\]

In addition, \( \bar{\kappa}(\omega) \geq \kappa(\omega; \sigma) \) for \( \omega > c \) and \( \bar{\kappa}(\omega) \leq \kappa(\omega; \sigma) \) for \( \omega < c \).

**Proof.** Since \( \sigma \) is optimal given \( c \), then

\[
\sigma_\theta(s) = \begin{cases} 
0 & \text{if } c_\theta(s) > c \\
1 & \text{if } c_\theta(s) < c 
\end{cases} \tag{9}
\]

and equation (8) follows. In addition, for all \( \omega > c \),

\[
\kappa(\omega; \sigma) \leq \sum_{\theta \in \Theta} \phi(\theta) q_\theta(c_\theta(S) \leq c | \omega) \leq \sum_{\theta \in \Theta} \phi(\theta) q_\theta(c_\theta(S) < c | \omega) = \bar{\kappa}(\omega). 
\]

Similarly, for all \( \omega < c \), \( \kappa(\omega; \sigma) \geq \bar{\kappa}(\omega) \).

We now prove both statements in Proposition 1.

*If \( (c, \sigma) \) is an equilibrium, then \( c \in C^{eqm} \):* For all \( \omega > c \),

\[
\bar{\kappa}(\omega) \geq \kappa(\omega; \sigma) \geq \rho,
\]

where the first inequality follows from the fact that \( \sigma \) is optimal given \( c \) and Claim 1.1, and the second inequality follows because \( c \) is a cutoff given \( \sigma \). These inequalities imply that \( c \geq \inf \{ \omega : \bar{\kappa}(\omega) \geq \rho \} \). Similarly, for all \( \omega < c \),

\[
\bar{\kappa}(\omega) \leq \kappa(\omega; \sigma) \leq \rho,
\]
which implies that \( c \leq \sup \{ \omega : \bar{k}(\omega) \leq \rho \} \).

If \( c \in C^{eqm} \), then there exists \( \sigma \) such that \((\sigma, c)\) is an equilibrium: Let \( \sigma \) satisfy (9). It remains to specify \( \sigma_\theta(s) \) for \((\theta, s)\) such that \( c_\theta(s) = c \). First, suppose that \( c \notin \{-1, 1\} \). Then \((\theta, s)\) such that \( c_\theta(s) = c \) is indifferent between \( A \) and \( B \), and, therefore, \( \sigma_\theta(s) = \alpha \) is optimal for any \( \alpha \in [0, 1] \). Let \( \sigma_\alpha \) denote the strategy profile constructed above. We now pick \( \alpha \) such that \( c \) is an election cutoff given \( \sigma_\alpha \). Let \( \hat{k}(\alpha) \equiv \kappa(c; \sigma_\alpha) \). By Claim 1.1,

\[
\hat{k}(\alpha) = \sum_{\theta \in \Theta} \phi(\theta) [q_\theta(c_\theta(S_\theta) < c \mid c) + q_\theta(c_\theta(S_\theta) = c \mid c) \alpha],
\]

which is continuous in \( \alpha \). First, we establish that \( \hat{k}(0) \leq \rho \). Suppose not, so that \( \hat{k}(0) = \bar{k}(c) > \rho \). Since \( \bar{k} \) is left-continuous (Lemma 1), then there exists \( \omega' < c \) such that \( \bar{k}(\omega') > \rho \). But then \( c > \sup \{ \omega : \bar{k}(\omega) \leq \rho \} \), which contradicts \( c \in C^{eqm} \).

Second, we establish that \( \hat{k}(1) \geq \rho \). Suppose not, so that \( \hat{k}(1) = \lim_{\omega \downarrow 1} \bar{k}(\omega) < \rho \). Then, there exists \( \omega'' > c \) such that \( \bar{k}(\omega'') < \rho \). But then \( c < \inf \{ \omega : \bar{k}(\omega) \geq \rho \} \), which contradicts \( c \in C^{eqm} \). Since \( \hat{k}(0) \leq \rho \) and \( \hat{k}(1) \geq \rho \), by continuity of \( \hat{k} \) there exists \( \alpha^* \) such that \( \hat{k}(\alpha^*) = \kappa(c; \sigma_{\alpha^*}) = \rho \). Since \( \kappa(\cdot; \sigma_{\alpha^*}) \) is nondecreasing (because \( \sigma_{\alpha^*} \) is nondecreasing), then \( c \) is an election cutoff given \( \sigma_{\alpha^*} \). Hence, \((\sigma_{\alpha^*}, c)\) is a voting equilibrium.

Finally, consider the case \( c = -1 \) (the case \( c = 1 \) is similar and, therefore, omitted). By setting \( \alpha^* = 1 \), it follows that \( \sigma_{\alpha^*} \) is optimal given \( c \). In addition, it follows from above that \( \hat{k}(1) = \kappa(c; \sigma_{\alpha^*}) \geq \rho \). Since \( \kappa(\cdot; \sigma_{\alpha^*}) \) is nondecreasing, it follows that \( \kappa(\omega; \sigma_{\alpha^*}) \geq \rho \) for all \( \omega \), implying that \( c = -1 \) is a cutoff given \( \sigma_{\alpha^*} \).

Existence of equilibrium: By Lemma 1, \( \{ \omega \in \Omega : \bar{k}(\omega) \geq \rho \} \) and \( \{ \omega \in \Omega : \bar{k}(\omega) \leq \rho \} \) are nonempty. Since \( \Omega \subset \mathbb{R} \) is bounded, then the supremum and infimum exist, implying that \( C^{eqm} \neq \emptyset \). \( \square \)

Proof of Theorem 3. For this proof, define

\[
\bar{k}^\eta(\omega) \equiv \sum_{\theta \in \Theta} \phi(\theta) \sum_{s \in S_\theta} q_\theta(s \mid \omega) F_\theta^\eta (v_\theta(s; \omega)).
\]

Let \((\sigma, c)\) be a voting equilibrium with \( c \in (-1, 1) \). Because \( c \in (-1, 1) \) is an election cutoff given \( \sigma \) and \( \kappa(\cdot; \sigma) \) is continuous, then \( \kappa(c; \sigma) = \rho \). We split the proof into two cases: Either it is the case that all players vote for the same alternative
(which may be different for each player) irrespective of their private information—so that $\kappa(\cdot; \sigma)$ is a constant function—or not—so that $\kappa(\cdot; \sigma)$ is increasing.

Case 1 ($\kappa(\cdot; \sigma)$ is increasing): Rewrite $\overline{\kappa}^n$ as

$$
\hat{\kappa}^n(\omega) = \sum_{\theta \in \Theta} \phi(\theta) \left\{ \sum_{s : c_\theta(s) < c} q_\theta(s \mid \omega) F^\eta_\theta (v_\theta(s; \omega)) + \sum_{s : c_\theta(s) = c} q_\theta(s \mid \omega) F^\eta_\theta (v_\theta(s; \omega)) \right\} + \sum_{s : c_\theta(s) > c} q_\theta(s \mid \omega) F^\eta_\theta (v_\theta(s; \omega)) = T^1_\eta(\omega) + T^2_\eta(\omega) + T^3_\eta(\omega).
$$

Since $v_\theta(s; \cdot)$ is increasing and $c \in (-1, 1)$, then: for all $(\theta, s)$ such that $c_\theta(s) \geq c$, $v_\theta(s; \omega) < 0$ for all $\omega < c$ and, for all $(\theta, s)$ such that $c_\theta(s) \leq c$, $v_\theta(s; \omega) > 0$ for all $\omega > c$. Therefore, since $\{F^\eta\}_\eta$ is vanishing and, for all $\omega < c$ and $\lim_{\eta \to 0} T^1_\eta(\omega) + T^2_\eta(\omega) = \sum_{\theta \in \Theta} \phi(\theta) q_\theta(c_\theta(S_\theta) \leq c \mid \omega) \geq \kappa(\omega; \sigma)$ for all $\omega > c$. In addition, $T^1_\eta(\omega) \leq \kappa(\omega; \sigma)$ and $T^2_\eta(\omega) \geq 0$ for all $\omega$. Therefore, $\lim_{\eta \to 0} \overline{\kappa}^n(\omega) \leq \kappa(\omega; \sigma) < \kappa(c; \sigma) = \rho$ for all $\omega < c$ and $\lim_{\eta \to 0} \overline{\kappa}^n(\omega) \geq \kappa(\omega; \sigma) > \kappa(c; \sigma) = \rho$ for all $\omega > c$. Consequently, by continuity of $\kappa^\eta(\cdot)$, there exists $(c^\eta)_\eta$ such that $c^\eta \to c \in (-1, 1)$ and $\overline{\kappa}^n(c^\eta) = \rho$ for all sufficiently small $\eta$. By letting $\sigma^\eta_\theta(s) = F^\eta(v_\theta(s; c^\eta))$ for all $\theta, s$, it follows that $\kappa(c^\eta; \sigma^\eta) = \overline{\kappa}^n(c^\eta) = \rho$ for all sufficiently small $\eta$ and, by Theorem 1, that $\sigma^\eta$ is a limit equilibrium and $c^\eta$ its corresponding cutoff for all sufficiently small $\eta$. Finally, it remains to establish that $\sigma^\eta \to \sigma$. Consider a type and signal such that $c_\theta(s) < c$, so that $v_\theta(s; c) > 0$. By continuity of $v_\theta(s; \cdot)$ and the fact that $c^\eta \to c$, it follows that $v_\theta(s; c^\eta) > 0$ for all sufficiently small $\eta$ and, therefore, by feasibility of $\{F^\eta\}_\eta$, that $\lim_{\eta \to 0} \sigma^\eta_\theta(s) = 1 = \sigma_\theta(s)$, where the last equality follows since $\sigma$ is optimal given $c$—see equation (9). A similar argument establishes that $\lim_{\eta \to 0} \sigma^\eta_\theta(s) = 0 = \sigma_\theta(s)$ for types and signals such that $c_\theta(s) > c$. Therefore, if $\{s : c_\theta(s) = c\} = \emptyset$ for all $\theta$, we have shown that, for any family of vanishing perturbations, there exists a sequence of limit equilibria that converge to a voting equilibria. In the case where $\{s : c_\theta(s) = c\} \neq \emptyset$ for some $\theta$, we construct a specific family of perturbations $\{\hat{F}^\eta\}_\eta$ with the property that $\lim_{\eta \to 0} \hat{F}^\eta_\theta(v_\theta(s; c^\eta)) = \sigma_\theta(s)$ for all $(\theta, s)$ such that $c_\theta(s) = c$; the details that show existence of such a family are tedious but straightforward and are available from the authors upon request.

Case 2 ($\kappa(\omega; \sigma) = \rho$ for all $\omega$): Without loss of generality, suppose that $S_\theta \subset (0, \infty)$ for all $\theta$. Let $T_B = \{ (\theta, s) : v_\theta(s; c) < 0 \text{ or } (v_\theta(s; c) = 0 \& \sigma_\theta(s) = 0) \}$, $T_A =$
\{(\theta, s) : v_{\theta}(s; c) > 0 \text{ or } (v_{\theta}(s; c) = 0 \& \sigma_{\theta}(s) = 1)\}, \text{ and } T_0 = \{(\theta, s) : v_{\theta}(s; c) = 0 \& \sigma_{\theta}(s) \in (0, 1)\}. \text{ Note that, since } (\sigma, c) \text{ is a voting equilibrium, then } \sigma_{\theta}(s) = 0 \text{ if } (\theta, s) \in T_B \text{ and } \sigma_{\theta}(s) = 1 \text{ if } (\theta, s) \in T_A. \text{ Define } X_{B} \equiv \sum_{(\theta, s) \in T_B} \phi(\theta)q(s \mid c)s \geq 0, X_{A} \equiv \sum_{(\theta, s) \in T_A} \phi(\theta)q(s \mid c)s \geq 0, \text{ and } X_0 \equiv \sum_{(\theta, s) \in T_0} \phi(\theta)q(s \mid c)s \geq 0. \text{ The proof constructs a specific family of perturbations. For all } \eta \text{ and all } \theta \in \Theta \text{ and } s \in S_\theta \text{ let}
\[
F^n_{\theta}(v_{\theta}(s; c)) = \begin{cases} 
\zeta_B s\eta & \text{if } v_{\theta}(s; c) < 0 \text{ or } (v_{\theta}(s; c) = 0 \& \sigma_{\theta}(s) = 0) \\
\sigma_{\theta}(s) + \zeta_0 \eta & \text{if } \{v_{\theta}(s; c) = 0 \& \sigma_{\theta}(s) \in (0, 1)\} \\
1 - \frac{\zeta_A}{s} \eta & \text{if } v_{\theta}(s; c) > 0 \text{ or } (v_{\theta}(s; c) = 0 \& \sigma_{\theta}(s) = 1) 
\end{cases}
\]

By construction, for all } \zeta_j \in (0, \infty), j = A, B \text{ and } \zeta_0 \in [0, \infty), \text{ and for all } \eta \text{ sufficiently low, there exists a vanishing family } \{F^n\}_{\eta} \text{ that satisfies the above restrictions; note that, by MLRP, for each } \theta \text{ there is at most one signal that satisfies } v_{\theta}(s; c) = 0. \text{ Then, since } c \in (-1, 1),
\[
\bar{\kappa}^n(c) - \rho = \bar{\kappa}^n(c) - \kappa(c; \sigma) = \sum_{(\theta, s)} \phi(\theta)q(s \mid c) \left(F^n_{\theta}(v_{\theta}(s; c)) - \sigma_{\theta}(s)\right) \\
= \eta \left(-\zeta_A X_A + \zeta_B X_B + \zeta_0 X_0\right).
\]

It is straightforward to check that we can always pick } \zeta_A, \zeta_B, \zeta_0 \text{ such that } -\zeta_A X_A + \zeta_B X_B + \zeta_0 X_0 = 0 \text{ and, therefore, } \bar{\kappa}^n(c) = \rho \text{ for all } \eta \text{ sufficiently small. As in Case 1, by letting } \sigma^n_{\theta}(s) = F^n(v_{\theta}(s; c^n)) \text{ for all } (\theta, s), \text{ it follows that } \sigma^n \text{ is a limit equilibrium and } c \text{ its corresponding cutoff for all sufficiently small } \eta. \text{ The proof is completed by noting that, by construction, } \lim_{\eta \to 0} \sigma^n = \sigma. \square

References


_ and E.I. Vespa, “Hypothetical Thinking and Information Extraction: Pivotal Voting in the Laboratory,” 2011._


Online Appendix

This appendix contains additional proofs for “Conditional Retrospective Voting in Large Elections,” by Ignacio Esponda and Demian Pouzo.

Proofs in Section 4

Proof of Proposition 3. If \( \overline{c} < c^{FB} \), since \( \mathcal{W} \) is single-peaked, \( \rho \) is optimal if and only if \( c \in C^{eqm}(\rho) \). From proposition 1, it is easy to see that this is the case if and only if, \( \rho \geq 1 - \sum_{\theta, s_{c_\theta}(s) = \overline{c}} \phi(\theta)q_\theta(s | \overline{c}) \). The case \( c > c^{FB} \) is similar.

If information is aggregated, then there exists an electoral rule \( \rho \) such that \( c^{FB} \in C^{eqm}(\rho) \), and by proposition 1, this is equivalent to \( \rho \in [\pi(c^{FB}), \lim_{\omega \downarrow c^{FB}} \pi(\omega)] \). \( \square \)

Proof of Proposition 4. First, we show that for any environment that satisfies Condorcet informativeness, there exists a rule \( \rho \) such that \( \max_{c \in C^{eqm}(\rho)} \mathcal{W}(c) < \min_{c' \in C^{eqm}(1/2)} \mathcal{W}(c') \). Without loss of generality, suppose that MLRP holds strictly (if A3 holds with equality for two signals \( s'_\theta \) and \( s''_\theta \), then relabel one of these signals as the other). Since \( c^{FB} \in (-1, 1) \), Condorcet informativeness and MLRP (A3) imply that for every type \( \theta \), \( S_\theta \) has at least two elements (so that \( s^H_\theta > s^L_\theta \)) and there is an optimal decision policy where \( \sigma(\theta) = 1 \) (by Condorcet informativeness, \( \sigma(\theta) = 1 \) is optimal), and, therefore,

\[
q_\theta(s^H_\theta | \omega) < 1/2 \text{ for all } \omega < c^{FB} \tag{10}
\]

A similar argument establishes that

\[
1 - q_\theta(s^L_\theta | \omega) > 1/2 \text{ for all } \omega > c^{FB}. \tag{11}
\]

Denote by \( \Theta \) and \( \overline{\Theta} \) the set of types that have personal cutoffs \( c \) and \( \overline{c} \), respectively (generically, these sets will be singletons). For notational simplicity, let \( q_\Theta(s^j | \cdot) \equiv \sum_{\theta \in \Theta^j} \phi(\theta)q_\theta(s^j | \cdot) \), for all \( j = L, H \) and \( \Theta' \subset \Theta \). Given MLRP, equation (10) implies that \( \lim_{\omega \downarrow c^{FB}} \overline{\mathcal{K}}(\omega) = q_{\overline{\Theta}}(s^H | c) < 1/2 \) if \( c < c^{FB} \) and (11) implies that \( \overline{\mathcal{K}}(\overline{c}) = 1 - q_{\overline{\Theta}}(s^L | \overline{c}) > 1/2 \) if \( \overline{c} > c^{FB} \). Therefore, by Lemma 1, Proposition 1, and Corollary 1, for any equilibrium cutoff under majority rule, \( c^M \), it follows that \( c^M > c \) if \( c < c^{FB} \) and \( c^M < c \) if \( c > c^{FB} \). Moreover, for any \( \rho_L < \lim_{\omega \downarrow c^{FB}} \overline{\mathcal{K}}(\omega) \)
and $\rho_H > \bar{\kappa}(\bar{c})$, then $\max_{c \in C^{eqm}(\rho_L)} W(c) = W(\xi)$ and $\max_{c \in C^{eqm}(\rho_H)} W(c) = W(\bar{c})$.

We now divide the proof into three cases. (1) $c^{FB} \leq \xi < \bar{c}$: Then $c^M < \bar{c}$ implies that $W(\bar{c}) < W(c^{M})$; (2) $\xi < \bar{c} \leq c^{FB}$: Then $c^M > \xi$ implies $W(\xi) < W(c^{M})$; (3) $\xi < c^{FB} < \bar{c}$: Then $c^M \in (\xi, \bar{c})$ implies that $W(\bar{c}) < W(c^{M})$ if $c^M > c^{FB}$ and $W(\xi) < W(c^{M})$ if $c^M \leq c^{FB}$.

Finally, we establish that majority rule is the unique rule satisfying the above property: i.e., for any $\rho \neq 1/2$, there exists a voting environment that satisfies Condorcet informativeness such that $\max_{c \in C^{eqm}(\rho)} W(c) \leq \min_{c \in \bar{\Theta}} W(c).$ Fix $\rho < 1/2$ (the proof for $\rho > 1/2$ is similar and, therefore, omitted). Consider the following environment with $\Theta = \{\theta\}$, $u(A, \omega) = \omega$, $u(B, \omega) = 0$, $W \sim U[-1, 1]$, $S_\theta = \{s^L_\theta, s^H_\theta\}$ and $q_\theta(s^H_\theta | \omega) = \frac{1}{2} + \theta \omega$, where $\theta = \frac{1}{2} - \rho$. Note that A1-A4 and Condorcet informativeness are satisfied and that $c^{FB} = 0$. In addition, the facts that $E(u(A, W) | W > c, S_\theta = s^L_\theta) > 0$ for all $c \geq 0 = c^{FB}$ and $E(u(A, W) | S_\theta = s^H_\theta) > 0$ implies that $-1 = c_\theta(s^H_\theta) < c_\theta(s^L_\theta) < c^{FB}$. Since $q_\theta(s^H_\theta | -1) = \frac{1}{2} - \theta = \rho$, then Proposition 1 implies that $-1$ is the unique equilibrium cutoff under $\rho$, therefore yielding the lowest possible equilibrium payoff. □

**Proof of Proposition 5.** By assumption (iv), there is a unique type $\bar{\theta}$ such that $c_\theta(s^L_\theta) = \bar{c}$. By assumption (ii) and the fact that, for all $\theta$, $E(u(A, W) | W > c, S_\theta = s^L_\theta) > u(B)$ for all $c \geq c^{FB}$, it follows that $-1 < \bar{c} < c^{FB}$. Proposition 2 implies that information cannot be aggregated. Then Proposition 3 implies that $\rho$ is optimal if and only if $\rho \geq \rho^* \equiv 1 - \phi(\bar{\theta})q_\theta(c(S_\theta) = \bar{c} | \bar{c}) > 1/2$, where the last inequality follows from assumption (iii). □

**Proof of Lemma 2.** For any $c \in \Omega$ and $s \in S$,

$$v(s; c; \mathcal{I}) = E_T(u(A, W) | W \geq c, S = s) - E_T(u(B, W) | W \leq c, S = s)$$

$$= E^T_A(u(A, W) | S = s) - E^T_B(u(B, W) | S = s),$$

where expectations are taken with respect to information structure $\mathcal{I}$ and, in addition, the expectation $E_A$ is taken with respect to the primitives $\Omega_A = [c, 1]$ and $g_A(W) = g(W)/(1 - G(c))$ and $E_B$ is taken with respect to $\Omega_B = [-1, c]$ and $g_B(W) = g(W)/G(c)$. A similar expression holds for $\mathcal{I}'$. By monotone utility (A2)
and MLRP (A3),
\[ v(s^H; c; I) \geq v(s; c; I) \text{ and } v(s^H; c; I') \geq v(s'; c; I') \]
for all \( s \in S \) and \( s' \in S' \); where \( s^H \) and \( s'^H \) are the highest signals in \( S \) and \( S' \) respectively (see Milgrom 1981 for a proof). Since \( I' \) is more informative than \( I \), then
\[ v(s^H; c; I) = \sum_{s'' \in S'} v(s' ; c; I') \cdot \frac{m_{s's''} \Pr(s')}{\sum_{s'' \in S'} m_{s's''} \Pr(s)} \leq v(s'^H; c; I'). \quad (12) \]
Then, by equation (2),
\[ \zeta(I') = c(s'^H; I') \leq c(s^H; I) = \zeta(I). \]
Finally, if \(|S'| \geq 2\), then A2-A3 imply that \( v(s^H; c; I') > v(s'; c; I') \) for all \( s' \neq s^H \). Hence, if \( m_{s's} > 0 \) for all \( s, s' \), then (12) holds with strict inequality. Therefore, \( c(s^H; I') < c(s^H; I) \) provided, of course that \( c(s^H; I) \neq -1 \). A similar argument establishes that \( \zeta(I') \geq \zeta(I) \), with strict inequality under the given additional conditions. □

**Proof of Proposition 6.** By Proposition 1, any \( \rho \in \left(0, \min \{q(s^H | c(s^H, I)), q'(s^H | c(s'^H, I'))\}\right) \) yields unique equilibrium cutoffs \( c(s^H, I) > c(s'^H, I') \) under \( I \) and \( I' \), respectively, where the inequality follows because \( I' \) is strictly more informative than \( I \) and \( c(s, I) \in (-1, 1) \) for all \( s \in S \) (Lemma 2).

Similarly, any \( \rho' \in \left(1 - \min \{q(s^H | c(s^H, I)), q'(s^H | c(s'^H, I'))\}, 1\right) \) yields unique equilibrium cutoffs \( c(s^L, I) < c(s'^L, I') \) under \( I \) and \( I' \), respectively. Suppose that \( c(s^H, I) \leq c^{FB} \). Then the fact that \( \mathcal{W}(\cdot) \) is increasing for \( c < c^{FB} \) implies that \( \mathcal{W}(c(s^H, I')) < \mathcal{W}(c(s^H, I)) \), and, therefore, that equilibrium welfare is strictly higher under \( I \) for any electoral rule \( \rho \) in the first interval. Similarly, in the case \( c(s^L, I) \geq c(s^H, I) > c^{FB} \), the fact that \( \mathcal{W}(\cdot) \) is decreasing for \( c > c^{FB} \) implies that \( \mathcal{W}(c(s^L, I')) < \mathcal{W}(c(s^L, I)) \), and, therefore, that equilibrium welfare is strictly higher under \( I \) for any electoral rule \( \rho' \) in the second interval. □

**Proof of Proposition 7.** For the case \( c^{FB} \in \left[c(s'^H; I'), c(s'^L; I')\right] \), then Corollary 1 implies that there exists \( \rho^* \) such that \( c^{FB} \) is an equilibrium cutoff under \( I' \), and therefore the first-best outcome is achieved under \( I' \). Consider the case where \( c^{FB} \) does not belong to the above interval. There are two cases to consider. If
\( c^{FB} < c(s^{H}; \mathcal{T}^{'}) \), then choose a rule \( \rho^* \) that results in equilibrium cutoff \( c(s^{H}; \mathcal{T}^{'}) \) under \( \mathcal{T}^{'}. \) Since \( c(s^{H}; \mathcal{T}^{'}) < c(s^{H}; \mathcal{T}) \) (by Lemma 2), then the fact that \( \mathcal{W}(\cdot) \) is decreasing for \( c > c^{FB} \) implies that \( \mathcal{W}(c(s^{H}; \mathcal{T}^{'})) > \mathcal{W}(c(s^{H}; \mathcal{T})) \geq \max_{c \in C^{\text{eqm}}(\rho; \mathcal{T})} \mathcal{W}(c), \) where the last inequality follows from the fact that only cutoffs in the interval \([c(s^{H}; \mathcal{T}), c(s^{L}; \mathcal{T})]\) can be obtained in equilibrium for some electoral rule (Corollary 1). Finally, if \( c^{FB} > c(s^{L}; \mathcal{T}^{'}) \), then a similar argument can be made by choosing a rule \( \rho^* \) that results in equilibrium cutoff \( c(s^{H}; \mathcal{T}^{'}) \) under \( \mathcal{T}^{'}. \)

**Proof of Proposition 8.** To simplify notation, we denote by a “prime” all objects that depend on \( \mathcal{T}^{'}, \) and omit the prime for objects that depend on \( \mathcal{T}. \) By assumption, \((\sigma, c)\) is an equilibrium under \( \mathcal{T} \) where \( \theta^* \) is a partisan for party \( B \) (the case for party \( A \) is similar and, therefore, omitted). Then \( \sigma_{\theta^*}(s) = 0 \) for all \( s \in \mathbb{S}_{\theta^*}, \) and, since \( \sigma \) is optimal given \( c, \) then \( c_{\theta^*}(s^{H}_{\theta^*}) \geq c. \) In addition,

\[
\bar{\kappa}(c_{\theta^*}(s^{H}_{\theta^*})) = \kappa(c_{\theta^*}(s^{H}_{\theta^*}); \sigma) \\
\geq \kappa(c; \sigma) \\
= \rho,
\]

(13)

where the first line follows because \( \sigma_{\theta^*}(s^{H}_{\theta^*}) = 0 \) and \( c_{\theta}(s) \neq c_{\theta^*}(s^{H}_{\theta^*}) \) for all \( s \in \mathbb{S}_{\theta}, \theta \neq \theta^*, \) the second line follows because \( c_{\theta^*}(s^{H}_{\theta^*}) \geq c \) and \( \kappa(\cdot; \sigma) \) is nondecreasing, and the third because \( c \in (-1, 1) \) is a cutoff given \( \sigma \) and \( \kappa(\cdot; \sigma) \) is continuous.

Suppose, in order to obtain a contradiction, that \((\sigma^{'}, c^{'})\) is an equilibrium under \( \mathcal{T}^{'}, \) such that \( \sigma_{\theta^{'}}(s) = 1 \) for all \( s \in \mathbb{S}_{\theta^{'}} \) and, therefore, \( c_{\theta^{'}}(s^{L}_{\theta^{'}}) \leq c'. \) Since \( \mathcal{T}_{\theta^{'}} \) is more informative than \( \mathcal{T}_{\theta^{'}}, \) Lemma 2 implies that \( c_{\theta^{'}}(s^{L}_{\theta^{'}}) \leq c_{\theta^{'}}(s^{H}_{\theta^{'}}), \) and, since \( c_{\theta^{'}}(\cdot) \) is nondecreasing, \( c_{\theta^{'}}(s^{H}_{\theta^{'}}) \leq c'. \) In addition,

\[
\lim_{\omega \downarrow c_{\theta^{'}}(s^{H}_{\theta^{'}})} \bar{\kappa}'(\omega) = \lim_{\omega \downarrow c_{\theta^{'}}(s^{H}_{\theta^{'}})} \sum_{\theta \neq \theta^{'}} \phi(\theta) q_{\theta}(c_{\theta}(S_{\theta}) < \omega \mid \omega) + \phi(\theta^{'}) q_{\theta^{'}}(c_{\theta^{'}}(S_{\theta^{'}}) < \omega \mid \omega) \\
= \bar{\kappa}(c_{\theta^{'}}(s^{H}_{\theta^{'}})) + \lim_{\omega \downarrow c_{\theta^{'}}(s^{H}_{\theta^{'}})} \phi(\theta^{'}) q_{\theta^{'}}(c_{\theta^{'}}(S_{\theta^{'}}) < \omega \mid \omega) \\
> \rho,
\]

(14)

where the first line follows by definition, the second because \( c_{\theta}(s) \neq c_{\theta^{'}}(s^{H}_{\theta^{'}}) \) for all \( s \in \mathbb{S}_{\theta}, \theta \neq \theta^{'}, \) and the third by equation (13) and because Lemma 2 implies that \( c_{\theta^{'}}(s^{H}_{\theta^{'}}) \leq c_{\theta^{'}}(s^{H}_{\theta^{'}}) \) and, therefore, by A4(ii), \( \lim_{\omega \downarrow c_{\theta^{'}}(s^{H}_{\theta^{'}})} \phi(\theta^{'}) q_{\theta^{'}}(c_{\theta^{'}}(S_{\theta^{'}}) < \omega \mid \omega) \)
\( \omega > 0. \)

By Proposition 1, equation (14) implies that \( c' \leq c_{\theta^*}(s_{\theta^*}^H). \) Hence, \( c' = c_{\theta^*}(s_{\theta^*}^H). \) Therefore,

\[
\kappa'(c' ; \sigma') \geq \sum_{\theta \neq \theta^*} \phi(\theta)q_\theta(c_\theta(S_\theta) < c' | c') + \phi(\theta^*) \\
> \sum_{\theta \neq \theta^*} \phi(\theta)q_\theta(c_\theta(S_\theta) < c' | c') \\
\geq \kappa(c' ; \sigma) \\
\geq \kappa(c ; \sigma) = \rho,
\]

(15)

where the first line follows because \( \sigma^*_\theta(s) = 1 \) for all \( s \in S_\theta^* \) (in fact, this line holds with equality), the second line because \( \phi(\theta^*) > 0 \), the third line because \( I_\theta = I'_\theta \) for all \( \theta \neq \theta^* \) and \( \sigma^*_\theta(s) = 0 \) for all \( s \in S_\theta^* \), and the remaining lines by the fact that \( c' = c_{\theta^*}(s_{\theta^*}^H) \) and by applying the same arguments in equation (13). Since, by assumption, \( c' \) is a cutoff given \( \sigma' \), then Equation (15) and continuity of \( \kappa'(\cdot ; \sigma') \) imply that \( c' = -1 \). Then, since \( c' = c_{\theta^*}(s_{\theta^*}^H) \geq c \), it follows that \( c = -1 \), which contradicts the assumption that \( c \in (-1, 1) \).

\[\Box\]

**Proof of Proposition 9.** Fix any \( \beta \geq 0 \). Simple algebra yields that \( v^0_\theta(s ; c) = v^0_\theta(s ; c) \) if \( c \geq c^{FB} \) and

\[
v^0_\theta(s ; c) = v^0_\theta(s ; c) - \beta \frac{Pr(c < W < c^{FB} | s)}{Pr(c < W | s)} \left[ h(c^{FB}) - E(h(W) | c < W < c^{FB}, S_\theta = s) \right] \\
< v^0_\theta(s ; c)
\]

if \( c < c^{FB} \). One implication is that \( c^\beta_\theta(s) = c^0_\theta(s) \) if \( c^0_\theta(s) \geq c^{FB} \). Then, if \( c^0 > c^{FB} \), it follows that \( c^\beta_\theta(s) = c^0_\theta(s) \) for all \( \theta \) and \( s \in S_\theta \), so that \( \kappa^\theta(\cdot) = \kappa^0(\cdot) \), and, by Proposition 1, \( C^{eqm,\beta}(\rho) = C^{eqm,0}(\rho) \) for all \( \rho \). Another implication of the way we wrote \( v^\beta_\theta \) above is that \( \lim_{\beta \rightarrow \infty} c^\beta_\theta(s) = c^{FB} \) if \( c^0_\theta(s) < c^{FB} \). Then, if \( c^0 > c^{FB} \), it follows that \( c^\beta_\theta(s) < c^{FB} \) and \( \lim_{\beta \rightarrow \infty} c^\beta_\theta(s) = c^{FB} \in (-1, 1) \) for all \( \theta \) and \( s \in S_\theta \). Fix any sequence \( (c^\beta_\theta) \) where \( c^\beta \) is the equilibrium cutoff under an optimal electoral rule \( \rho^\beta \). Since \( c^\beta_\theta(s) < c^{FB} \) for all \( \theta, s \), then \( c^\beta = \bar{c^\beta} < c^{FB} \) for all \( \beta \) (by Proposition 1) and
\[ \lim_{\beta \to \infty} c^\beta = c^{FB}. \] By definition of personal cutoffs \( c^\beta_\theta(s) \), it follows that
\[
E\left(u^\beta(A, W) \mid W > c^\beta, S_\theta = s\right) \geq E\left(u(B, W) \mid W < c^\beta, S_\theta = s\right)
\]
\[
\rightarrow E\left(u(B, W) \mid W < c^{FB}, S_\theta = s\right)
\]
for all \( \theta \) and all \( s \in S_\theta \). Therefore, choosing any \( \theta^* \in \Theta \),
\[
\lim_{\beta \to \infty} W^\beta(c^\beta) = \lim_{\beta \to \infty} \sum_{s \in S_{\theta^*}} \Pr(s) \left[ E\left(u^\beta(A, W) \mid W > c^\beta, S_{\theta^*} = s\right) \Pr(W > c^\beta \mid S_{\theta^*} = s) + E\left(u(B, W) \mid W < c^\beta, S_{\theta^*} = s\right) \Pr(W < c^\beta \mid S_{\theta^*} = s) \right] \geq \sum_{s \in S_{\theta^*}} \Pr(s) E\left(u(B, W) \mid W < c^{FB}, S_{\theta^*} = s\right) = E\left(u(B, W) \mid W < c^{FB}\right).
\]
Since information is aggregated, Proposition 2 implies that either \( c^0 > c^{FB} \) or \( c^0 < c^{FB} \), so that there are no other cases to consider. □

**Proof of Proposition 10.** First, note that, since there is no private information, Proposition 1 implies that, for any electoral rule \( \rho \in (0,1) \), the voting equilibrium cutoff \( c(p_A, p_B) \) for \( p_A + p_B > 0 \) is the unique solution to \( v(c(p_A, p_B)) = 0 \), where
\[
v(c) \equiv p_A E\left(u_A(W) \mid W \geq c\right) - p_B E\left(u_B(W) \mid W \leq c\right)
\]
for all \( c \in \Omega \). Suppose that \( p_A > 0 \). Then, by B2, \( v(c_A) = -p_B E\left(u_B(W) \mid W < c_A\right) < 0 \) and \( v(c_B) = p_A E\left(u_A(W) \mid W > c_B\right) > 0 \). Therefore, the voting equilibrium cutoff \( c(p_A, p_B) \in (c_A, c_B) \). By the implicit function theorem, the sign of \( \frac{\partial v(p_A, p_B)}{\partial p_B} \) is given by the sign of \( E\left(u_B(W) \mid W < c(p_A, p_B)\right) > 0 \) where the inequality follows since \( c(p_A, p_B) < c_B \). Since party B maximizes the probability of being elected by choosing a policy that results in the highest possible election cutoff, her best response is to choose \( p_B = 1 \). A similar argument yields that a best response to \( p_B > 0 \) is to choose \( p_A = 1 \). Therefore, \( p_A = p_B = 1 \) is a Nash equilibrium and, since \( E\left(u_B(W) \mid W < c(p_A, p_B)\right) > 0 \) and \( E\left(u_A(W) \mid W > c(p_A, p_B)\right) > 0 \), equilibrium welfare is higher than under the neutral policy that yields a payoff of zero.

It remains to show that \( p_A = p_B = 0 \) is not a Nash equilibrium. Suppose that \( p_A = 0 \). Then, for any \( p_B > 0 \), the equilibrium cutoff is \( c_B \) and the probability that B
wins the election is $G(c_B)$. For $p_B = 0$, then party B wins the election with probability .5. Therefore $p_B = 0$ is a best response if and only if $G(c_B) \leq .5$. A similar argument shows that $p_A = 0$ is a best response to $p_B = 0$ if and only if $1 - G(c_A) \leq .5$. But then $G(c_A) \geq G(c_B)$, which contradicts assumptions (B2) and (A4(i)). □

**Proof of Lemma 3**

The proof of Lemma 3 relies on the following results, which are proven at the end of this section.

**Lemma 3.1.** Suppose that there exists $\alpha$ and $\Xi'$ with $\Phi(\Xi') > 0$ such that $\alpha$ is $\Xi'$-asymptotically interior and for all $\xi \in \Xi'$

$$\lim_{n \to \infty} ||\sigma^n(\xi; \alpha) - \sigma|| = 0,$$

where $\sigma$ is increasing. Then, $\alpha$ is $\Xi'$-asymptotically $c$-cutoff, where $\kappa(c; \sigma) = \rho$, and for all $i, s_i$,

$$\lim_{n \to \infty} \Delta_i(P^n(\alpha(\xi)), s_i) = v_{\theta_i}(s_i; c)$$

almost surely in $\Xi'$.

**Lemma 3.2.** There exists $\bar{\varepsilon}$ such that for all $\varepsilon < \bar{\varepsilon}$: If $\sigma$ is a limit $\varepsilon$-equilibrium, then it is increasing.

**Proof of Lemma 3:** Let $\varepsilon \leq \bar{\varepsilon}$, where $\bar{\varepsilon}$ is defined by Lemma 3.2. Suppose that $\sigma$ is a limit $\varepsilon$-equilibrium with corresponding $\varepsilon$-equilibrium mapping $\alpha$ and convergence in a set $\Xi'$. By Lemma 3.2, $\sigma$ is increasing. Therefore, all the hypotheses of Lemma 3.1 are satisfied, implying that $\sigma$ is a $c$-cutoff limit $\varepsilon$-equilibrium, where $\kappa(c; \sigma) = \rho$. In addition, Lemma 3.1 implies that $\lim_{n \to \infty} \Delta_i(P^n(\alpha(\xi)), s_i) = v_{\theta_i}(s_i; c)$ a.s.-$\Xi'$ and, by continuity of $F_{\theta_i}$ (A5), that $\lim_{n \to \infty} F_{\theta_i}(\Delta_i(P^n(\alpha(\xi)), s_i)) = F_{\theta_i}(v_{\theta_i}(s_i; c))$ a.s.-$\Xi'$.
Therefore, there exists $n_\varepsilon$ such that for all $n \geq n_\varepsilon$, all $i, s_i$
\[
|\alpha_i^n(\xi)(s_i) - F_\theta(v_\theta(s_i; c))| \leq |\alpha_i^n(\xi)(s_i) - F_\theta(\Delta_i(P^n(\alpha(\xi)), s_i))| \\
+ |F_\theta(\Delta_i(P^n(\alpha(\xi)), s_i)) - F_\theta(v_\theta(s_i; c))| \\
\leq 2\varepsilon
\]
a.s.-$\Xi'$, where for the first term in the RHS, we have used the fact that $\alpha$ is an $\varepsilon$-equilibrium mapping. Moreover, the previous inequality and equation (6) imply that for all $n \geq n_\varepsilon$, all $\theta, s_\theta$,
\[
|\sigma^n_\theta(\xi; \alpha)(s_\theta) - F_\theta(v_\theta(s_\theta; c))| \leq 2\varepsilon.
\]
Finally, the previous result and the fact that $\lim_{n \to \infty} \sigma^n_\theta(\xi; \alpha) = \sigma$ for all $\xi \in \Xi'$ imply that there exists $n'_\varepsilon \geq n_\varepsilon$ such that for $n \geq n'_\varepsilon$, all $\theta, s_\theta$,
\[
|\sigma_\theta(s_\theta) - F_\theta(v_\theta(s_\theta; c))| \leq |\sigma_\theta(s_\theta) - \sigma^n_\theta(\xi; \alpha)(s_\theta)| \\
+ |\sigma^n_\theta(\xi; \alpha)(s_\theta) - F_\theta(v_\theta(s_\theta; c))| \\
\leq 3\varepsilon.
\]

Lemma 3 then follows by letting $\gamma(\varepsilon) = 3\varepsilon$. □

In order to prove Lemmas 3.1 and 3.2, we use the following notation. We let $\kappa^n_i(\omega; \xi) \equiv P^n(x_i = A \mid \omega)$ (we use $\kappa^n_i(\omega)$ when $\xi$ is omitted) be the probability that player $i = 1, \ldots, n$ votes for A conditional on the state being $\omega$, and let $\kappa^n(\omega; \xi) \equiv \frac{1}{n} \sum_{i=1}^n \kappa^n_i(\omega; \xi)$ (we use $\kappa^n_i(\omega)$ when $\xi$ is omitted) be the average over all players. We also omit $\alpha$ from the notation: $P^n(\xi) \equiv P^n(\alpha(\xi))$ and $\sigma^n_\theta(\xi)$ denotes the average strategy profile of type $\theta$.

**Proof of Lemma 3.1**

Recall that to show this lemma we assume that: (a) $\alpha$ is $\Xi'$ asymptotically interior, (b) $\lim_{n \to \infty} \sigma^n_\theta(\xi) = \sigma_\theta$ a.s. in $\Xi'$, and (c) $\sigma$ is increasing. The proof relies on the following claims.

**Claim 3.1.1:** $\kappa(\cdot; \sigma)$ is increasing and therefore $\{\omega : \kappa(\omega; \sigma) = \rho\}$ is either empty or
a singleton.

**Proof.** We show that $\kappa(\cdot; \sigma)$ is increasing given that $\sigma$ is increasing. First note that by Bayes theorem and A4(i)(ii), for all $\omega' > \omega$, for all $\theta$, and $s' > s$

$$\frac{q_\theta(s'|\omega')}{q_\theta(s'|\omega)} > \frac{q_\theta(s|\omega')}{q_\theta(s|\omega)} \iff \frac{g_\theta(\omega'|s')}{g_\theta(\omega'|s)} > \frac{g_\theta(\omega|s)}{g_\theta(\omega|s)}.$$  

(where $g_\theta$ is the conditional pdf of $\omega$). Therefore, A6 implies the RHS. Moreover, by Proposition 1 in Milgrom (1981a), $q_\theta(S_\theta \leq s \mid \omega') < q_\theta(S_\theta \leq s \mid \omega)$ for all $s < s^H_\theta$. Note also that, casting $S_\theta = \{s^1_\theta, \ldots, s^H_\theta\}$, it follows that

$$\sum_{s \in S_\theta} \sigma_\theta(s)q_\theta(s|\omega) = \sum_{j=1}^{H_\theta} A_\theta(s^j_\theta - 1) \left( \sum_{s \leq s^j_\theta} q_\theta(s|\omega) \right),$$

where $A_\theta(s^j_\theta) = \sigma_\theta(s^j_\theta) - \sigma_\theta(s^{j+1}_\theta)$ and $A_\theta(s^H_\theta) = \sigma_\theta(s^H_\theta)$. Hence

$$\sum_{s \in S_\theta} \sigma_\theta(s)\{q_\theta(s|\omega') - q_\theta(s|\omega)\} = \sum_{j=1}^{H_\theta-1} A_\theta(s^j_\theta) \left( \sum_{s \leq s^j_\theta} (q_\theta(s|\omega') - q_\theta(s|\omega)) \right).$$

Since $\sigma$ is nondecreasing, $A_\theta(s^j_\theta) < 0$, then the expression above is strictly positive. Since $\phi(\theta) > 0$ all $\theta$, the desired result follows. 

**Claim 3.1.2:** For all $\omega \in \Omega$, $\lim_{n \to \infty} \kappa^n(\omega; \xi) = \kappa(\omega; \sigma)$ a.s. in $\Xi'$.

**Proof.** First, note that

$$\kappa^n(\omega; \xi) = \frac{1}{n} \sum_{i=1}^{n} \sum_{\theta \in \Theta} \sum_{s \in S_\theta} q_\theta(s|\omega)1\{\theta_i(\xi) = \theta\} \alpha^n_i(\xi)(s)$$

$$\quad = \sum_{\theta \in \Theta} \sum_{s \in S_\theta} q_\theta(s|\omega) \left\{ \frac{1}{n} \sum_{i=1}^{n} 1\{\theta_i(\xi) = \theta\} \alpha^n_i(\xi)(s) \right\}$$

$$\quad = \sum_{\theta \in \Theta} \sum_{s \in S_\theta} q_\theta(s|\omega) \left\{ \sigma^n_\theta(\xi)(s) \times \left( \frac{1}{n} \sum_{i=1}^{n} 1\{\theta_i(\xi) = \theta\} \right) \right\}$$

$$\quad \to \sum_{\theta \in \Theta} \sum_{s \in S_\theta} q_\theta(s|\omega) \sigma_\theta(s) \phi(\theta) = \kappa(\omega; \sigma),$$
where convergence is a.s. in $\Xi'$ and follows from (i) the assumption that $\lim_{n \to \infty} \sigma^n_\theta(\xi) = \sigma_\theta$ a.s. in $\Xi'$, (ii) the strong law of large numbers applied to $\frac{1}{n} \sum_{i=1}^{n} 1\{\theta_i(\xi) = \theta\}$, and (iii) the fact that $1\{\cdot\}$ and $\sigma^n_\theta$ are uniformly bounded.

Claim 3.1.3:

$$\lim_{n \to \infty} P^n(\xi)(\omega = A | \omega) = \begin{cases} 0 & \text{if } \rho > \kappa(\omega; \sigma) \\ 1 & \text{if } \rho < \kappa(\omega; \sigma) \end{cases} \quad \text{a.s. in } \Xi'$$

Proof. It follows that

$$P^n(\xi)(\omega = A | \omega) = \Pr \left( n^{-1} \sum_{i=1}^{n} 1\{x^n_i = A\} \geq \rho | \omega \right)$$

$$= \Pr \left( n^{-1/2} \sum_{i=1}^{n} (1\{x^n_i = A\} - \kappa^n_i(\xi | \omega)) \geq \sqrt{n}(\rho - \kappa^n(\xi | \omega)) | \omega \right)$$  \hspace{1cm} (16)

Moreover, by the Markov inequality,

$$\Pr \left( n^{-1/2} \sum_{i=1}^{n} (1\{x^n_i = A\} - \kappa^n_i(\omega; \xi)) \geq \sqrt{M} | \omega \right) \leq (nM)^{-1} \sum_{i=1}^{n} E \left[ (1\{x^n_i = A\} - \kappa^n_i(\omega; \xi))^2 | \omega \right]$$

$$\leq 4M^{-1}$$  \hspace{1cm} (17)

goesto zero as $M \to \infty$.

Suppose that $\rho > \kappa(\omega; \sigma)$. By Claim 3.1.2, $\sqrt{n}(\rho - \kappa^n(\omega; \xi)) \to \infty$ a.s. in $\Xi'$. Therefore, by equations (16) and (17), $\lim_{n \to \infty} P^n(\xi)(\omega = A | \omega) = 0$ a.s. in $\Xi'$. Similarly, if $\rho < \kappa(\omega; \sigma)$ then $\sqrt{n}(\rho - \kappa^n(\omega; \xi)) \to -\infty$ and $\lim_{n \to \infty} P^n(\xi)(\omega = A | \omega) = 1$ a.s. in $\Xi'$.

Proof of Lemma 3.1. First, Claim 3.1.3 and the facts that $\kappa(\cdot; \sigma)$ is increasing (Claim 3.1.1) and continuous (by A4(iii)) imply that there exists $c \in [-1, 1]$ such that $c \in \arg \min_{\omega \in \Omega} |\kappa(\omega; \sigma) - \rho|$ and

$$\lim_{n \to \infty} P^n(\xi)(\omega = A | \omega) = 1\{\omega > c\} \quad \text{a.s. in } \Xi.$$  \hspace{1cm} (18)

Suppose that $c = 1$. Then $\lim_{n \to \infty} P^n(\xi)(\omega = A) = 0$ a.s. in $\Xi'$, therefore contradicting that $\alpha$ is asymptotically interior. A similar argument rules out $c =
−1. Therefore, \( c \in (-1, 1) \), implying that \( \alpha \) is \( \Xi' \)-asymptotically \( c \)-cutoff and that \( \kappa(c; \sigma) = \rho \).

Second, note that,

\[
P^n(\xi)(o = A | \omega) = \sum_{s_i \in \mathcal{S}_{\theta_i}} P^n(\xi)(o = A | \omega, s_i)q_{\theta_i}(s_i | \omega),
\]

hence, under A4(ii), for any \( \omega \in \Omega \) such that \( \lim_{n \to \infty} P^n(\xi)(o = A | \omega) = 0(= 1) \), it must be the case that \( \lim_{n \to \infty} P^n(\xi)(o = A | \omega, s_i) = 0(= 1) \) for all \( s_i \in \mathcal{S}_{\theta_i} \).

Therefore, a.s. in \( \Xi' \),

\[
\lim_{n \to \infty} E_{P^n(\xi)}(u_{\theta_i}(A, W) | o = A, s_i) = \lim_{n \to \infty} \frac{\int_{\Omega} P^n(\xi)(o = A | W, s_i)q_{\theta_i}(s_i | W)u_{\theta_i}(A, W)G(dW)}{\int_{\Omega} P^n(\xi)(o = A | W, s_i)q_{\theta_i}(s_i | W)G(dW)}
= \frac{\int_{\Omega} \lim_{n \to \infty} P^n(\xi)(o = A | W, s_i)q_{\theta_i}(s_i | W)u_{\theta_i}(A, W)G(dW)}{\int_{\Omega} \lim_{n \to \infty} P^n(\xi)(o = A | W, s_i)q_{\theta_i}(s_i | W)G(dW)}
= \frac{\int_{\Omega} 1\{W > c\}q_{\theta_i}(s_i | W)u_{\theta_i}(A, W)G(dW)}{\int_{\Omega} 1\{W > c\}q_{\theta_i}(s_i | W)G(dW)}
= E (u_{\theta_i}(A, W) | W \geq c, s_i),
\]

where the expectation is well-defined because A4(ii) and the fact that \( \alpha \) is asymptotically interior imply that the denominator is greater than zero, where the second line follows from the dominated convergence theorem (since \( u_i \) is assumed to be uniformly bounded), where the third line follows from Claim 3.1.3, and where the last line uses A1 to replace \( 1\{\omega > c\} \) by \( 1\{\omega \geq c\} \).

**Proof of Lemma 3.2**

Throughout the proof let \( \Xi' \) be the set in Definition 11 and fix \( \xi \in \Xi' \) and a strategy mapping \( \overline{\alpha} \) such that 1.-3. in Definition 11 are satisfied. We drop \( \xi \) and \( \overline{\sigma} \) from the notation, let \( P^n \equiv P^n(\overline{\alpha}(\xi)) \) and, for each strategy \( \alpha^n_i \), let \( P^n_{\alpha_i} \equiv P^n(\alpha^n_i, \overline{\alpha} - i(\xi)) \).

The proof relies on the following claims; the proofs of the first three claims appear at the end of this section.

**Claim 3.2.1:** For all \( \delta > 0 \) and \( \omega \in \Omega \), there exits \( n_{\delta, \omega} \) such that for all \( n \geq n_{\delta, \omega} \),

\[
|P^n_{\alpha_i}(o = A | \omega, s_i) - P^n_{\alpha_i'}(o = A | \omega, s'_i)| < \delta \text{ uniformly over } i, s_i, s'_i, \alpha^n_i, \hat{\alpha}^n_i.
\]
Claim 3.2.2: For all $\delta > 0$ there exist $n_\delta$ such that for all $n \geq n_\delta$, $|\Delta_i(P^n, s_i) - \Delta_i(P^n_{\alpha_i}, s_i)| < \delta$ uniformly over $i, s_i, \alpha_i^n$.

Claim 3.2.3: There exists $c > 0$ and $n_c$ such that for all $n \geq n_c$, $\Delta_i(P^n_{\alpha_i}, s'_i) - \Delta_i(P^n_{\alpha_i}, s_i) \geq c$ for all $i$ and $s'_i > s_i$ such that $\alpha^n_i(s'_i) = \alpha^n_i(s_i)$.

Claim 3.2: There exists $c' > 0$ and $n_{c'}$ such that for all $n \geq n_{c'}$, $\Delta_i(P^n, s'_i) - \Delta_i(P^n, s_i) \geq c'$ for all $i$ and $s'_i > s_i$.

Proof of Claim 3.2: Fix any $\alpha^n_i$ such that $\alpha^n_i(s'_i) = \alpha^n_i(s_i)$. By Claims 3.2.2 and 3.2.3, for all $n \geq \max \{n_c, n_\delta\}$

$$
\Delta_i(P^n, s'_i) - \Delta_i(P^n, s_i) \geq \left(\Delta_i(P^n_{\alpha_i}, s'_i) - \delta\right) - \left(\Delta_i(P^n_{\alpha_i}, s_i) + \delta\right) \\
\geq c - 2\delta.
$$

The claim follows by setting $\delta = c/4$ and $c' = c/2 > 0$. $\square$

Proof of Lemma 3.2. The definition of $\varepsilon$-equilibrium implies that for all $i$, $s'_i > s_i$, $n \geq n_\varepsilon$,

$$
\overline{\pi}^n_i(s'_i) - \overline{\pi}^n_i(s_i) \geq F_i(\Delta_i(P^n, s'_i)) - F_i(\Delta_i(P^n, s_i)) - 2\varepsilon. \\
+ F_i(\Delta_i(P^n, s_i) + c') - F_i(\Delta_i(P^n, s_i) + c'),
$$

where we have added and subtracted the same term to the RHS. Let $c' > 0$ be as defined in Claim 3.2. Since $F_i$ is absolutely continuous, then

$$
F_i(\Delta_i(P^n, s_i) + c') - F_i(\Delta_i(P^n, s_i)) = \int_{\Delta_i(P^n, s_i)}^{\Delta_i(P^n, s_i) + c'} f_i(t) \, dt \geq d \cdot c' \equiv c'' > 0,
$$

where the inequality follows from A5. Hence, the sum of the second and fourth terms in the RHS of (19) is at least $c'' > 0$. By Claim 3.2, the sum of the first and last terms in the RHS of (19) is positive. Therefore, for all $i$, $s'_i > s_i$, $n \geq n_\varepsilon$,

$$
\overline{\pi}^n_i(s'_i) - \overline{\pi}^n_i(s_i) \geq c'' - 2\varepsilon > 0.
$$

Since $\sigma^n_\theta(\xi, \alpha)$ are averages of the strategies, then for all $\theta, s' > s$, and $n \geq n_\varepsilon$, it follows that $\sigma^n_\theta(s') - \sigma^n_\theta(s) \geq c'' - 2\varepsilon$. Since $\lim_{n \to \infty} \sigma^n = \sigma$, then it follows that
$\sigma(s') - \sigma(s) \geq c'' - 2\varepsilon > 0$, thus establishing that limit $\varepsilon$-equilibrium are increasing as long as $0 < \varepsilon < \bar{\varepsilon} \equiv c''/2 > 0$. \hfill \Box

Proof of Claim 3.2.1. The proof is divided into 3 steps.

Step 1. We first show that the probability of being pivotal goes to zero; i.e., for all $\omega \in \Omega$, for all $i$, $\lim_{n \to \infty} Piv_n^n = 0$, where

$$Piv_n^n \equiv P_i^n (o = A | \omega) - P_0^n (o = A | \omega),$$

where the “1” and “0” are understood as vectors of the same dimension as $\alpha_i$.

By simple algebra,

$$Piv_n^n = P^n \left( \frac{n}{\sqrt{n-1}} K_n^n + \frac{n - 1}{\sqrt{n-1}} \right) \leq \frac{n}{\sqrt{n-1}} \sum_{j \neq i} Z_{j,\omega}^n + \frac{n - 1}{\sqrt{n-1}} K_{i,\omega}^n,$$

where $Z_{j,\omega}^n \equiv \{1\{x_j^n = A\} - \kappa_{j,\omega}^n\}/V_{\omega}^n$, $V_{\omega}^n \equiv \sqrt{\frac{1}{n-1} \sum_{j \neq i} \kappa_{j,\omega}^n (1 - \kappa_{j,\omega}^n)}$, and $K_{i,\omega}^n \equiv \frac{\rho - \kappa_{i,\omega}^n}{V_{\omega}^n}$. Note that, for a given $n$, $\{Z_{j,\omega}^n\}_{j \neq i}$ are independent, they have zero mean and unit variance. Moreover, by Step 3 below, $\liminf_{n \to \infty} V_{\omega}^n > 0$, so that

$$\sum_{j \neq i} E \left[ \frac{Z_{j,\omega}^n}{\sqrt{n-1}} \right] \leq \frac{2}{\sqrt{n-1}} \to 0 \text{ as } n \to \infty,$$

Hence by Lindeberg-Feller CLT, it follows that, given $\omega$, $\sum_{j \neq i} \frac{Z_{j,\omega}^n}{\sqrt{n-1}} \to N(0,1)$ as $n \to \infty$.

We divide the remainder of the proof in 3 cases: (a) $\frac{n}{\sqrt{n-1}} K_{i,\omega}^n \to -\infty$, (b) $\frac{n}{\sqrt{n-1}} K_{i,\omega}^n \to K \in (-\infty, \infty)$ or (c) $\frac{n}{\sqrt{n-1}} K_{i,\omega}^n \to \infty$ (if necessary, we take a subsequence that converges, which exists since $\{V_{\omega}^n(\xi)\}_n$ and $(\kappa_{\omega}^n(\xi))_n$ are uniformly bounded).

We first explore case (a) (case (c) is symmetrical). Note that, since $\liminf_{n \to \infty} V_{\omega}^n > 0$, then $\frac{\kappa_{\omega}^n}{V_{\omega}^n} \to 0$. Therefore, $\frac{n}{\sqrt{n-1}} K_{i,\omega}^n + \frac{\kappa_{\omega}^n}{V_{\omega}^n} \to -\infty$, so that we can take $n \geq n_{M,\epsilon}$ such that $\sqrt{n} K_{i,\omega}^n + \frac{\kappa_{\omega}^n}{V_{\omega}^n} \leq -M$, where $\mathbb{L}_N(-M) < 0.5\epsilon$ (where $\mathbb{L}_N$ is the standard Gaussian cdf) for any $\epsilon$. Therefore, for all $\epsilon > 0$ there exists $n_{\epsilon,\omega}$ such that for all $n \geq \max\{n_{\epsilon,\omega}, n_{M,\epsilon}\}$:

$$Piv_n^n \leq P^n \left( \sum_{j \neq i} \frac{Z_{j,\omega}^n}{\sqrt{n-1}} \leq -M | \omega \right) \leq 0.5\epsilon + \mathbb{L}_N(-M) < \epsilon,$$
where the first inequality follows from the fact that \( n \geq n_{M, \epsilon} \) and the second follows from CLT and our choice of \( M \).

For case (b) (i.e., \( K \) finite). Let \( \delta > 0 \) be such that \( L_N(K + \delta) - L_N(K - \delta) < 0.5 \epsilon \). Note that since \( \limsup_{n \to \infty} (V_n^\omega \sqrt{n - 1})^{-1} = 0 \), there exists a \( n_{\delta, \omega} \) such that \( (V_n^\omega \sqrt{n - 1})^{-1} < \delta \) for all \( n \geq n_{\delta, \omega} \). Then, it follows for all \( \epsilon > 0 \), there exists \( n_{\epsilon, \omega} \) such that for all \( n \geq \max\{n_{\epsilon, \omega}, n_{\delta, \omega}\} \):

\[
Piv_n^\omega \leq P_n \left( \frac{n}{\sqrt{n - 1}} K_n^\omega - \frac{1}{V_n^\omega \sqrt{n - 1}} \leq \frac{\sum_{j \neq i} Z_{j, i}^n}{V_n^\omega \sqrt{n - 1}} < \frac{n}{\sqrt{n - 1}} K_n^\omega + \frac{1}{V_n^\omega \sqrt{n - 1}} \mid \omega \right) \leq P_n \left( K - \delta < \frac{\sum_{j \neq i} Z_{j, i}^n}{\sqrt{n - 1}} \leq K + \delta \mid \omega \right) \leq 0.5 \epsilon + L_N(K + \delta) - L_N(K - \delta) < \epsilon,
\]

where the third inequality follows from the CLT. We showed that for any convergent subsequence \( (K_n^\omega)_n \), the associated subsequences of probabilities converge to zero, thus this result must hold for the whole sequence.

**Step 2.** Note that:

\[
P_{\alpha_i}^n (o = A \mid \omega, s_i) = \alpha_i^n(s_i)P_{1}^n (o = A \mid \omega) + (1 - \alpha_i^n(s_i))P_0^n (o = A \mid \omega) = P_0^n (o = A \mid \omega) + \alpha_i^n(s_i) (P_1^n (o = A \mid \omega) - P_0^n (o = A \mid \omega)) \equiv P^n (o = A \mid \omega) + \alpha_i^n(s_i)Piv_n^\omega
\]

Therefore

\[
|P_{\alpha_i}^n (o = A \mid \omega, s_i) - P_{\alpha_i'}^n (o = A \mid \omega, s_{i'})| \leq |\alpha_i^n(s_i) - \hat{\alpha}_i^n(s_i)| \cdot |Piv_n^\omega|.
\]

By step 1, it follows that for all \( n \geq n_{\delta, \omega} \): \(|Piv_n^\omega| \leq \delta\). Since \(|\alpha_i^n(s_i) - \hat{\alpha}_i^n(s_i)| \leq 1\) the desired result follows.

**Step 3.** We now show that for all \( \omega \in \Omega \),

\[
\liminf_{n \to \infty} \frac{1}{n - 1} \sum_{j \neq i \neq i} \kappa_{j, i}^n (1 - \kappa_{j, i}^n) > 0. \quad (20)
\]

Fix any \( n \) and \( j \leq n \). By assumption, \( \alpha_j^n(s_j) \in [F_j (-2K), F_j (2K)] \subset (0, 1) \) for all
Therefore, $0 < \kappa_{j,\omega}^n < 1$ for all $\omega$, thus implying equation (20).

**Proof of Claim 3.2.2.** We prove that

$$\lim_{n \to \infty} \left( \mathbb{E}P_n(u_{\theta_i}(A, W) \mid o = A, S = s_i) - \mathbb{E}P_{\alpha_i}^n(u_{\theta_i}(A, W) \mid o = A, S = s_i) \right) = 0;$$

the proof for $o = B$ is similar and therefore omitted. We first show that, for all $i, s_i, \alpha_i$,

$$\mathbb{E}P_{\alpha_i}^n(u_{\theta_i}(A, W) \mid o = A, S = s_i) = \frac{\int_{\Omega} P_{\alpha_i}^n(o = A \mid W, s_i) q_{\theta_i}(s_i \mid W) u_{\theta_i}(A, W) G(dW)}{\int_{\Omega} P_{\alpha_i}^n(o = A \mid W, s_i) q_{\theta_i}(s_i \mid W) G(dW)}$$

is well-defined for sufficiently large $n$. Fix any $i$. A4(ii) and the fact that $\bar{\alpha}$ is asymptotically interior imply that there exists $\pi$ such that for all $n \geq \pi$, there exists $s_i^*$ such that

$$P^n(o = A, s_i^*) = \int_{\Omega} P^n(o = A \mid W, s_i^*) q_{\theta_i}(s_i^* \mid W) G(dW) \geq c > 0,$$

which implies that $\int_{\Omega} P^n(o = A \mid W, s_i^*) G(dW) \geq c > 0$. By Claim 3.2.1, for each $s_i, \alpha_i^n$, $P^n(o = A \mid \omega, s_i^*) - P^n_{\alpha_i}(o = A \mid \omega, s_i)$ converges to zero as $n \to \infty$. Since both probabilities are bounded by one, then the dominated convergence theorem implies that $\int_{\Omega}(P^n(o = A \mid W, s_i^*) - P^n_{\alpha_i}(o = A \mid W, s_i)) G(dW) \to 0$ as $n \to \infty$, uniformly over $\alpha_i$. Therefore, there exists $n, 5c$ such that $\sup_{\alpha_i} |\int_{\Omega}[P^n(o = A \mid W, s_i^*) - P^n_{\alpha_i}(o = A \mid W, s_i)] G(dW)| < .5c$ for all $n \geq n, 5c$. So for all $n \geq \max \bar{n}, n, 5c \equiv \bar{n}, c$,

$$\int_{\Omega} P^n_{\alpha_i}(o = A \mid W, s_i) q_{\theta_i}(s_i \mid W) G(dW) \geq d \int_{\Omega} P^n_{\alpha_i}(o = A \mid W, s_i) G(dW) > .5d c > 0.$$

Hence, $\mathbb{E}P^n_{\alpha_i}(u_{\theta_i}(A, W) \mid o = A, S = s_i)$ is well defined.

By simple algebra, and letting $\Delta P^n_{\alpha_i}(A, \omega, s_i) \equiv P^n(o = A \mid \omega, s_i) - P^n_{\alpha_i}(o = A \mid \omega, s_i)$,
Proof of Claim 3.2.3. For each $O \in \{A, B\}$: Let $g^n_{\alpha_i}(\omega \mid O, s_i) \equiv P^n_{\alpha_i}(d\omega \mid o = O, s_i)$ denote the density of $\omega$ conditional on $o = O$ and $s_i$, and let $G^n_{\alpha_i}(\omega \mid O, s_i) \equiv P^n_{\alpha_i}([W \leq \omega] \mid o = O, s_i)$ denote the cdf. Let $\Delta g^n_{\alpha_i}(\omega \mid O, s'_i, s_i) \equiv g^n_{\alpha_i}(\omega \mid O, s'_i) - g^n_{\alpha_i}(\omega \mid O, s_i)$ and $\Delta G^n_{\alpha_i}(\omega \mid O, s'_i, s_i) \equiv G^n_{\alpha_i}(\omega \mid O, s'_i) - G^n_{\alpha_i}(\omega \mid O, s_i)$.

Then

$$
\Delta_i \left( P^n_{\alpha_i}, s'_i \right) - \Delta_i \left( P^n_{\alpha_i}, s_i \right) = \int_{\Omega} \left( u_{\theta_i}(A, W) \Delta g^n_{\alpha_i}(W \mid A, s'_i, s_i) - u_{\theta_i}(B, W) \Delta g^n_{\alpha_i}(W \mid B, s'_i, s_i) \right) dW
$$

$$
= \int_{\Omega} \left( u'_{\theta_i}(A, W) \Delta G^n_{\alpha_i}(W \mid A, s_i, s'_i) - u'_{\theta_i}(B, W) \Delta G^n_{\alpha_i}(W \mid B, s_i, s'_i) \right) dW
$$

$$
\geq \int_{\Omega \cap \{W \leq \omega\}} u'_{\theta_i}(A, W) \Delta G^n_{\alpha_i}(W \mid A, s_i, s'_i) dW
$$

$$
\geq c_m \cdot c_M \inf_{O \in \{A, B\}, W \in \Omega} u'_{\theta_i}(A, W)
$$

$$
\equiv c > 0
$$

for all $n \geq n'$ (where $\Omega^n$, $c_m \cdot c_M > 0$, and $n'$ are all defined in Claim 3.2.3.1 below), where the first line follows by definition, the second by integration by parts (note how the signals are inverted), the third by Claim 3.2.3.1(i) (see below) and the facts that that $u'_{\theta_i}(A, \omega) > 0$ and $u'_{\theta_i}(B, \omega) < 0$ for all $\omega$, the fourth by Claim 3.2.3.1(ii), and the fifth line by the facts that $c_m \cdot c_M > 0$ and $\inf_{\omega \in \Omega} u'_{\theta_i}(A, \omega) > 0$ (because $u_{\theta_i}$ is continuously differentiable in a compact set $\Omega$ and $u'_{\theta_i}(A, \omega) > 0$ for all $\omega$).
Claim 3.2.3.1: For all \( i \) and \( s_i' > s_i \) such that \( \alpha^n_i(s_i) = \alpha^n_i(s'_i) \): (i) For all \( n \), \( \Delta G^n_{\alpha_i}(\omega \mid O, s_i, s'_i) \geq 0 \) for all \( \omega \) and \( O \in \{A, B\} \); (ii) There exists \( n' \) and \( (\Omega^n)_n \) with \( \Omega^n = [l_n, u_n] \subseteq \Omega \) and \( \lim \inf_{n \to \infty} u_n - l_n = \beta_2 > 0 \) such that for all \( n \geq n' \) and all \( \omega^* \in \Omega^n \setminus \{-1, 1\} \),

\[
\Delta G^n_{\alpha_i}(\omega \mid A, s_i, s'_i) \geq C_M > 0.
\]

Proof of Claim 3.2.3.1. There exists \( z > 0 \) such that for all \( n \) and all \( \omega' > \omega \),

\[
g^n_{\alpha_i}(\omega' \mid O, s'_i)g^n_{\alpha_i}(\omega \mid O, s_i) - g^n_{\alpha_i}(\omega' \mid O, s_i)g^n_{\alpha_i}(\omega \mid O, s'_i) \\
= \frac{P^n_{\alpha_i}(O \mid \omega', s_i)P^n_{\alpha_i}(O \mid \omega, s_i)g(\omega')g(\omega)}{P^n_{\alpha_i}(O, s'_i)P^n_{\alpha_i}(O, s_i)} \left[ q_{\theta_i}(s'_i \mid \omega') q_{\theta_i}(s_i \mid \omega) - q_{\theta_i}(s_i \mid \omega') q_{\theta_i}(s'_i \mid \omega) \right] \\
\geq \frac{z}{P^n_{\alpha_i}(O \mid \omega', s_i)P^n_{\alpha_i}(O \mid \omega, s_i)} \left[ g(\omega')g(\omega)q_{\theta_i}(s'_i \mid \omega) q_{\theta_i}(s_i \mid \omega) (\omega' - \omega) \right] \\
\geq 0 \tag{21}
\]

where the first line uses the fact that \( P^n_{\alpha_i}(O \mid \hat{\omega}, s_i) = P^n_{\alpha_i}(O \mid \hat{\omega}, s'_i) \) for all \( \hat{\omega} \) (because of conditional independence and the fact that \( \alpha^n_i(s_i) = \alpha^n_i(s'_i) \)), the second line follows from A6, and the third line follows because \( z > 0 \) and \( \omega' > \omega \). Therefore, it follows from Milgrom (1981, Proposition 1) that, for all \( n \), \( \Delta G^n_{\alpha_i}(\omega \mid O, s_i, s'_i) \geq 0 \) for all \( \omega \).

(ii) From the proof of Claim 3.2.2, there exists \( n' \) and \( c' > 0 \) such that, for all \( n \geq n' \),

\[
\int_{\Omega} P^n_{\alpha_i}(o = A \mid W, s_i)G(dW) \geq c'
\]

for all \( i, \alpha_i, s_i \). For \( a \in (0, 1) \), let

\[
\omega^n_a = \min \left\{ \omega' : \int_{W \leq \omega'} P^n_{\alpha_i}(o = A \mid W, s_i)G(d\omega) \geq a \cdot c' \right\} \in \Omega.
\]

Fix any \( n \geq n' \). Then

\[
c'/A = \int_{\omega^n_{a,25} \leq W \leq \omega^n_{a,50}} P^n_{\alpha_i}(o = A \mid W, s_i)G(dW) \leq G\left(\omega^n_{a,50}\right) - G\left(\omega^n_{a,25}\right).
\]

Therefore the fact that \( G \) has no mass points (A1) implies that \( \omega^n_{a,50} - \omega^n_{a,25} \geq c_L > 0 \). A similar argument establishes that \( \omega^n_{a,75} - \omega^n_{a,50} \geq c_R > 0 \).
Let \( \Omega^n = [\omega_{0.50}^n - c_m/2, \omega_{0.50}^n + c_m/2] \), where \( c_m \equiv \min\{c_L, c_R\} > 0 \). Then, \( u_n - l_n = c_m > 0 \). In addition, fix any \( \omega^* \in \Omega^n \). Then, by construction,

\[
\int_{\omega < \omega^* - c_m/2} P^n_{\alpha_i}(o = A \mid W, s_i) G(dW) \geq c'/4 \tag{22}
\]

and

\[
\int_{\omega > \omega^* + c_m/2} P^n_{\alpha_i}(o = A \mid W, s_i) G(dW) \geq c'/4. \tag{23}
\]

By integrating each side of (21) twice, first with respect to \( G(d\omega) \) over \( \omega \leq \omega^* \) and second with respect to \( G(d\omega') \) over \( \omega' > \omega^* \), we obtain

\[
\Delta G^n_{\alpha_i}(\omega^* \mid A, s_i, s'_i) = \\
= \frac{z}{P^n_{\alpha_i}(A, s'_i) P^n_{\alpha_i}(A, s_i)} \int_{W' > \omega^*} \int_{W' < \omega^*} P^n_{\alpha_i}(A \mid W', s_i) P^n_{\alpha_i}(A \mid W, s_i) g(W') g(W) q_{\theta_i}(s'_i \mid W) q_{\theta_i}(s_i \mid W) (W' - W) dG(W) dG(W')
\]

\[
\geq z \int_{W' > \omega^* + c_m/2} \int_{W' < \omega^* - c_m/2} P^n_{\alpha_i}(A \mid W', s_i) P^n_{\alpha_i}(A \mid W, s_i) g(W') g(W) q_{\theta_i}(s'_i \mid W) q_{\theta_i}(s_i \mid W) (W' - W) dG(W) dG(W')
\]

\[
\geq z \cdot c_m \cdot d^2 \int_{W' > \omega^* + c_m/2} P^n_{\alpha_i}(A \mid W', s_i) G(dW') \int_{W' < \omega^* - c_m/2} P^n_{\alpha_i}(A \mid W, s_i) G(dW)
\]

\[
\geq z \cdot c_m \cdot d^2 \cdot \left( \frac{c'}{4} \right)^2 \equiv c_M > 0,
\]

where the first inequality follows from \( P^n_{\alpha_i}(A, s'_i) P^n_{\alpha_i}(A, s_i) \leq 1 \), the second from A4, and the third from (22) and (23).