Conditional Retrospective Voting in Large Elections*

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

January 27, 2014

Abstract

We introduce a solution concept in the context of elections with private information by embedding a model of boundedly rational voters into an otherwise standard equilibrium setting. Voters evaluate alternatives based on past performance, but, since counterfactual outcomes remain unobserved, the sample from which they learn is potentially biased. A retrospective voting equilibrium formalizes the idea that voters learn from a biased sample and have systematically biased beliefs in large elections. This approach provides several novel insights regarding the preference and information aggregation properties of elections. When applied to a Downsian setting of two-party competition, we find that, in contrast to the standard Nash equilibrium case, parties have an incentive to exacerbate the degree to which their policy platforms differ. Moreover, this incentive to polarize increases the welfare of the median voter.

*We thank Heski Bar-Isaac, Ernesto Dal Bó, Erik Eyster, Frederico Finan, Gregory Huber, Navin Kartik, Alessandro Lizzier, Santiago Oliveros, Wolfgang Pesendorfer, Andy Skrzypacz, and several seminar participants for their comments. Esponda: Stern School of Business, New York University, 44 W Fourth Street KMC 7-76, New York, NY 10012, iesponda@stern.nyu.edu; Pouzo: Department of Economics, UC Berkeley, 530 Evans Hall #3880, Berkeley, CA 94704, dpouzo@econ.berkeley.edu.
1 Introduction

In the economics literature, voters are often portrayed as sophisticated individuals who have well-defined preferences, can solve complicated signal-extraction problems, and have correct expectations about the distribution of (counterfactual) payoffs.\footnote{In the context of elections with private information that we consider in this paper, see, e.g., Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1997).} The empirical evidence, on the other hand, often finds that voters are poorly informed and have little understanding of ideology and policy.\footnote{See, e.g., Delli Carpini and Keeter (1997) and Converse (2000).} Consistent with the evidence, political scientists often view voters as boundedly rational individuals who vote “retrospectively” and 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.”

Many of the “rational” assumptions in the voting literature are implicitly embodied in the notion of Nash equilibrium. In particular, these assumptions often seem driven by a methodological tradition rather than by a conviction in their empirical validity. Of course, equilibrium analysis and comparative statics lies at the heart of economics and has provided numerous insights in voting and several other contexts.

The main contribution in this paper is to embed a model of boundedly rational voters who learn from the previous performance of the policies into an otherwise standard equilibrium setting. This approach provides several novel insights regarding the preference and information aggregation properties of elections. We then apply our boundedly rational equilibrium framework to a Downsian model of two-party competition and find that, in contrast to the standard Nash equilibrium case, parties have an incentive to exacerbate the degree to which their policy platforms differ. Moreover, this incentive to polarize can actually increase the welfare of the electorate.

Our model captures an important feature of elections that is often overlooked in the literature. To illustrate this feature, consider an election between a Republican and a Democratic candidate in the United States. Voters are likely to use information about the past performance of the parties to determine which party to vote for. For example, voters who are currently unemployed may favor a Democratic candidate if they have experienced better results from previously elected Democratic administrations, compared to Republican administrations, when also unemployed in the
past. This tendency to learn from the past is not limited to political elections. When shareholders vote on takeover proposals, they benefit from learning the outcome of previous takeovers in the same or comparable firms. A similar phenomenon occurs with legislators choosing whether to vote along party lines, union members voting to accept or reject negotiated contracts, and residents voting whether to approve additional funding for school districts.

A key feature in these examples is that, when using past information to evaluate alternatives, voters only observe the performance of the elected alternatives, so that counterfactual outcomes are not observable. For example, we will never find out how Romney would have performed had he been elected President of the U.S. in 2012 instead of Obama. Similarly, shareholders will not learn the benefits of a takeover that is not approved. Consequently, the sample from which voters learn is potentially biased. The reason is that the selection of alternatives is not randomized: To the extent that voters have some private information, they will elect alternatives that are likely to perform better. It is reasonable to assume that voters will not be able to control for unobserved counterfactual outcomes, thus ending up with systematically biased beliefs (see Section 2 for the evidence).

In our setup, there is a continuum of voters and two alternatives. One of the alternatives wins the election if it receives a high enough proportion of votes; otherwise the other alternative wins. Voters have (possibly heterogeneous) payoffs that are increasing in the state of the world for one alternative and decreasing for the other. In addition, voters have some information about the state of the world. For example, in an election between two political parties, the state can represent the fundamentals of the economy. One of the parties might be best at governing during recessions and the other during booms (perhaps because of their different positions on monetary and fiscal policy).

We propose a new solution concept, \textit{retrospective voting equilibrium} (RVE), to formalize the idea that voters learn from a biased sample and have systematically biased beliefs. An RVE consists of a strategy profile and an election cutoff that satisfy two conditions: (i) there is a tie at the cutoff, with one alternative being elected above and the other below the cutoff; (ii) the strategy profile must be optimal given the election cutoff. Optimality 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. This parsimonious characterization of retrospective voting in large elections is a major advantage of the framework.\(^3\)

We compare RVE to the two standard solution concepts in the literature, sincere voting (SV) and Nash equilibrium (NE), and find that RVE exhibits the more realistic properties of these two solution concepts: Behavior is endogenous and depends not only on the characteristic of an individual voter but also on the aggregate characteristics of the electorate (as in NE, but unlike SV); outcomes are significantly affected by both the electoral rule and the precision of information (as in SV, but unlike NE); and there is a significant fraction of nonpartisans (as in SV, but unlike NE).

We also highlight a tension that often precludes information aggregation under RVE but is not present under NE or SV: If information were aggregated, then, in general, one of the alternatives would be observed to have the best performance. But then everyone would want to vote for that alternative irrespective of their private information, which precludes information aggregation in the first place. Contrary to the typical calls for informing the electorate, mistakes are driven by a biased sample and, therefore, welfare might even decrease with better information. The extent of mistakes, however, is limited by the fact that voters, while not fully sophisticated, still tend to penalize alternatives that yield bad outcomes. Finally, the bias tends to run in the direction of overestimating the benefits of the more risky alternative, thus justifying conservatism as a means to mitigate mistakes.\(^4\)

We then turn to the main application of the framework, which is to embed the voting model in a Downsian model of two-party competition, modified to allow for state-contingent payoffs. There are two political parties, \textit{Left} and \textit{Right}, and each of these parties is committed to a “left” and “right” platform, respectively, but they can choose their degree of polarization. To illustrate, suppose that the state of the economy ranges from recession to boom. The \textit{Left} party is ideologically constrained to favor expansionary fiscal policy while the \textit{Right} party is constrained to favor contractionary policy. Expansionary policy does best in a recession but hurts in a boom, while the opposite is true for a contractionary policy. There is also a neutral, hands-off policy that neither helps nor hurts the economy. The \textit{Left} and \textit{Right} parties

\(^3\)In a parallel with the definition of a competitive equilibrium, the role of prices is played here by the election cutoff. Voters take the cutoff as given when optimizing, and the cutoff is determined endogenously in equilibrium.

choose the degree of expansion or contraction in their policies, respectively.

The parties simultaneously choose their policies in order to maximize the chance of being elected. Voters take these policies as given and play an equilibrium of the voting game. We compare the cases where voters play Nash equilibrium (NE) and retrospective voting equilibrium (RVE).

When voters play NE, the policy platforms converge to a neutral policy, and the logic is similar to the standard convergence result (Downs, 1957). The idea is that polarization hurts the chances of a party not only in states that are in the opposite extreme of the policy but also in intermediate states. Thus, the parties end up converging to a common, middle policy.

When voters play RVE, they evaluate parties based on observed, not counterfactual, performance, and the standard logic no longer applies. A 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 its performance. This incentive to polarize under RVE implies that there is a better match between states and policies under RVE compared to NE and, therefore, welfare (of the median voter, under majority rule) will be higher. In the example with RVE voting, the Left party chooses an expansionary policy and tends to be elected during recessions and the Right party chooses a contractionary policy and tends to be elected during booms. Under NE, in contrast, both parties choose a neutral policy that does not respond to a fluctuating economy.

These results highlight an important benefit of polarization that is analogous to the idea of specialization: parties specialize in certain policies and they are elected into office when these policies tend to be best. This feature of two-party competition is present in a context where there is uncertainty about the best policy, the best policy depends on the state of the world, and parties are committed to different ideological platforms (e.g., the Democratic and Republican parties in the U.S.; the Labour and Conservative parties in the U.K.).

Another implication of these results is that, in order to evaluate the functioning of an electoral system, it is misleading to focus exclusively on whether voters are sophisticated or well informed and to ignore the incentives of the parties. Under NE,

---

5 The empirical evidence from the literature on the political business cycle is consistent with the predictions of the model. For example, Hibbs (1977) and Alesina and Roubini (1992) find that left-wing governments tend to expand the economy when elected and right-wing governments, which are presumably more concerned with inflation rather than unemployment, cause a slowdown.
voters are as sophisticated as possible, but the rational response of parties results in a relatively worse outcome compared to a model where voters follow a simple retrospective voting rule.

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 Sanet (2007), Jehiel and Koessler (2008), and Esponda (2008))). Papers that study elections with non-Nash solution concepts include Osborne and Rubinstein (2003), Eyster and Rabin (2005), Costinot and Kartik (2007), and Martinelli (2011).6

Bendor et al. (2010, 2011) postulate a dynamic model of retrospective voting where voters follow a satisficing rule and vote for the incumbent if it has performed well given their endogenous aspiration level. Spiegler (2013) studies a dynamic model of reforms in which an infinite sequence of policy makers care about the public evaluation of their interventions. The public follows a simple attribution rule and (mistakenly) attributes changes in outcomes to the most recent intervention. In these papers, voters have no private information and, therefore, there is no role for the type of sample bias studied in this paper.7

Our model is inspired by the notion of retrospective voting advanced by Key (1966) and Fiorina (1981), among others. Our work, however, is conceptually very different to the formal literature in retrospective voting, beginning with Barro (1973) and Ferejohn (1986), which studies elections as incentive mechanisms that hold politicians accountable. Instead, our model follows Downs’ (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).

Motivated by the empirical evidence on polarization (e.g., McCarty et al. (2006)), a large literature relaxes the assumptions of the Downsian framework to explain nonconvergent policies.8 While this literature restricts attention to a private values

---

6The original literature on the political business cycle also assumed boundedly rational voters (Nordhaus, 1975).

7Callander (2011) studies a model of dynamic policy-making where “rational” voters learn the mapping between policies and outcomes.

8Some explanations include: policy motivated candidates with uncertainty about median voter preferences (Wittman (1977), Calvert (1985)), the threat of entry by a third party (Palfrey, 1984), the effect of executive-legislative compromise (Alesina and Rosenthal, 2000), lack of policy commitment (Alesina, Osborne and Slivinski (1996), Besley and Coate (1997)), candidates with “valence” attributes, (Aragonés and Palfrey (2002), Gul and Pesendorfer (2009), Kartik and McAfee (2007)), differentiation in the presence of multiple constituencies (Eyster and Kittsteiner, 2007), and convex
setting, McMurray (2013) recently considers a pure value setting and shows that Nash voting leads to convergent policies when parties can commit and are office-motivated, which is inefficient because policies do not match the state of the world. In contrast, we allow for both private and common value elements in voter preferences and show that polarization obtains when voters are boundedly rational.

In Section 2, we discuss the behavioral assumptions underlying our solution concept and the related evidence. In Section 3, we introduce the framework and the solution concept. In Section 4, we study the degree of polarization under two-party competition and, in Section 5, we provide a game-theoretic foundation for our solution concept. We conclude in Section 6 by mentioning possible extensions.

2 Foundation for conditional retrospective voting

In the dynamic environment that motivates our solution concept, voters evaluate alternatives by their past performance. For every election, each voter first observes a private signal that is correlated with the performance of the alternatives. After observing their signals, voters simultaneously cast a vote. The election outcome is then determined and voters observe the performance of the elected alternative. The environment is stationary, in the sense that the signals and performances of the alternatives are drawn from the same distribution in every election.

Table 1 shows data for eight elections from the point of view of one particular voter. Suppose that there is an election in period 9 in which this voter observes signal $r$ prior to voting. A retrospective voter behaves as follows. First, she uses past elections to form a belief about the performance of each alternative conditional on signal $r$. In this example, alternative $L$ always delivers a payoff of zero, while alternative $R$ delivers an average payoff of $(-1 + 1 + 1)/3 = 1/3$ when the signal is $r$ (i.e., in periods 1, 2, and 6). Then, in period 9, the voter votes for $R$, which is the best alternative given the current evidence.

This retrospective voting rule formally captures the idea of retrospective voting in the political science literature (Key (1966), Fiorina (1981)). This idea has received both empirical (e.g., Kramer (1971), Fiorina (1978), Lewis-Beck and Stegmaier voter preferences (Kamada and Kojima, 2012).

Kartik et al. (2012) show that elections are also inefficient when candidates have socially valuable information about the state of the world.
Table 1: Illustration of retrospective voting rule

(2000), Martorana and Mazza, 2012) and experimental support (Woon (2012), Huber et al. (forthcoming)). More generally, the evidence shows that the electorate is often poorly informed and has little understanding of ideology and policy (e.g., Delli Carpini and Keeter (1997) and Converse (2000)), and voter mistakes do not tend to cancel in the aggregate (e.g., Bartels, 1996).

One novelty with respect to the literature on retrospective voting is that we allow for private information—hence the use of the term conditional retrospective voting. An important insight is that the introduction of private information gives rise to sample selection problems. In Table 1, the voter also observes signal r in periods 3 and 7, but, since L is elected, the voter does not observe the performance of R in those periods. If L and R were randomly chosen each period, the fact that the performance of R is not observed in periods 3 and 7 should not affect beliefs in the long run. The problem, however, is that the election outcome depends on private information that is correlated with performance. In particular, it is likely that the reason why R was not elected in periods 3 and 7 is that voters obtained signals that were relatively unfavorable to R. So, if our voter had been somehow able to observe the counterfactual performance of R in periods 3 and 7, she would have probably observed a relatively bad performance. Thus, the fact that counterfactual performances are not observed likely leads to overestimating the value of electing R. Our model provides a tractable way to account for the systematic bias in beliefs that
results from the selection problem.

As a starting point, we make the stark assumption that voters do not try to control for the selection problem, either because they do not understand the selection problem or because they do not know how to control for it. For example, even a sophisticated voter might have trouble thinking how Romney would have performed as president of the U.S. and including that counterfactual assessment in his overall evaluation of Republican candidates.

The idea that voters do not try to correct for unobserved counterfactuals is consistent with the empirical findings of Achen and Bartels (2004), Leigh (2009), and Wolfers (2009), who find that voters punish politicians for events that are outside of their control. Healy and Malhotra (2010) find that punishment is related to the politician’s response to these events. Our model allows voters to be fairly sophisticated and to condition their learning on private signals, such as campaign platforms, media reports, and economic indicators. This type of naiveté also underlies the winner’s curse in common value auctions and has received robust support in experimental settings (e.g., Thaler (1988), Kagel and Levin (2002), and Charness and Levin (2009)).

In this paper, we focus on the steady state of the previous dynamic voting environment when voters follow the retrospective voting rule described above. This steady state is formally captured by the notion of a (naive) behavioral equilibrium (Esponda, 2008). This solution concept captures the failure of players to account for selection problems and differs from the standard notion of Nash equilibrium.

One of the contributions of the current paper is to characterize the (naive) behavioral equilibrium, which captures the retrospective behavior we want to analyze, as the number of voters goes to infinity. This exercise is analogous to that carried out by Feddersen and Pesendorfer (1997), who characterize Nash equilibrium as the number

---

10 For voting experiments, Guarnaschelli et al. (2000) conclude that subjects’ votes do not deviate much from Nash equilibrium play, although Eyster and Rabin (2005) find that these deviations can be systematically attributed to naiveté. Recently, Esponda and Vespa (2011) show that about half of their subjects are sophisticated. These experiments are only indirect tests of our assumption because they inform subjects of the primitives of the game.

11 A formal proof of this statement is provided by Esponda and Pouzo (2012).

12 In Esponda and Pouzo (2012), we also show that there is a rule that provides a foundation for Nash equilibrium. This pivotal rule is identical to the retrospective rule described above, with the exception that learning only takes place from those elections in which a voter was pivotal. We believe, however, that this rule is unrealistic in the context of this paper, where there is a large electorate, the probability of being pivotal is very small, and, hence, there would be essentially no learning opportunities from the past under this rule. Thus, we believe that the existence of a behavioral foundation for Nash equilibrium under a large electorate remains an open question.
of voters goes to infinity. The formal results appear in Section 5. It turns out that this characterization takes a very convenient and intuitive form, analogous to the notion of a competitive equilibrium in market economies. Thus, we begin by directly postulating this convenient characterization as our definition of retrospective voting equilibrium in Section 3.

3 Voting framework

3.1 Setup

A continuum of voters participate in an election between two alternatives, $R$ (right) and $L$ (left). A state $\omega \in \Omega = [-1, 1]$ 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 $R$ or $L$. Votes are aggregated according to an electoral rule $\rho \in (0, 1)$: Alternative $R$ is elected if the proportion of votes in favor of $R$ is greater or equal than $\rho$; otherwise, $L$ is elected.

We model heterogeneity (in preferences and information) by assuming that each voter is of a particular type $\theta$, where $\phi$ is the full-support probability distribution over the set of types $\Theta \subset \mathbb{R}$. 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 $q_\theta(s \mid \omega)$; let $s_\theta^L$ and $s_\theta^R$ denote the lowest and highest signals in $S_\theta$. The payoff of type $\theta$ is given by $u_\theta(o, \omega)$, where $o \in \{L, R\}$ is the winner of the election.

Let $\sigma_\theta : S_\theta \to [0, 1]$ denote the strategy of type $\theta$, where $\sigma_\theta(s)$ is the probability that type $\theta$ votes for alternative $R$ 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 $\theta$.

We maintain the following assumptions throughout the paper, for all $\theta \in \Theta$:

A1. (i) $u_\theta(R, \cdot) : \Omega \to \mathbb{R}$ is nondecreasing and $u_\theta(L, \cdot) : \Omega \to \mathbb{R}$ is nonincreasing, and one of them is strictly monotone; (ii) $u_\theta(R, \cdot)$ and $u_\theta(L, \cdot)$ are both continuously differentiable, except possibly in a finite number of points, and $\sup_{\theta, o, \omega} |u_\theta(o, \omega)| \leq K < \infty$. 


A2. MLRP: For all $\omega' > \omega$, and $s' > s$:

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

A3. (i) $G$ has a density function $g$, where $\inf_{\Omega} g(\omega) > 0$; (ii) there exists $d > 0$ such that $q_\theta(s|\omega) > d$ for all $\theta \in \Theta, s \in S$ and $\omega \in \Omega$; (iii) $q_\theta(s \mid \cdot)$ is continuous for all $s \in S$.

A4. $\Theta \subset \mathbb{R}$ is a compact interval (a singleton is a special case) and $S_\theta = S$; $u_\theta(R, \omega)$, $u_\theta(L, \omega)$, and $q_\theta(s \mid \omega)$ are jointly continuous in $\Theta \times \Omega$ for all $s \in S$.

Assumptions A1-A2 provide an ordering between states, information, and players’ preferences. Note that A2 is trivially satisfied for types with a unique signal (i.e., no private information). As mentioned in Section 5, we view the case without private information as the limit of environments with private information as information precision vanishes. Assumption A3 rules out “strong signals” in the sense of (Milgrom, 1979). Assumption A4 guarantees uniqueness of the equilibrium outcome and is made only for convenience. Thus, the voting environment essentially coincides with the standard setup in Feddersen and Pesendorfer (1997).

Example 1. (Homogenous informed voters) Consider an election between two candidates with different proposals to lower unemployment. The Left candidate is committed to spending more resources in education and training while the Right candidate is committed to lowering corporate taxes to incentivize employment. The state of the world captures the true underlying cause for unemployment. In high states of the world, unemployment is mostly due to weak demand; in low states of the world, it is due to workers lacking the right skills.

Formally, the state is uniformly distributed in $[-1, 1]$ and there is a unique voter type with payoffs $u(R, \omega) = \omega - 1/2$, $u(L, \omega) = -\omega - 1/2$, so that the payoff from the Left [Right] policy is increasing [decreasing] in the state. In particular, $c^{FB} = 0$ is the first-best election cutoff, i.e., everyone prefers $R$ in states $\omega > c^{FB}$ and $L$ in states $\omega < c^{FB}$. In addition, voters privately observe binary signals $S = \{s^L, s^R\}$ with probability $q(s^R \mid \omega) = .5 + \iota \omega$, where $\iota \in (0, .5]$ is the precision of information. □

13One difference is that we require $u_\theta(L, \cdot)$ and $u_\theta(R, \cdot)$ to be separately monotone, rather than only their difference to be increasing.
3.2 Retrospective voting equilibrium

Let
\[ \kappa(\omega; \sigma) = \int_\Theta \sum_{s \in S} q_\theta(s \mid \omega) \sigma_\theta(s) \phi(d\theta) \]
denote the proportion of votes in favor of alternative \( R \). Assumption A2 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: \( R \) is elected if and only if \( \kappa(\omega; \sigma) \geq \rho \), or, equivalently, for all sufficiently high states. This observation motivates the following definition.14

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

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 \( \omega \in \Omega \), the difference in benefits from electing \( R \) over \( L \) that is perceived by a voter of type \( \theta \) who observes signal \( s \) is

\[ v_\theta(s; \omega) \equiv E(u_\theta(R, W) \mid W \geq \omega, S_\theta = s) - E(u_\theta(L, W) \mid W < \omega, S_\theta = s) . \tag{1} \]

To interpret the above expression, note that, for election cutoff \( \omega \), alternative \( R \) is elected whenever \( W \geq \omega \), so that a voter’s retrospective evaluation of \( R \) 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 that she sincerely believes to have the highest perceived benefit.

**Definition 2.** A strategy profile \( \sigma \) is optimal given an election cutoff \( \omega \) if, for all \( \theta \in \Theta \) and \( s \in S \), \( v_\theta(s; \omega) > 0 \) implies \( \sigma_\theta(s) = 1 \) and \( v_\theta(s; \omega) < 0 \) implies \( \sigma_\theta(s) = 0 \).

14When \( \sigma \), 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-A2, \( v_\theta(\cdot; \omega) \) is increasing. Therefore, any strategy that is optimal given some cutoff must be nondecreasing.

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

A retrospective voting equilibrium requires players to optimize given an election cutoff that is endogenously determined by players’ strategies. In particular, unlike the standard notion of sincere voting, voting behavior now depends endogenously on the primitives of the environment. Moreover, the definition of equilibrium is reminiscent of the definition of a competitive equilibrium in market economies. In the voting context, the role of prices is played by the election cutoff. Voters take the election cutoff as given when they optimize, and their consequent behavior yields that election cutoff. In Section 5, we provide a foundation for Definition 3 by showing that a retrospective voting equilibrium characterizes the naive behavioral equilibrium (Esponda, 2008) of the voting game as the number of voters goes to infinity.

### 3.3 Characterization of RVE

We now characterize retrospective voting equilibrium. For each type \( \theta \) and signal \( s \), define the *personal cutoffs*

\[
c_\theta(s) \equiv \arg \min_{\omega \in \Omega} \left| v_\theta(s; \omega) \right|,
\]

which depend only on the primitives of the environment. Since \( \Omega \) is compact and \( v_\theta(s; \cdot) \) is continuous and increasing (by A1-A3), there exists a unique solution \( c_\theta(s) \) that is nonincreasing in \( s \). Moreover, A4 implies that \( v_\theta(s; \omega) \) is jointly continuous in \((\theta, \omega)\) and, by the Theorem of the Maximum, \( c_\theta(s) \) is continuous in \( \theta \). Thus, we can define \( \underline{c} \equiv \min_\theta c_\theta(s^R) \) and \( \overline{c} = \max_\theta c_\theta(s^L) \) as the lowest and highest personal cutoffs across all types.

If we knew the equilibrium election cutoff \( \omega^* \), then it would be straightforward to characterize the equilibrium strategy: a type \( \theta \) with signal \( s \) such that \( c_\theta(s) < \omega^* \) must satisfy \( v_\theta(s; \omega^*) > 0 \) and, therefore, she will optimally vote for \( R \); similarly, if...
Assuming \( c_\theta(s) > \omega^* \), then she will optimally vote for \( L \). Consequently, we now characterize the set of equilibrium cutoffs. For a possible election cutoff \( \omega \in \Omega \),

\[
\pi(\omega) \equiv \int_\Theta \sum_{\{s : c_\theta(s) < \omega\}} q_\theta(s | \omega) \phi(d\theta)
\]

may be interpreted as the proportion of players that vote for \( R \) in state \( \omega \) when the cutoff is also given by \( \omega \).

**Lemma 1.** \( \pi : \Omega \to [0, 1] \) is left-continuous, increasing over the subdomain \((c, \bar{c})\), and satisfies: \( \pi(\omega) = 0 \) if \( \omega \leq c \) and \( \pi(\omega) = 1 \) if \( \omega > \bar{c} \).

**Proof.** See the Appendix.

**Theorem 1.** For any electoral rule \( \rho \in (0, 1) \), there exists a unique retrospective voting equilibrium cutoff and it is given by \( \bar{\pi}^{-1}(\rho) \in [c, \bar{c}] \).

**Proof.** See the Appendix.

Theorem 1 says that there is a unique equilibrium cutoff and that it is essentially given by the intersection of the function \( \pi \) with the electoral rule \( \rho \). The following examples illustrates how to find a retrospective voting equilibrium.

**Example 1, continued.** Simple algebra yields

\[
v(s; \omega) = E(W | W \geq \omega, s) - E(-W | W < \omega, s) = \frac{1}{2} (1 - \omega^2) + I_{sR} (1 - \omega^3) + \frac{1}{2} (\omega^2 - 1) + I_{sL} (\omega^3 + 1)
\]

where \( I_{sR} = -I_{sL} = 1 \). Since \( v(sR; 0) > 0 > v(sL; 0) \), it is easy to see that the personal cutoffs \( c(s) \), which solve \( v(s; c(s)) = 0 \), satisfy \( c(sR) < c^{FB} < c(sL) \). Then

\[
\pi(\omega) = \begin{cases} 0 & \text{if } \omega \leq c(sR) \\ 0.5 + \omega & \text{if } c(sR) < \omega \leq c(sL) \\ 1 & \text{if } \omega > c(sL) \end{cases}
\]

---

15 The interpretation is exact except when there is a unique type (i.e., \( \Theta \) is a singleton) and \( \omega \) is one of its personal cutoffs.

16 \( \bar{\pi}^{-1} : (0, 1) \to [c, \bar{c}] \) is defined as \( \bar{\pi}^{-1}(\rho) = \inf\{\omega \in \Omega : \pi(\omega) \geq \rho\} \).
By Theorem 1, the equilibrium cutoff as a function of the electoral rule \( \rho \) is given by

\[
\omega^* = \begin{cases} 
  c(s^R) & \text{if } \rho \leq 0.5 - (-\iota c(s^R)) \\
  \frac{1}{\iota} (\rho - 0.5) & \text{if } 0.5 - (-\iota c(s^R)) < \rho < 0.5 + \iota c(s^L) \\
  c(s^L) & \text{if } \rho \geq 0.5 + \iota c(s^L)
\end{cases}
\]

Thus, the first-best outcome can be obtained with our boundedly rational voters if and only if the electoral rule is \( \rho = 1/2 \). In contrast, a rule that requires a supermajority to elect one of the alternatives will inefficiently elect the other alternative too often in equilibrium. The equilibrium is depicted in the left panel of Figure 1 for the case \( \rho > 1/2 \) and \( \iota = 1/2 \). In equilibrium, alternative \( R \) is elected in states higher than \( \omega^* \) and \( L \) is elected in lower states. □

3.4 Implications and comparison to other solution concepts

We use examples to compare retrospective voting equilibrium (RVE) to the two standard solution concepts in the literature: sincere voting (SV) and Nash equilibrium (NE). The main finding is that RVE exhibits the more realistic features of the other solution concepts: behavior is endogenous (as in NE, but unlike SV), but outcomes depend on both the electoral rule and the precision of information, and individual
voting behavior depends on private information for a significant fraction of the electorate (as in SV, but unlike NE). In the Online Appendix, we provide a more general treatment of the implications of retrospective voting.

As a reminder, under SV voters choose the alternative with the highest expected payoff given their private signal. Under NE, voters choose the alternative with the highest expected payoff conditional on their private signal and conditional on the event that their vote is pivotal in equilibrium. Feddersen and Pesendorfer (1997) establish full information equivalence: the NE outcome in a sufficiently large election is essentially the outcome that would arise if voters had perfect information about the state of the world.\(^\text{17}\)

### 3.4.1 Homogenous preferences

We start by considering the case where all voters have the same preferences, so that the main role of elections is to aggregate information.

**Example 1, continued.** Under SV, voters vote for \(R\) if they observe \(s^R\) and for \(L\) if they observe \(s^L\). The election cutoff is given by the intersection of the electoral rule with the proportion of voters choosing \(R\) in each state, which is given by \(q(s^R | \cdot)\). The right panel of Figure 1 shows that SV is efficient (i.e., aggregates information) if and only if majority rule is used, \(\rho = 1/2\). This result is an illustration of Condorcet’s famous “jury theorem”. Under NE, by full information equivalence, the NE outcome is efficient and \(R\) is elected for \(\omega > 0\) and \(L\) is elected for \(\omega < 0\). The striking aspect of NE is that this is true for any (non-unanimous) election rule and for any precision of information \(\iota > 0\), no matter how small. In contrast, these changes in the primitives affect outcomes both under SV and RVE, as shown by the right panel of Figure 1.

**Changes in election rules.** Under SV, any cutoff in the \((-1, 1)\) interval is an equilibrium for some electoral rule; in particular, welfare can be relatively low for extreme cutoffs. Under RVE, in contrast, equilibrium cutoffs always remain in the interval \([c(s^R), c(s^L)]\). The result that RVE mitigates welfare losses from naive voting is fairly general and stems from the fact that behavior changes endogenously with the cutoff: As the cutoff, say, increases beyond zero, alternative \(L\) is elected in positive states of the world, where its performance is relatively weak, thus making voters less

\(\text{\textsuperscript{17}}\text{We restrict attention to the symmetric Nash equilibrium that is characterized by Feddersen and Pesendorfer (1997).}\)
willing to vote for $L$. Under SV, in contrast, voting behavior is exogenous and will not react to changes in observed performance.

Changes in information precision. Under SV, a decrease in information precision flattens $q(s^R \mid \cdot)$, leading to more extreme election cutoffs and lower welfare. The situation is more subtle under RVE. On the one hand, a flatter $q(s^R \mid \cdot)$ leads to a flatter $\bar{\kappa}(\cdot)$ over some range. This effect leads to more extreme equilibrium cutoffs. On the other hand, the personal cutoffs $c(s^R)$ and $c(s^L)$ get closer to zero as information decreases, therefore bringing the equilibrium cutoff closer to the first-best cutoff of zero. Thus, information has an ambiguous welfare effect. This result makes sense because voters learn from a biased sample and have systematically biased beliefs, so there is no reason why better information should mitigate this bias. □

The next example highlights a systematic bias present under RVE. The electorate faces a choice between a safe alternative, which delivers a payoff that does not depend on the state of the world, and a risky alternative, that is superior in some states but inferior in other states.

**Example 2.** (Risky vs. safe alternatives) The state is uniformly distributed in $[-1,1]$ and there is a unique voter type with payoffs $u(R, \omega) = \omega$ and $u(L, \omega) = 0$ for all $\omega \in \Omega$; we call $R$ the risky alternative and $L$ the safe alternative. Once again, $c^{FB} = 0$ is the first-best election cutoff. There are three signals, $S = \{s^L, s^M, s^R\}$, and, for simplicity, we omit the details about the precision of each signal. Consider the case of retrospective voting. First, note that $v(s;0) < 0$ for all $s$, which implies
that all personal cutoffs are negative. Thus, the efficient outcome cannot be an RVE for any voting rule: if the cutoff were efficient, at $\omega^* = 0$, then, since $u(R, \omega) > 0$ for all $\omega > 0$, all voters would prefer, irrespective of their signals, to vote for $R$ rather than the zero-payoff alternative $L$; but then the election cutoff would be $-1$ rather than 0. In equilibrium, the perceived benefits of the risky alternative will have to be lowered by inefficiently electing it in some wrong states of the world, thus providing incentives for voters to be willing to vote for the safe alternative for some signal. Figure 2 depicts the $\bar{\kappa}()$ function for this example. While efficiency cannot be achieved, welfare is maximized by choosing any rule higher than $\rho^*$. The intuition behind this result is that, since voters tend to elect the risky alternative in those states in which it is best, they will overestimate its value and will be biased towards voting for the risky alternative. This bias is partially mitigated by choosing a high enough threshold for electing the risky alternative. These features of the election depend crucially on the selection problem faced by voters who learn from past elections and does not arise either under NE (where the outcome is always efficient) or SV (where there is a rule under which the outcome is efficient). □

Example 2 also illustrates how enriching the theory of retrospective voting with private information can address a particular empirical challenge to the literature. Using data from U.S. presidential elections since 1948, Bartels (2010) finds that certain economic indicators have been better under Democratic presidencies. Campbell (2011) points out that this finding, when combined with the frequent success of Republican candidates, poses a challenge to theories of retrospective voting. How can parties with poorer performance be elected with higher probability? In our retrospective model without private information, the average performance of both parties must be equal in equilibrium (otherwise, voters would always vote for the party with the higher average performance). But, if we allow for private information, it is possible that the party with the worse average performance is elected more than half of the time. For a simple example, suppose that the electoral rule in Example 2 is set low enough so that the equilibrium cutoff is $c(s^R)$. Then $R$ has the worse average performance, $E(u(R, W) \mid W > c(s^R)) < 0$, but it is elected in more than half of the states.
3.4.2 Heterogeneous preferences

The case of heterogeneous preferences illustrates the differences in individual voting behavior. Under SV, a significant fraction of voters are nonpartisan in the sense of responding to their signals, but their behavior is exogenously determined. Under NE, behavior is endogenous, leading to subtle demographic and composition effects, but only a negligible fraction of the electorate is nonpartisan. Under RVE, behavior is endogenous and a significant fraction of the electorate is nonpartisan.

**Example 3.** (Heterogeneous voters) The environment is the same as in Example 1, with the exception that preferences are heterogeneous: Payoffs of type \( \theta \) are
\[
  u_{\theta}(R, \omega) = \omega - 1/2 \quad \text{and} \quad u_{\theta}(L, \omega) = -\omega - 1/2 + \theta
\]
for all \( \omega \in \Omega \), and types are drawn uniformly from \([-1, 1]\). In particular, higher types get higher payoffs under \( R \). For concreteness, suppose that information precision is \( \iota = .5 \) and that the alternative is chosen by majority rule, \( \rho = .5 \).

Under SV, a voter votes \( R \) whenever she either observes \( s^R \) and has type \( \theta > -2/3 \) or she observes \( s^L \) and has type \( \theta > 2/3 \). Thus, the proportion of \( R \)-votes in state \( \omega \) is
\[
  \kappa_{SV}(\omega) = .5 + \omega/3.
\]
The left panel of Figure 3 depicts the election cutoff, \( \omega_{SV} = 0 \). Types below \(-2/3\) always vote \( L \), types above \( 2/3 \) always vote \( R \), and types in between \(-2/3 \) and \( 2/3 \) vote their signal. Under NE, since type \( \theta \) prefers \( R \) whenever
\[
  u_{\theta}(R, \omega) > u_{\theta}(L, \omega),
\]
or, equivalently, \( \omega > -5\theta \), then full information equivalence implies that the proportion of votes for \( R \) is
\[
  \kappa_{NE}(\omega) = .5 + \omega \quad \text{for} \quad \omega \in [-.5, .5].
\]
The left panel of Figure 3 also illustrates the NE cutoff, \( \omega_{NE} = 0 \). Moreover, Feddersen and Pesendorfer (1997) show that voting behavior is as follows. Given a sufficiently
large number of voters, there is a sufficiently small $\varepsilon$-neighborhood of types around $\theta = 0$ such that voters in that neighborhood vote their signal, and everyone else always votes $R$ (if $\theta > \varepsilon$) or always votes $L$ (if $\theta < -\varepsilon$).

Under RVE, recall that the personal cutoffs $c_\theta(s)$ solve $v_\theta(s;c) = 0$. Then

$$\pi(\omega) = q(s^R | \omega) \Pr(c_\theta(s^R) < \omega) + q(s^L | \omega) \Pr(c_\theta(s^L) < \omega).$$

The right panel of Figure 3 depicts the numerical solution for $\pi(\cdot)$ and the RVE cutoff, $\omega_{RVE} = 0$. All types lower than $-.22$ always vote $L$ and all types higher than $.22$ always vote $R$, but all types between $-.22$ and $.22$ vote their signal.

Shift in the distribution of preferences. Suppose that there is an increase (in the first order stochastic dominance sense) in the distribution of voters who prefer $R$. For concreteness, let the new distribution of types be $\phi'(\theta) = .5(1 + \theta)$ for $\theta \in [-1, 1]$. Figure 3 shows the new functions $\bar{\kappa}'_{SV}$, $\bar{\kappa}'_{NE}$, and $\bar{\kappa}'$, under SV, NE, and RVE. As expected, in all cases there is an upward shift in the proportion of votes for $R$, and the cutoff moves to the left, so that $R$ is chosen in more states of the world. But the effect on voting behavior is different in each case.

Under SV, voting behavior is exogenous and so every type behaves exactly as before. The increase in the proportion of votes for $R$ is driven by the fact that there are more high types and these types vote for $R$. In contrast, behavior is endogenously affected by the change in equilibrium cutoff under both NE and RVE. Under NE, the new “marginal type” is $\theta = .42$: lower types always vote $L$ and higher types always vote $R$ (while a vanishing fraction of types around $.42$ vote their signal). Under RVE, types lower than $.09$ always vote $L$, types higher than $.56$ always vote $R$, and types in between $.09$ and $.56$ vote their signal. In particular, only those with more extreme types continue to support $R$ under NE and RVE, but this “extremism” actually helps mitigate the extent to which $R$ is elected after the preference shift.\footnote{In the case of NE, the reason for more extreme supporters of $R$ is that the pivotal voter believes that the state is given by the cutoff state, since this is where the proportions voting for $R$ and $L$ are equal. When preferences shift and the cutoff decreases, then the pivotal voter believes the state is lower. In order to be indifferent between voting for $L$ or $R$, then its type must be higher. In the case of RVE, the reason for more extreme supporters of $R$ is that, if $R$ is elected more often, then it must be elected in worse states and its observed performance must be lower. Therefore, types that were marginally willing to vote for $R$ will no longer desire to vote for $R$.} □

Example 3 also illustrates the more general point that voting behavior under both NE and RVE is influenced by the composition of the electorate. The presence
of composition effects has been documented in empirical work (e.g., Gelman et al. (2008), Leigh (2005)). One novel implication is that individual voting behavior might differ in local vs. national elections, even if the underlying alternatives are similar, because the composition of the electorate is different (see Fiorina (1992), Chari et al. (1997), and Franck and Tavares (2008) for evidence and alternative explanations.)

4 Endogenous policies and the degree of polarization

In this section, we apply the framework to study the degree of polarization in a two-party system. In the first stage, candidates commit to certain policies. In the second stage, voters play a voting equilibrium—we study both Nash (NE) and retrospective voting equilibrium (RVE).

We assume that there are two parties, that these two parties are ideologically constrained to choose policies from different platforms, but that they are free to choose the degree of polarization within their platform (e.g., a Republican candidate chooses whether to be closer to the center or far to the right). These assumptions seem consistent with constraints faced by candidates in the real world (e.g., because of their affiliation to different national parties), and these constrains are arguably unrelated to the type of equilibrium (Nash or RVE) played by voters.\footnote{The two-party system is often attributed to the fact that there is only one winner of the election (e.g., Duverger, 1954). The constraint on policies is often attributed to the weight of the national parties’ distinctive ideologies (e.g., Ansolabehere et al. (2001)) and the influence of closed primaries (e.g., Gerber and Morton (1998)). These assumptions guarantee that the monotonicity assumptions on preferences hold for all possible policy choices.}

We also assume that there is no private information (to be interpreted as the limiting case of a negligible amount of information). This assumption is made for tractability and because it constitutes a relatively small departure from the standard setting of two-party competition (Downs, 1957).\footnote{In fact, with no private information (not even a negligible amount), our setting corresponds to a Downsian setting with random payoffs.} The main insight that arises is that parties have incentives to polarize under RVE, but not NE, and that this polarization actually increases welfare.

The environment is described by \( \{ I, \Omega, g, \Theta, \phi, X, u_\theta \} \), where: \( I = \{ \text{Left}, \text{Right} \} \) is the set of players; \( \Omega = [-1, 1] \) is the state space; \( g \) is the density function over states satisfying \( \inf_{\Omega} g(\omega) > 0 \); \( \Theta \subset \mathbb{R} \) is a compact interval (a singleton is a special case) representing the set of preference types; \( \phi \) is the probability distribution over
types; \( X = [-1,1] \) is the set of policies, with the interpretation that \( x = L < 0 \) represents a Left policy, \( x = R > 0 \) represents a Right policy, and \( x = 0 \) is a Neutral policy; \( u_\theta : X \times \Omega \rightarrow \mathbb{R} \) is the payoff function of type \( \theta \), which is assumed to be bounded, continuously differentiable in \( \Omega \) (except possible in a finite number of points), and jointly continuous in \( X \times \Omega \). We assume (for simplicity) that the election is decided by majority rule and denote the median type by \( \theta_M \). We make the following assumptions.\(^{21}\)

**B1.** For all \( \theta \in \Theta \): \( u_\theta(L, \cdot) \) is decreasing for all \( L < 0 \), \( u_\theta(R, \cdot) \) is increasing for all \( R > 0 \), and \( u(0, \omega) = 0 \) for all \( \omega \in \Omega \).

**B2.** For all \( L < 0 \) and \( R > 0 \), \( u_\theta(L, \omega) \) is decreasing in \( \theta \) and \( u_\theta(R, \omega) \) is increasing in \( \theta \).

**B3.** (i) \( u_{\theta_M}(x, 0) < 0 \) for all policies \( x \neq 0 \); (ii) There exist policies \( L < 0 \) and \( R > 0 \) such that \( E(u_{\theta_M}(L, W)|W \leq 0) > 0 \) and \( E(u_{\theta_M}(R, W)|W \geq 0) > 0 \).

Assumption B1 says that, the higher the state, then the higher the payoff from Right policies and the lower the payoff from Left policies. In addition, there is a Neutral policy \( x = 0 \) with a constant payoff (normalized to zero) that captures the potential for policy convergence among the parties. Assumption B2 says that types are also ordered: higher types have higher payoffs from Right policies and lower payoffs from Left policies. Finally, assumption B3 says that polarized policies are bad in “neutral” states but can be beneficial in “extreme” states from the point of view of the median voter. If the state is \( \omega = 0 \), then the Neutral policy is best and polarized policies \( L < 0 \) and \( R > 0 \) result in lower payoffs. There exist, however, polarized policies \( L < 0 \) and \( R > 0 \) that do on average better than the Neutral policy when evaluated in the Left (\( \omega < 0 \)) and Right (\( \omega > 0 \)) states, respectively.

We fix the previous environment and consider a policy game between the two players or parties, \( \text{Left} \) and \( \text{Right} \). In the first stage, the parties simultaneously choose and commit to policies \( (L, R) \) with the objective of maximizing their probability of winning the election. The \( \text{Left} \) party is restricted to choose a Left or Neutral policy, \( L \leq 0 \), and the \( \text{Right} \) party is restricted to choose a Right or Neutral policy, \( R \geq 0 \). In the second stage, voters play an equilibrium of the voting game, where the policies in the first stage determine voters’ payoffs. We consider two notions of voting equilibrium for the second stage: Nash (NE) and retrospective voting equilibrium (RVE).

\(^{21}\)Formally, the median type is \( \theta_M = \min \{ \theta' : \phi \{ \{ \theta : \theta \leq \theta' \} \} \geq 1/2 \} \).
We assume that parties are sophisticated and know which notion of voting equilibrium is played in the second stage, and we focus on the Nash equilibrium policies of the policy game, which we refer to as policy equilibrium.

Voting behavior under NE and RVE is characterized by Feddersen and Pesendorfer (1997) and the results in Section 3, respectively, for all policies \((L, R) \neq (0, 0)\). For the case \((L, R) = (0, 0)\), we make the natural assumption that, under both NE and RVE, the Left party is elected in states \(\omega < 0\) and the Right party is elected in states \(\omega > 0\).

Finally, to compare outcomes we will use the welfare of the median type, which, for a fixed policy profile \((L, R)\), is

\[
W_{\theta_M}(L, R) \equiv \Pr(W \geq \omega^*) E(u_{\theta_M}(R, W) | W \geq \omega^*) + \Pr(W < \omega^*) E(u_{\theta_M}(R, W) | W < \omega^*),
\]

where \(\omega^*\) is the election cutoff that results either under under NE or RVE.

The following examples satisfy the previous assumptions and illustrate the range of environments to which our results apply.

**Example 1, continued.** Suppose that, in addition to a Left policy that focuses on training and a Right policy that focuses on corporate subsidies, there is also a Neutral policy that mitigates the costs of unemployment by a magnitude that does not depend on whether unemployment is due to weak demand or poor skills. A typical example is welfare policy intended to bring people out of poverty. In particular, the Neutral policy does not depend on the state of the world, and we normalize its payoff to zero. Then \(x = L \leq 0\) represents the weight given to education policies relative to welfare policies, with \(L = -1\) representing all the weight on education: \(u(L, \omega) = L(\omega + 1/2)\). Similarly, \(x = R \geq 0\) represents the weight given to corporate subsidies relative to welfare policies, with \(R = 1\) representing all the weight on subsidies: \(u(R, \omega) = R(\omega - 1/2)\). □

**Example 4.** Two candidates compete in a local union election. The Left and Right candidates adhere to a tough and soft bargaining platform, respectively. All

\[\]
workers/voters have the same quadratic utility

$$\Pi(x, \omega) = -(x - \omega)^2.$$  

The Left candidate commits to a relatively tough demand $x = L \leq 0$ and the Right candidate commits to a relatively soft demand $x = R \geq 0$. The interpretation is that, the higher the state of the world, the higher the firm’s bargaining power and, therefore, the softer is the optimal demand by the union. Workers also observe, as a benchmark, the payoffs $\Pi(0, \omega)$ of a non-unionized sector that is equivalent to implementing a Neutral policy $x = 0$. Workers evaluate their union representative against this benchmark:

$$u(x, \omega) = \Pi(x, \omega) - \Pi(0, \omega) = -x^2 + 2x\omega.$$  

In particular, higher states make it more desirable to adopt softer demands. □

**Example 5.**$^{23}$ The natural rate of unemployment is given by a function $\bar{U}(\omega)$ that is decreasing in the state. The actual unemployment rate $U_\theta$ of a voter of type $\theta > 0$ is

$$U_\theta(x, \omega) = \bar{U}(\omega) + x/\theta > 0,$$

where $x \in [-1, 1]$ is the policy. A policy $x = L < 0$ is a fiscal stimulus and decreases unemployment; a policy $x = R > 0$ is contractionary (e.g., expenditure reduction) and increases unemployment. The Neutral policy $x = 0$ results in the natural unemployment rate. Voters dislike both unemployment and government expenditure, and their utility is given by

$$\Pi_\theta(x, \omega) = -(U_\theta(x, \omega))^2 + x.$$  

At the beginning of a period, the Neutral policy of $x = 0$ is in place and voters observe the effects of this benchmark policy. Then, the party in power implements its chosen policy and voters observe the effects of this policy. Voters’ then assess the extent to which the policy implemented by the party was beneficial; thus

$$u_\theta(x, \omega) = \Pi_\theta(x, \omega) - \Pi_\theta(0, \omega) = -x^2/\theta^2 - 2x\bar{U}(\omega)/\theta + x.$$  

We assume that the median voter prefers the Neutral policy if she believes the state

---

$^{23}$This example is based on a model by Persson and Tabellini (2000, p. 426).
to be $\omega = 0$, i.e., $\theta_M = 2\hat{U}(0)$.

In particular, higher states represent better economic fundamentals and make $x = R > 0$ policies more desirable; similarly, $x = L < 0$ policies are more desirable when fundamentals are bad. Higher types are less affected by economic policy and prefer less stimulus and more expenditure reduction. Finally, the median type prefers the Neutral policy in state $\omega = 0$ but prefers some $x = R > 0$ policy if the economy is better than average and some $x = L < 0$ if it is worse than average. □

The next result shows that, because of the ordering over types, equilibrium is determined by the preferences of the median type.

**Lemma 2.** For any policy profile $(L, R) \neq (0, 0)$, the unique NE election cutoff is given by

$$\omega_{NE}(L, R) = \arg \min_{\omega \in [-1, 1]} |u_{\theta_M}(R, \omega) - u_{\theta_M}(L, \omega)|$$

and the unique RVE election cutoff is given by the personal cutoff of the median type,

$$\omega_{RVE}(L, R) = \arg \min_{\omega \in [-1, 1]} |v_{\theta_M}(\omega)|.$$

Moreover, welfare of the median type under RVE is given by $W_{\theta_M}(L, R) = E(u_{\theta_M}(R, W)|W \geq \omega^*) \geq E(u_{\theta_M}(L, W)|W \leq \omega^*)$ if $\omega^* < 1$ and

$$W_{\theta_M}(L, R) = E(u_{\theta_M}(L, W)|W \leq \omega^*) \geq E(u_{\theta_M}(R, W)|W \geq \omega^*)$$

if $\omega^* > -1$, where $\omega^* = \omega_{RVE}(L, R)$.

**Proof.** See the Appendix. □

The last part of Lemma 2, which will be used repeatedly in the sequel, says that observed equilibrium performance must be equalized, from the perspective of the median type, when voters follow an RVE with an interior election cutoff. The next result compares equilibrium policies and welfare under NE and RVE.

**Proposition 1.** The Neutral policy profile $(0, 0)$ is the unique policy equilibrium when voters play NE, but it is not an equilibrium when voters play RVE. Moreover, any policy equilibrium under RVE yields strictly higher welfare to the median type than the unique policy equilibrium under NE.
**Proof.** Suppose voters play NE. Then B1, B3, and Lemma 2 imply that the \textit{Left} party can guarantee a non-negative election cutoff by choosing $L = 0$ and the right party can guarantee a non-positive election cutoff by choosing $R = 0$. Therefore, when voters play NE, the election cutoff under any policy equilibrium $(L, R)$ is $\omega_{NE}(L, R) = 0$.

Suppose, in order to obtain a contradiction, that $(L, R) \neq (0, 0)$ is a policy equilibrium under NE voting. It follows by B1 and B3 that $u_{\theta_M}(L, 0) = u_{\theta_M}(R, 0) < 0 = u_{\theta_M}(0, 0)$. But then either party has an incentive to deviate to the Neutral policy and move the election cutoff either to the right or left. Finally, we show that $(L, R) = (0, 0)$ is a policy equilibrium under NE. In this case, $\omega_{NE}(0, 0) = 0$. By the previous argument, the \textit{Left} party can guarantee a non-negative election cutoff by choosing $L = 0$; therefore, the \textit{Right} party has no profitable deviation. Similarly, the \textit{Left} party has no profitable deviation.

Now suppose voters play RVE and that $(L, R) = (0, 0)$, so that $\omega_{RVE}(0, 0) = 0$. By B3, there exists $\bar{R}$ such that $E(\theta_M(R, W) | W \geq 0) > 0 = u(0, \cdot)$. Thus, $\bar{R}$ is a profitable deviation for party \textit{Right} because, by B1 and 2, $\omega_{RVE}(\bar{R}, 0) < \omega_{RVE}(0, 0)$.

Finally, consider any policy equilibrium under RVE, $(L, R)$, with election cutoff $\omega^* \equiv \omega_{RVE}(L, R) \leq 0$. By Lemma 2, the equilibrium welfare of the median type is $W_{\theta_M}(L, R) = E(\theta_M(R, W) | W \geq \omega^*)$. Suppose, in order to obtain a contradiction, that $W_{\theta_M}(L, R) \leq 0$, where 0 is the payoff in the policy equilibrium under NE. Then, by B3, there exists $\bar{L}$ such that

$$E(\theta_M(R, W) | W \geq \omega^*) \leq 0 < E(\theta_M(\bar{L}, W) | W \leq 0) \leq E(\theta_M(\bar{L}, W) | W \leq \omega^*) ,$$

where the last inequality follows by B1. Then $\bar{L}$ is a profitable deviation for the \textit{Left} party because, by B1, $\omega_{RVE}(\bar{L}, R) > \omega_{RVE}(L, R)$, thus contradicting that $(L, R)$ is an equilibrium under RVE. The case where $\omega_{RVE}(L, R) > 0$ is similar and, therefore, omitted.

The first result in Proposition 1 says that, when voters play NE, both parties choose the same, Neutral policy. This result extends the standard Downsian logic of the median voter theorem to a setting where there is uncertainty about the best alternative. The idea is that polarization hurts the chances of a party not only

---

24McMurray (2013) recently shows this convergence result in a pure common value setting.
in states that are in the opposite extreme but also in intermediate states. Thus, parties end up converging to a common, middle platform. This simple logic does not apply under RVE. The reason is that retrospective voters evaluate parties based on observed, not counterfactual, performance. A party which chooses a polarized policy wins in those extreme states in which the policy is best, and, therefore, voters will assess the party to have a relatively high performance. In equilibrium, this advantage over the other party cannot persist and will have to be mitigated by electing the party even in intermediate states in which its policy is not superior. Thus, the party with a polarized policy ends up elected in both extreme and intermediate states of the world.

Proposition 1 also says that the welfare of the median type is strictly higher under RVE compared to NE. The reason is that, under RVE, the parties choose different policies and there is a better match between policies and the state of the world. Thus, while Nash voting is efficient in aggregating information for fixed policies, it does poorly if policies are endogenous when compared to retrospective voting.

Example 4 (continued). Figures 4 and 5 illustrate how to find a policy equilibrium for this example under the assumption that the state is uniformly distributed. First, consider the case where voters follow NE. Suppose that \( L = -1/2 \) and \( R = 1/2 \), so that the election cutoff is at \( \omega_{NE}(L, R) = 0 \), which is where the two quadratic utility functions intersect. Then, as shown in the left panel of Figure 4, the Right party can deviate to \( R' = 0 \), move the election cutoff to \( \omega_{NE}(L, R') = -1/4 \), and, therefore, increase its chances of being elected. The unique Nash equilibrium, \((0, 0)\), is shown in the right panel of Figure 4. Any deviation (such as \( R' > 0 \) in the figure) makes a party worse off.

The left panel of Figure 5 shows that \((0, 0)\) is not an equilibrium under RVE because the Right party can decrease the election cutoff to \( \omega^*(0, R') < 0 \) by deviating to policy \( R' \). This is clear from the picture because the integral of \( \Pi(R', \omega) \) is higher than \( \Pi(0, \omega) \) over \( \omega > 0 \), which implies that \( v(0) > 0 \) and, therefore, the election cutoff decreases.\(^{25}\) The right panel of Figure 5 depicts the unique equilibrium under RVE: \((L, R) = (-1/2, 1/2)\) with election cutoff \( \omega^*(-1/2, 1/2) = 0 \). Any deviation, such as \( R' > 0 \) shown in the figure, results in a lower integral and, therefore, a decrease in the \( v(\cdot) \) function and an increase in the election cutoff. Inspection of the right panels

\(^{25}\)It suffices to look at \( \Pi \), rather than \( u \), because the same benchmark term \( E(u(L, W) \mid W \leq 0) = E(u(R, W) \mid W \geq 0) \) is subtracted from payoffs whenever \( L = -R \) and the election cutoff is zero.
Figure 4: Example 4. NE voting: \((-1/2, 1/2)\) is not an equilibrium (left panel); \((0, 0)\) is the unique policy equilibrium (right panel).

Figure 5: Example 4. RVE voting: \((0, 0)\) is not an equilibrium (left panel); \((-1/2, 1/2)\) is the unique policy equilibrium (right panel).
of Figures 4 and 5 also reveals that welfare is strictly higher under RVE compared to NE. In both cases, Left is elected for \( \omega < 0 \) and Right is elected for \( \omega > 0 \). The difference is that the average performance of policy \(-1/2\) is strictly higher than the performance of policy \(0\) in states \( \omega < 0 \), and similarly for policies \(1/2\) and \(0\) in states \( \omega > 0 \).

In the previous example, welfare is not only higher under RVE compared to NE, but it is also efficient, in the sense that a planner who has to commit to two policies and who wants to maximize voter welfare would choose policies \(L = -1/2\) and \(R = 1/2\) and would then decide to elect \(L\) for \( \omega < 0 \) and \(R\) for \( \omega > 0 \). Efficiency does not hold in general, but we now show that a policy equilibrium \((L, R)\) under RVE is constrained-efficient in the sense that it maximizes the welfare of the median type when the election is decided by retrospective voting.

**Proposition 2.** Suppose that voters play RVE.

1. If \((L, R)\) maximizes \(W_{\theta_M}\), then \((L, R)\) is a policy equilibrium.
2. If \((L, R)\) is a policy equilibrium with interior election cutoff, then \((L, R)\) maximizes \(W_{\theta_M}\).

**Proof.** (1) Suppose that \((L, R)\) with election cutoff \(\omega^* \equiv \omega_{RVE}(L, R)\) maximizes \(W_{\theta_M}\). Suppose that Right has a profitable deviation to \(R'\), so that \(\omega' \equiv \omega_{RVE}(L, R') < \omega^* \leq 1\). Then, by Lemma 2,

\[
W_{\theta_M}(L, R') = E(u_{\theta_M}(R', W) \mid W \geq \omega') \geq E(u_{\theta_M}(L, W) \mid W \leq \omega') > E(u_{\theta_M}(L, W) \mid W \leq \omega^*) = W_{\theta_M}(L, R),
\]

where the strict inequality follows by B1. But the above expression contradicts the assumption that \((L, R)\) maximizes \(W_{\theta_M}\). Therefore, Right has no profitable deviation. A similar proof establishes that Left has no profitable deviation. Therefore, \((L, R)\) is a policy equilibrium.

(2) Suppose not, so that \((L', R')\) with election cutoff \(\omega' \equiv \omega_{RVE}(L', R')\) gives strictly higher welfare to type \(\theta_M\) than \((L, R)\) with election cutoff \(\omega^* \equiv \omega_{RVE}(L, R) \in (-1, 1)\), where we assume that \(\omega' \leq \omega^*\) (the case \(\omega' > \omega^*\) is similar and, therefore,
omitted). Then

\[ W_{θM}(L, R) = \mathbb{E}(u_{θM}(L, W) \mid W ≤ ω^*) = \mathbb{E}(u_{θM}(R, W) \mid W ≥ ω^*) \]

\[ < \mathbb{E}(u_{θM}(R', W) \mid W ≥ ω') = W_{θM}(L', R') \]

\[ ≤ \mathbb{E}(u_{θM}(R', W) \mid W ≥ ω^*), \]

where the first two lines follow by Lemma 2 and the assumption that \( W_{θM}(L', R') > W_{θM}(L, R) \) and the last line follows by B1 and the fact that \( ω' ≤ ω^* \). In particular, \( \mathbb{E}(u_{θM}(L, W) \mid W ≤ ω^*) < \mathbb{E}(u_{θM}(R', W) \mid W ≥ ω^*) \) and, therefore, \( ω_{RVE}(L, R') < ω^* \) and, therefore, \((L, R)\) is not an equilibrium because player Right can increase its chances of being elected by deviating to \( R' \). □

In Lemma 4 in the Appendix, we show that there always exists a policy profile \((L, R)\) that maximizes \( W_{θM} \) under RVE. Thus, a corollary of Proposition 2 is that a policy equilibrium under RVE always exists. We conclude by illustrating how Proposition 2 can also be useful for finding a policy equilibrium.

**Example 5 (continued).** Suppose that the state is uniformly distributed and that the natural rate of unemployment is given by \( \bar{U}(ω) = 1 − ω/2 \) and the median type is \( \theta_M = 2\bar{U}(0) = 2 \). First, we argue that there is no equilibrium policy profile with a corner election cutoff. Suppose, for example, that \( R \) is always elected in equilibrium. Then \( \mathbb{E}(u_{θM}(R, W)) = −R^2/4 < 0 \) and, therefore, player Left could profitably deviate to \( L = 0 \) and be elected with probability one. Thus, from now on we analyze cases where the election cutoff is interior. For every policy profile \((L, R)\),

\[ v_{θM}(ω) = \mathbb{E}(u_{θM}(R, W) \mid W ≥ ω) − \mathbb{E}(u_{θM}(L, W) \mid W ≤ ω) \]

\[ = \frac{1}{4} \left( L^2 − R^2 + R + L + ω(R − L) \right). \]

Then, for every \((L, R) \neq (0, 0)\) with interior cutoff, the equilibrium cutoff is given by

\[ c_{RVE}(L, R) = L + R − \frac{R + L}{R − L}. \]

Moreover, by Lemma 2, \( W_{θM}(L, R) = \mathbb{E}(u_{θM}(R, W) \mid W ≥ c_{RVE}(L, R)) \). It is straightforward to check that \((L, R) = (-1/2, 1/2)\) is the unique policy profile that maximizes
$W_{\theta_m}(L, R)$. Then, by Proposition 2, $(-1/2, 1/2)$ is the unique policy equilibrium under RVE. Thus, the Left party chooses to stimulate the economy and is elected in states where unemployment is high, while the Right party chooses a contractionary policy and is elected in states where unemployment is low. On the other hand, the unique equilibrium under NE voting is $(0, 0)$, so that both parties choose the same Neutral policy, which yields the natural rate of unemployment. Welfare is therefore higher for the median type under RVE compared to both NE voting and to a single-party system which implements the ex-ante optimal policy for the median voter, $x = 0$. □

5 Foundation for voting equilibrium

We provide a game-theoretic foundation for RVE by applying the notion of a naive behavioral equilibrium (Esponda, 2008) to a voting environment with a finite number of players.\footnote{A behavioral equilibrium 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.} We begin by defining naive equilibrium for a voting environment where players’ payoffs are now perturbed to guarantee that each alternative is elected with positive probability in equilibrium; Esponda and Pouzo (2012) show that a naive equilibrium corresponds to the steady state of an explicit learning environment. We then turn to our contribution in this section, which is to study the naive equilibrium concept as the number of players goes to infinity. The main result is that the definition of RVE in Section 3 corresponds to the limit of naive equilibria as, first, the number of players goes to infinity and, then, the payoff perturbations vanish.

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 $i = 1, ..., n$, with types $(\theta_1, ..., \theta_n)$, where we now assume that $\Theta$ is a finite set rather than a compact interval (see the end of the section for a discussion). Player $i$’s payoff when the election outcome is $o \in \{R, L\}$ is now

$$u_{\theta_i}(o, \omega) + 1 \{o = L\} \nu,$$

where
where $\nu \in \mathbb{R}$ is a privately-observed payoff perturbation drawn independently for each player from a probability distribution $F_{\theta_i}$. Recall that $K$ is the uniform bound on payoffs postulated in assumption A1. In addition to A1-A3, 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 $f$ satisfies $\inf_{x \in [-2K, 2K]} f_{\theta}(x) > 0$.

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

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

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 $n \to \infty$.\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. Esponda and Pouzo (2012) show that the perturbations are also important for providing a learning foundation for naive equilibrium.}

Assumption A6 is a strengthening of MLRP that establishes a uniform bound on the rate at which the likelihood ratio changes.

Following Harsanyi (1973), for each player there is a threshold perturbation above which the player will vote for $L$ and below which she will vote for $R$. Thus, integrating over such perturbations and noting that $F_{\theta}$ is absolutely continuous, we obtain a (mixed) strategy for each player $i$, $\alpha_i : S \to [0, 1]$, where $\alpha_i(s)$ is the probability of voting for $R$ after observing signal $s$. In addition, each strategy profile $\alpha = (\alpha_1, ..., \alpha_n)$, together with the primitives of the game, induces a distribution $P^n(\alpha)$ over the outcomes of the game $\{R, L\} \times S^n \times \Omega$.

To gain intuition for the notion of a naive 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 $R$ is observed only whenever $R$ is chosen, player $i$ will come to observe that, conditional on observing signal $s$, alternative $R$ yields in expectation $E_{P^n(\alpha)} (u_{\theta_i}(R, W) \mid o = R, S_i = s)$.\footnote{Whenever an expectation $E_P$ has a subscript $P$, this means that the probabilities are taken with respect to the distribution $P$.} A similar expression holds for alternative $L$.\footnote{Esponda and Pouzo (2012) show that the perturbations are also important for providing a learning foundation for naive equilibrium.}
A naive player who observes $\nu$ and $s$ believes that expected utility is maximized by voting for $R$ whenever $\Delta_i(P^n(\alpha), s) - \nu > 0$ and voting for $L$ otherwise, where

$$\Delta_i(P^n(\alpha), s) \equiv E_{P^n(\alpha)}(u_{\theta_i}(R, W) | o = R, S_i = s) - E_{P^n(\alpha)}(u_{\theta_i}(L, W) | o = L, S_i = s)$$

is well-defined because of the payoff perturbations.

**Definition 4.** A strategy profile $\alpha = (\alpha_1, ..., \alpha_n)$ is a (naive) equilibrium of the voting game if for every player $i = 1, ..., n$ and for every $s \in S$,

$$\alpha_i(s) = F_{\theta_i}(\Delta_i(P^n(\alpha), s)).$$

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. Naive players, however, do not account for the correlation between others’ votes and the state of the world (conditional on their own private information); see Esponda and Pouzo (2012) for the proof that equilibrium exists and additional discussion.

### 5.2 Large number of players

Our technical contribution is to analyze (naive) equilibrium as the number of voters goes to infinity. We do so by studying sequences of voting games. We build such sequences by independently drawing infinite sequences of types $\xi = (\theta_1, \theta_2, ..., \theta_n, ...) \in \Xi$ according to the probability distribution $\phi \in \Delta(\Theta)$; we denote the distribution over $\Xi$ by $\Phi$ and we let $\theta_i(\xi)$ denote the type of player $i$, i.e., the $i$th component of $\xi$. 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 $\alpha(\xi) = (\alpha^1(\xi), ..., \alpha^n(\xi), ...)$, where

$$\alpha^n(\xi) = (\alpha^1_n(\xi), ..., \alpha^n_n(\xi))$$

is the strategy profile that is played in the $n$-player game with types $\theta_1, ..., \theta_n$. Let $P^n(\alpha(\xi))$ be the probability distribution over $\{R, L\} \times S^n \times \Omega$ induced by the strategy
profile $\alpha^n(\xi)$ in the $n$-player game. We define two properties of strategy mappings.\footnote{The a.e. in “for a.e. $\xi \in \Xi$” stands for “almost every” and means that there is a set $\Xi'$ with $\Phi(\Xi') = 1$ such that a condition is true for all $\xi \in \Xi'$. The results continue to hold if we only require $\Phi(\Xi') > 0$.}

**Definition 5.** A strategy mapping $\alpha$ is an $\varepsilon$-equilibrium mapping if for a.e. $\xi \in \Xi$ there exists $n_{\varepsilon, \xi}$ such that for all $n \geq n_{\varepsilon, \xi}$

$$\|\alpha_i^n(\xi) - F_{\theta_i(\xi)}(\Delta_i(P^n(\alpha(\xi)), \cdot))\| \leq \varepsilon$$

for all $i = 1, ..., n$. A strategy mapping $\alpha$ is asymptotically interior if, for a.e. $\xi \in \Xi$,

$$\liminf_{n \to \infty} P^n(\alpha(\xi))(o = R) > 0 \quad \text{and} \quad \limsup_{n \to \infty} P^n(\alpha(\xi))(o = R) < 1.$$\footnote{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.}

The first property in Definition 5 requires that, for large enough $n$, players play strategies that constitute an $\varepsilon$ equilibrium. Our notion of limit equilibrium will require 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.\footnote{The a.e. in “for a.e. $\xi \in \Xi$” stands for “almost every” and means that there is a set $\Xi'$ with $\Phi(\Xi') = 1$ such that a condition is true for all $\xi \in \Xi'$. The results continue to hold if we only require $\Phi(\Xi') > 0$.}

The second property requires that the probabilities of choosing $R$ and $L$ remain bounded away from zero as the number of players increases. We will provide a full characterization of equilibria that have such a property.

In addition to characterizing the equilibrium $c$-cutoff, our objective is to characterize the profile of equilibrium strategies. Given a strategy mapping $\alpha$ and a sequence of types $\xi \in \Xi$, let $\sigma^n(\xi; \alpha) : \Theta \to [0, 1]^S$ represent the average strategy of each type in the $n$-player game. Formally, for all $\theta \in \Theta$ and $s \in S$,

$$\sigma^n_\theta(\xi; \alpha)(s) = \frac{\sum_{i=1}^n 1\{\theta_i(\xi) = \theta\} \alpha_i^n(\xi)(s)}{\sum_{i=1}^n 1\{\theta_i(\xi) = \theta\}}$$

whenever $\sum_{i=1}^n 1\{\theta_i(\xi) = \theta\} > 0$, and arbitrary otherwise. We call any element $\sigma : \Theta \to [0, 1]^S$ an average strategy profile and say that $\sigma$ is increasing if $s' > s$ implies $\sigma_\theta(s') > \sigma_\theta(s)$ for every type $\theta \in \Theta$.\footnote{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.}
**Definition 6.** An average strategy profile \( \sigma : \Theta \to [0, 1]^S \) is a *limit \( \varepsilon \)-equilibrium* if there exists an asymptotically interior \( \varepsilon \)-equilibrium mapping \( \alpha \) such that
\[
\lim_{n \to \infty} \| \sigma^n(\xi; \alpha) - \sigma \| = 0 \quad \text{for a.e. } \xi \in \Xi.
\]
An average strategy profile \( \sigma \) is a *limit equilibrium* if it is a limit \( \varepsilon \)-equilibrium for all \( \varepsilon > 0 \).

**Lemma 3.** Let \( \alpha \) be such that \( \lim_{n \to \infty} \| \sigma^n(\xi; \alpha) - \sigma \| = 0 \) for a.e. \( \Xi \), where \( \sigma \) is increasing. Then there exists \( c \in \arg \min_{\omega \in \Omega} |\kappa(\omega; \sigma) - \rho| \) such that, for a.e. \( \xi \in \Xi \) and all \( \omega \in \Omega \),
\[
\lim_{n \to \infty} P^n(\alpha(\xi))(o = R | \omega) = 1\{\omega > c\}. \quad (8)
\]
Moreover, if \( c \in (-1, 1) \), then for a.e. \( \xi \in \Xi \) and for all \( \varepsilon > 0 \) there exists \( n_{\xi, \varepsilon} \) such that for all \( n \geq n_{\xi, \varepsilon} \),
\[
\left\| \Delta_i(P^n(\alpha(\xi)), \cdot) - v_{\theta,i}(\xi)(\cdot; c) \right\| \leq \varepsilon \quad (9)
\]
for all \( i = 1, \ldots, n \).

**Proof.** See the Appendix. \( \square \)

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 \( R \), conditional on \( \omega \), converges to \( \kappa(\omega; \sigma) \). By standard asymptotic arguments, the proportion of votes for \( R \) becomes concentrated around \( \kappa(\omega; \sigma) \). So, for states where \( \kappa(\omega; \sigma) > \rho \), the probability that \( R \) is elected converges to 1. Similarly, for states where \( \kappa(\omega; \sigma) < \rho \), the probability that \( R \) is elected converges to 0. Since \( \sigma \) is increasing, then there is at most one (measure zero) state such that \( \kappa(\omega; \sigma) = \rho \), so that the election outcome and, therefore, the beliefs can be characterized by a cutoff.

We now use Lemma 3 to characterize limit equilibria.

**Theorem 2.** \( \sigma \) is a limit equilibrium if and only if there exists \( c \in (-1, 1) \) such that \( \kappa(c; \sigma) = \rho \) and \( \sigma_\theta(s) = F(v_\theta(s; c)) \) for all \( \theta \in \Theta \) and \( s \in S \).

**Proof.** Only if: Let \( \sigma \) be a limit equilibrium, so that \( \sigma \) is a limit \( \varepsilon \)-equilibrium for all \( \varepsilon > 0 \). Lemma OA in the Online Appendix shows that \( \sigma \) must be increasing. Fix any \( \varepsilon > 0 \) and let \( \alpha \) be the corresponding \( \varepsilon \)-equilibrium mapping that is asymptotically interior. By Lemma 3, equation (8) holds. Moreover, because \( \alpha \) is asymptotically...
interior, then \( c \in (-1, 1) \) and, therefore, (9) also holds. Then, for all \( \theta \in \Theta \), there exists \( \xi \in \Xi \) and \( n' \) such that for all \( n \geq n' \),

\[
\| \sigma - F_{\theta}(v_{\theta}(\cdot; c)) \| \leq \| \sigma - \sigma^n(\xi; \alpha) \| \\
+ \left\| \sum_{i=1}^n \{ \theta_i(\xi) = \theta \} \alpha^n_i(\xi)(s) - \sum_{i=1}^n \{ \theta_i(\xi) = \theta \} F_{\theta}(\Delta_i(P^n(\alpha(\xi)), s)) \right\| \\
+ \left\| \sum_{i=1}^n \{ \theta_i(\xi) = \theta \} F_{\theta}(\Delta_i(P^n(\alpha(\xi)), s)) - F_{\theta}(v_{\theta}(\cdot; c)) \right\|
\]

\[
\leq \varepsilon + \varepsilon + \varepsilon,
\]

where the first inequality follows from definitions and the second inequality follows from: (i) \( \sigma \) being a limit equilibrium implies that \( \lim_{n \to \infty} \| \sigma^n(\xi; \alpha) - \sigma \| = 0 \) for a.e. \( \xi \in \Xi \); (ii) \( \alpha \) is an \( \varepsilon \)-equilibrium mapping; and (iii) equation (9) and continuity of \( F_{\theta} \) (A5). Since the above relationship holds for every \( \varepsilon > 0 \), then \( \| \sigma - F_{\theta}(v_{\theta}(\cdot; c)) \| = 0 \) for all \( \theta \).

**If:** Consider the strategy mapping \( \alpha \) defined by letting players of type \( \theta \) always play \( \sigma_{\theta} \)-i.e., for all \( \xi, s, n \), and \( i \leq n \), \( \alpha^n_i(\xi)(s) = \sigma_{\theta_i(\xi)}(s) \). First, note that \( \sigma^n = \sigma \) converges trivially to \( \sigma \), and \( \sigma \) is increasing because \( F_{\theta} \) and \( v_{\theta}(\cdot; c) \) are increasing (by A1-A3 and A5). Moreover, \( c \in (-1, 1) \) by assumption. Then, equation (8) in Lemma 3 and the dominated convergence theorem imply that \( \alpha \) is asymptotically interior. In addition, for a.e. \( \xi \in \Xi \) and for every \( \varepsilon > 0 \), there exists \( n_{\xi, \varepsilon} \) such that for all \( n \geq n_{\xi, \varepsilon} \),

\[
\left\| \alpha^n_i(\xi) - F_{\theta_i(\xi)}(\Delta_i(P^n(\alpha(\xi)), \cdot)) \right\| = \left\| \sigma_{\theta_i(\xi)} - F_{\theta_i(\xi)}(\Delta_i(P^n(\alpha(\xi)), \cdot)) \right\|
\]

\[
= \left\| F_{\theta_i(\xi)}(v_{\theta_i(\xi)}(\cdot; c)) - F_{\theta_i(\xi)}(\Delta_i(P^n(\alpha(\xi)), \cdot)) \right\| \leq \varepsilon
\]

for all \( i = 1, \ldots, n \), where the first line follows by construction of the strategy and the second line follows by (9) and continuity of \( F_{\theta} \) (A5). Thus, \( \sigma \) is a limit equilibrium. \( \square \)

### 5.3 Vanishing perturbations

We now consider sequences of equilibria where the perturbations vanish. We index games by a parameter \( \eta \) that affects the distribution \( F^n_\theta \) from which perturbations are drawn.
Definition 7. A family of perturbations \( \{ F^n_\eta \}_{\eta \in \mathbb{N}} \), where \( F^n_\eta = \{ F^n_\eta^\theta \}_{\theta \in \Theta} \), is vanishing if for all \( \theta \in \Theta \) and \( \eta \): assumption A5 is satisfied and
\[
\lim_{\eta \to 0} F^n_\eta^\theta(\nu) = \begin{cases} 0 & \text{if } \nu < 0 \\ 1 & \text{if } \nu > 0 \end{cases}
\]

The next two results provide a foundation for the notion of RVE introduced in Section 3.

Theorem 3. Suppose that there exists a vanishing family of perturbations \( \{ F^n_\eta \}_\eta \) and a sequence \((\sigma^n_\eta, c^n_\eta)_\eta \) such that \( \lim_{\eta \to 0}(\sigma^n_\eta, c^n_\eta) = (\sigma, c) \) and where \( \sigma^n \) is a limit equilibrium and \( c^n \) its corresponding cutoff for all \( \eta \). Then \((\sigma, c)\) is a voting equilibrium.

Proof. Theorem 2 implies that \( \sigma^\theta(s) = \lim_{\eta \to 0} \sigma^n_\eta^\theta(s) = \lim_{\eta \to 0} F^n_\eta^\theta(v_\theta(s; c)) \) for all \( \theta \in \Theta \) and \( s \in \mathbb{S} \). Since \( F^n_\eta \) 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 \( c^n \to c \), there exists \( \bar{\eta} \) such that, for all \( \eta < \bar{\eta} \), \( \omega' < c^n \), and, by Theorem 2, \( \kappa(\omega'; \sigma^n) \leq \rho \). Since \( \sigma^n \to \sigma \), 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 4. Suppose that \((\sigma, c)\) is a voting equilibrium with \( c \in (-1, 1) \). Then there exists a vanishing family of perturbations \( \{ F^n_\eta \}_\eta \) and a sequence \((\sigma^n_\eta, c^n_\eta)_\eta \) such that \( \lim_{\eta \to 0}(\sigma^n_\eta, c^n_\eta) = (\sigma, c) \) and where \( \sigma^n \) is a limit equilibrium and \( c^n \) its corresponding cutoff for all \( \eta \).

Proof. See the Appendix.\[^{31}\]

We conclude by making three 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,

\[^{31}\]As shown in the proof, the argument holds for any family of perturbations if \( c \) is the unique equilibrium cutoff and \( \phi(\{ \theta : c_\theta(s) = c \}) = 0 \).
by a sequence of non-arbitrary beliefs. Second, our game-theoretic foundation uses assumption A6, which is stronger than assumption A2 in Section 3. In particular, A2 allows for the case where voters 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 \( \lim_{r \rightarrow \infty} z^r = 0 \). Therefore, the case of no information studied in Sections 3 and ?? must be viewed as the limiting case of an information structure that satisfies A6 but where informativeness vanishes. Finally, in Section 3 we assumed that \( \Theta \) was a compact interval, rather than a finite set, in order to obtain uniqueness of equilibrium and facilitate the application of the framework. However, we can view the case where \( \Theta \) is a compact interval as the limiting case of a sequence of environments where the finite number of elements in \( \Theta \) goes to infinity.

6 Conclusion

We provided a framework that formalizes a previously ignored feature of many elections: voters learn to make decisions by observing past outcomes, but they cannot observe the consequences of outcomes that are never chosen. Voters hold systematically biased beliefs, but these beliefs arise endogenously from the biased sample that derives from the aggregate behavior of all voters. The framework is easy to apply and yields several new insights about large elections. When embedded into a setting of two-party competition, the model predicts that parties with differentiated platforms will tend to exacerbate their differences. This polarization, however, increases the welfare of the median voter.

The model can be generalized in several nontrivial directions: considering ways in which voters could account for selection; allowing for more than two alternatives and letting voters be strategic, as in Myatt, 2007; considering nonstationary environments; and relaxing the monotonicity assumptions on preferences and information that drive our characterization results.\(^{32}\)

\(^{32}\)The definition of behavioral equilibrium that we use in Section 5 also applies to contexts without monotonicity, so the question of characterizing equilibrium with a large electorate for all possible settings is well defined. Recently, there has been interest in relaxing these assumptions under Nash equilibrium (e.g., Bhattacharya (2008)).
References


7 Appendix

Proof of Lemma 1. \( \bar{\kappa}(\cdot) \) is left-continuous: By the Dominated Convergence Theorem, it suffices to show left-continuity of 
\[
q_{\theta} (c_{\theta}(s) < \omega \mid \omega) \equiv \sum_{\{s : c_{\theta}(s) < \omega\}} q_{\theta} (s \mid \omega)
\]
for all \( \theta \in \Theta \). Fix any \( \theta \in \Theta \). Since there are a finite number of personal cutoffs for type \( \theta \) (defined by equation (2)), then for each \( c \in (-1, 1) \) there exists \( \omega'_{\theta} < c \) such that all personal cutoffs of \( \theta \) are outside the interval \([\omega'_{\theta}, c)\). Then, for all \( \hat{\omega} \),
\[
q_{\theta} (c_{\theta}(s) < \omega \mid \hat{\omega}) = q_{\theta} (c_{\theta}(s) < c \mid \hat{\omega}) \quad \text{for all } \omega \in [\omega'_{\theta}, c].
\]
In addition, \( q_{\theta}(s \mid \cdot) \) is continuous by A3(iii). Therefore, \( \lim_{\omega \uparrow c} q_{\theta} (c_{\theta}(s) < \omega \mid \omega) = q_{\theta} (c_{\theta}(S_{\theta}) < c \mid c) \).

\( \bar{\kappa}(\cdot) \) is increasing over \((\underline{c}, \bar{c})\): Let \( \underline{c} < \omega < \omega' < \bar{c} \). Then
\[
\int_{\Theta} \sum_{\{s : c_{\theta}(s) < \omega'\}} q_{\theta} (s \mid \omega') \phi(d\theta) \geq \int_{\Theta} \sum_{\{s : c_{\theta}(s) < \omega\}} q_{\theta} (s \mid \omega') \phi(d\theta)
\]
\[
\geq \int_{\Theta} \sum_{\{s : c_{\theta}(s) < \omega\}} q_{\theta} (s \mid \omega) \phi(d\theta), \quad (10)
\]
where the last inequality follows because, since \( c_{\theta}(\cdot) \) is nondecreasing, the event \( \{c_{\theta}(s) < \omega\} \) is equivalent to \( \{s \leq s_{\theta}(\omega)\} \) for some threshold \( s_{\theta}(\omega) \), and, therefore, MLRP implies that \( \sum_{\{s : c_{\theta}(s) < \omega\}} q_{\theta} (s \mid \omega') \geq \sum_{\{s : c_{\theta}(s) < \omega\}} q_{\theta} (s \mid \omega) \) (see (Milgrom, 1981)).

Next, we show that the inequality in (10) holds strictly. This is trivially true if there exists a positive \( \phi \)-measure of types with personal cutoffs in \([\omega, \omega')\), so suppose that is not the case. Since, by A4, \( c_{\theta}(s^L) \) is continuous in \( \theta \) and \( \Theta \) is a compact interval,
the union of $c_\theta(s_L)$ over all $\theta \in \Theta$ is a compact interval. Given that there is no
positive measure of types with personal cutoffs in $[\omega, \omega']$, then, the facts that $\phi$ has
full support and $\omega' < \bar{c}$ implies that, for all $\theta \in \Theta$, $c_\theta(s_L) \geq \omega > \omega'$ and, therefore,$\{c_\theta(s) < \omega\} \neq \emptyset$. Then, because MLRP holds strictly (by A2), the second inequality
in (10) is strict.

Finally: If $\omega \leq \underline{c}$, then $\{c_\theta(s_\theta) < \omega\} = \emptyset$ for all $\theta$, so that $\bar{\kappa}(\omega) = 0$. Similarly, if $\omega > \bar{c}$, then $\{c_\theta(S_\theta) < \omega\} = S$ for all $\theta$, so that $\bar{\kappa}(\omega) = 1$. □

Proof of of Theorem 1. The proof relies on the following claim.

Claim 1.1 Suppose that $\sigma$ is optimal given election cutoff $c$. Then

$$\kappa(\omega; \sigma) = \int_{\Theta} \left( \sum_{\{s: c_\theta(s) < c\}} q_\theta(s \mid \omega) + \sum_{\{s: c_\theta(s) = c\}} q_\theta(s \mid \omega)\sigma_\theta(s) \right) \phi(d\theta).$$

(11)

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

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

(12)

and equation (11) follows. In addition, for all $\omega > c$, $\kappa(\omega; \sigma) \leq \int_{\Theta} \sum_{\{s: c_\theta(s) \leq c\}} q_\theta(s \mid \omega)\phi(d\theta)$

$$\leq \int_{\Theta} \sum_{\{s: c_\theta(s) < \omega\}} q_\theta(s \mid \omega)\phi(d\theta) = \bar{\kappa}(\omega).$$

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

We now prove Theorem 1. Fix $\rho \in (0, 1)$ and let

$$c^* \equiv \kappa^{-1}(\rho) = \inf\{\omega: \bar{\kappa}(\omega) \geq \rho\}.$$ 

(13)

Note that, by Lemma 1, $c^* \in [\underline{c}, \bar{c}]$. We begin by showing that there exists $\sigma^*$ such
that $(\sigma^*, c^*)$ is a voting equilibrium. Let $\sigma^*$ satisfy (12). It remains to specify $\sigma^*_\theta(s)$
for \((\theta, s)\) such that \(c_\theta(s) = c^*\). First, suppose that \(c^* \notin \{-1, 1\}\). If \(c^*\) is the election cutoff, then \((\theta, s)\) such that \(c_\theta(s) = c^*\) is indifferent between \(R\) and \(L\), 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 \(\hat{\kappa}(\alpha) \equiv \kappa(c^*; \sigma_\alpha^*)\). By Claim 1.1,

\[
\hat{\kappa}(\alpha) = \int_\Theta \left( \sum_{\{\theta: c_\theta(s) < c^*\}} q_\theta(s \mid c^*) + \sum_{\{\theta: c_\theta(s) = c^*\}} q_\theta(s \mid c^*) \alpha \right) \phi(d\theta),
\]

which is continuous in \(\alpha\). First, we establish that \(\hat{\kappa}(0) \leq \rho\). Suppose not, so that \(\hat{\kappa}(0) = \hat{\kappa}(c^*) > \rho\). Since \(\hat{\kappa}\) is left-continuous (Lemma 1), then there exists \(\omega’ < c^*\) such that \(\hat{\kappa}(\omega’ > \rho)\). But then (13) is contradicted. Second, we establish that \(\hat{\kappa}(1) \geq \rho\). Suppose not, so that \(\hat{\kappa}(1) = \lim_{\omega \uparrow c^*} \hat{\kappa}(\omega) < \rho\). Then, there exists \(\omega” > c^*\) such that \(\hat{\kappa}(\omega” < \rho)\). But, since \(\hat{\kappa}(\cdot)\) is increasing (Lemma 1), then (13) is contradicted. Since \(\hat{\kappa}(0) \leq \rho\) and \(\hat{\kappa}(1) \geq \rho\), by continuity of \(\hat{\kappa}\) there exists \(\alpha^*\) such that \(\hat{\kappa}(\alpha^*) = \kappa(c^*; \sigma_{\alpha^*}^*) = \rho\). 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. Next, suppose that \(c^* = -1\) (the case \(c^* = 1\) is similar and, therefore, omitted). Now let \(\alpha^* = 1\); in particular, \(\sigma_{\alpha^*}^*\) is optimal given \(c^*\) (note it would not necessarily be optimal for a different value of \(\alpha^*\)). In addition, we just established above that \(\hat{\kappa}(1) = \kappa(c^*; \sigma_{\alpha^*}^*) \geq \rho\). 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 \(\sigma_{\alpha^*}^*\). Finally, we show that, for all \(c \neq c^*\), there exists no \(\sigma\) such that \((\sigma, c)\) is a voting equilibrium. Suppose, in order to obtain a contradiction, that \((\sigma, c)\) is a voting equilibrium, where \(c < c^*\) (the case \(c > c^*\) is similar and, therefore, omitted). Let \(\omega^* \in (c, c^*)\). Then \(\hat{\kappa}(\omega^*) \geq \kappa(\omega^*; \sigma) \geq \rho\), where the first inequality follows from Claim 1.1 and the second from the fact that \(c\) is an election cutoff given \(\sigma\). But then (13) is contradicted. □

**Proof of Lemma 2.** The equilibrium cutoff under \(j = NE, RVE\) is unique and given by \(\omega_j(L, R) = \inf\{\omega \in \Omega: \hat{\kappa}_j(\omega) \geq 1/2\}\), where \(\hat{\kappa}_j(\omega) = \phi(\{\theta: c_{\theta,j} < \omega\})\), and where \(c_{\theta,NE} = \arg\min_{\omega \in [-1, 1]} |u_\theta(R, \omega) - u_\theta(L, \omega)|\) and \(c_{\theta,CRV}^*\) is the personal cutoff defined in (2). By Theorem 1, the above statement is correct for \(j = RVE\). For the case of Nash equilibrium, \(j = NE\), the statement follows from full information equivalence (Feddersen and Pesendorfer, 1997): at each state, the proportion of peo-
ple voting for an alternative is given by the proportion of people that prefer that
alternative at the given state. By B1-B2, $v_\theta(\omega)$ and $u_\theta(R, \omega) - u_\theta(L, \omega)$ are con-
tinuous in $\omega$ and increasing in $\theta$ and $\omega$; thus $c_{\theta,j}$ is continuous, nonincreasing, and it
is decreasing for all $\theta$ such that $c_{\theta,j} \in (-1, 1)$. Therefore, because $\theta_M$ is the median
type, it follows that $\omega_j(L, R) = c_{\theta_M,j} = \inf\{\omega \in \Omega : \phi(\{\theta : c_{\theta,j} < \omega\}) \geq 1/2\}.$
The last statement in the lemma follows because $v_{\theta_M}$ is increasing. For example, if $\omega^* \equiv \omega_{RVE}(L, R) \in (-1, 1)$, then, since $v_{\theta_M}$ is increasing, it follows that $v_{\theta_M}(\omega^*) = 0.$
Thus, $W_{\theta_M}(L, R) = E (u_{\theta_M}(R, W) | W \geq \omega^*) = E (u_{\theta_M}(L, W) | W \leq \omega^*).$ □

Lemma 4. There exists a policy profile $(L, R)$ that maximizes $W_{\theta_M}$ under RVE.

Proof. By B1, $v_\theta$ is continuous and, therefore, the Theorem of the Maximum implies
that $\omega_{RVE}(L, R)$ is continuous for all $(L, R) \neq (0, 0)$. Then, B1, Lemma 2, and the
Dominated convergence theorem imply that $W_{\theta_M}$ is continuous for all $(L, R) \neq (0, 0)$.
In addition, $W_{\theta_M}(0, 0) = 0$ and, for all $R$,

$$\limsup_{L \to 0} E (u_{\theta_M}(L, W) | W \leq \omega_{RVE}(L, R)) \leq \lim_{L \to 0} u_{\theta_M}(L, -1) = 0$$

where the first inequality follows because $u_{\theta_M}(L, \cdot)$ is decreasing (assumption B1)
and the last equality because $u_{\theta_M}(\cdot, -1)$ is continuous. A similar result holds for
$E (u_{\theta_M}(R, W) | W \geq \omega_{RVE}(L, R))$. Thus

$$\limsup_{(L, R) \to (0, 0)} W_{\theta_M}(L, R) \leq 0.$$ 

Therefore, $W_{\theta_M}$ is upper semi-continuous. Since $[-1, 0] \cup [0, 1]$ is compact, then the
maximum of $W_{\theta_M}$ is attained. □

Proof of Lemma 3. We use the following notation. Let $x_i \in \{R, L\}$ denote
the vote of player $i$, let $\kappa^n_i(\omega; \xi) \equiv P^n (x_i = R | \omega)$ be the probability that player $i = 1, \ldots, n$ votes for $R$ conditional on the state being $\omega$, and let $\kappa^n(\omega; \xi) \equiv \frac{1}{n} \sum_{i=1}^n \kappa^n_i(\omega; \xi)$ be the average over all players.
First, note that, for a.e. \( \xi \in \Xi \), for all \( \omega \in \Omega \),
\[
\lim_{n \to \infty} \kappa^n(\omega; \xi) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{s \in S} q_\theta(s|\omega) 1\{\theta_i(\xi) = \theta\} \alpha^n_\theta(\xi)(s)
\]
\[
= \lim_{n \to \infty} \sum_{\theta \in \Theta} \sum_{s \in S} q_\theta(s|\omega) \left\{ \frac{1}{n} \sum_{i=1}^{n} 1\{\theta_i(\xi) = \theta\} \alpha^n_\theta(\xi)(s) \right\}
\]
\[
= \sum_{\theta \in \Theta} \sum_{s \in S} q_\theta(s|\omega) \left\{ \lim_{n \to \infty} \sigma^n_\theta(\xi; \alpha)(s) \times \left( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} 1\{\theta_i(\xi) = \theta\} \right) \right\}
\]
\[
= \sum_{\theta \in \Theta} \sum_{s \in S} q_\theta(s|\omega) \sigma_\theta(s) \phi(\theta) = \kappa(\omega; \sigma),
\] (14)
where we have used the assumption that \( \lim_{n \to \infty} \|\sigma^n(\xi; \alpha) - \sigma\| = 0 \) a.s.-\( \Xi \) and the strong law of large numbers applied to \( \frac{1}{n} \sum_{i=1}^{n} 1\{\theta_i(\xi) = \theta\} \).

Second, let \( Y^n(\omega; \xi) \equiv n^{-1/2} \sum_{i=1}^{n} (1\{x_i^n = R\} - \kappa^n_i(\omega; \xi)) \). Fix \( \omega \) such that \( \rho > \kappa(\omega; \sigma) \). Then, for a.e. \( \xi \) and for all \( \varepsilon > 0 \), there exists \( n' \) such that, for all \( n \geq n' \),
\[
P^n(\alpha(\xi))(o = R | \omega) = P^n(\alpha(\xi)) (Y^n(\xi | \omega) \geq \sqrt{n}(\rho - \kappa^n(\xi | \omega)) | \omega)
\]
\[
\leq P^n(\alpha(\xi)) (Y^n(\omega; \xi) \geq \sqrt{n\varepsilon} | \omega)
\]
\[
\leq (n^2\varepsilon)^{-1} \sum_{i=1}^{n} E \left[ (1\{x_i^n = R\} - \kappa^n_i(\omega; \xi))^2 \right] | \omega]
\]
\[
\leq 4(n\varepsilon)^{-1},
\]
where the second line follows from (14) and the third from the Markov inequality. A similar argument for the case \( \rho < \kappa(\omega; \sigma) \) thus implies that, for a.e. \( \xi \in \Xi \), for all \( \omega \in \Omega \),
\[
\lim_{n \to \infty} P^n(\alpha(\xi))(o = R | \omega) = \begin{cases} 0 & \text{if } \rho > \kappa(\omega; \sigma) \\ 1 & \text{if } \rho < \kappa(\omega; \sigma) \end{cases}
\]
Third, the facts that \( \kappa(\cdot; \sigma) \) is increasing (because \( \sigma \) is increasing) and continuous (by A3(iii)) imply that there exists \( c \in [-1, 1] \) such that \( c \in \arg\min_{\omega \in \Omega} |\kappa(\omega; \sigma) - \rho| \) and, for a.e. \( \xi \in \Xi \) and all \( \omega \in \Omega \),
\[
\lim_{n \to \infty} P^n(\alpha(\xi))(o = R | \omega) = 1\{\omega > c\}.
\] (15)
Finally, suppose that \( c \in (-1,1) \). For all \( n \) and all \( \omega \in \Omega \),

\[
P^n(\alpha(\xi))(o = R \mid \omega) = \sum_{s \in \mathcal{S}} P^n(\xi)(o = R \mid \omega, S_i = s) q_{\theta_i(\xi)}(s \mid \omega) \tag{16}
\]

for all \( i \leq n \). By (15), (16), and A3(ii), for a.e. \( \xi \in \Xi \) and all \( s \in \mathcal{S} \),

\[
\lim_{n \to \infty} P^n(\alpha(\xi))(o = R \mid \omega, S_i = s) = 0 \quad (= 1) \tag{17}
\]

for \( \omega < c \) (\( \omega > c \)), where convergence is uniform in \( i \leq n \).\(^{33}\) Therefore, for a.e. \( \xi \in \Xi \) and all \( s \in \mathcal{S} \),

\[
\lim_{n \to \infty} E_{P^n(\alpha(\xi))}(u_{\theta_i(\xi)}(R, W) \mid o = R, S_i = s) = \int_{\Omega} \frac{\sum_{\mathcal{S}} P^n(\alpha(\xi))(o = R \mid W, S_i = s) q_{\theta_i(\xi)}(s \mid W) u_{\theta_i(\xi)}(R, W) G(dW)}{\sum_{\mathcal{S}} P^n(\alpha(\xi))(o = R \mid W, S_i = s) q_{\theta_i(\xi)}(s \mid W) G(dW)}
\]

\[
= \int_{\Omega} \lim_{n \to \infty} P^n(\alpha(\xi))(o = R \mid W, S_i = s) q_{\theta_i(\xi)}(s \mid W) u_{\theta_i(\xi)}(R, W) G(dW) \frac{\sum_{\mathcal{S}} 1\{W \geq c\} q_{\theta_i(\xi)}(s \mid W) u_{\theta_i(\xi)}(R, W) G(dW)}{\sum_{\mathcal{S}} P^n(\alpha(\xi))(o = R \mid W, S_i = s) q_{\theta_i(\xi)}(s \mid W) G(dW)}
\]

\[
= E\left(u_{\theta_i(\xi)}(R, W) \mid W \geq c, S_i = s\right), \tag{18}
\]

where convergence is uniform in \( i \leq n \). The first and fourth lines in (18) follow by definition, the second line follows from the dominated convergence theorem and the fact that \( u_{\theta_i} \) is bounded (and the denominator being greater than zero, as established next), and the third line follows from (17) and the fact that \( G \) is absolutely continuous, so we can ignore the case \( \{W = c\} \) (also, note the importance of \( c < 1 \) for the denominator to be well-defined). A similar argument holds for \( E_{P^n(\alpha(\xi))}(u_{\theta_i(\xi)}(L, W) \mid o = L, S_i = s) \), thus establishing the lemma. \( \square \)

**Proof of Theorem 4.** For this proof, define

\[
\tilde{\kappa}^n(\omega) \equiv \sum_{\theta \in \Theta} \phi(\theta) \sum_{s \in \mathcal{S}} q_{\theta}(s \mid \omega) F^n(\nu_{\theta}(s; \omega)).
\]

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

\(^{33}\)Formally, suppose that \( \omega < c \). Then for all \( \varepsilon > 0 \) there exists \( n_{\xi,\omega,\varepsilon} \) such that, for all \( n \geq n_{\xi,\omega,\varepsilon} \),

\[
P^n(\alpha(\xi))(o = R \mid \omega, S_i = s) q_{\theta_i(\xi)}(s \mid \omega) \leq \varepsilon \quad \text{for all } i \leq n \text{ and } s \in \mathcal{S}.
\]
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 \((\kappa(\cdot; \sigma) \text{ is increasing}):\) Rewrite \( \bar{\kappa}^n \) as

\[
\bar{\kappa}^n(\omega) = \sum_{\theta \in \Theta} \phi(\theta) \left\{ \sum_{s: c_\theta(s) < c} q_\theta(s \mid \omega) \Phi_\theta^n(v_\theta(s; \omega)) + \sum_{s: c_\theta(s) = c} q_\theta(s \mid \omega) F_\theta^n(v_\theta(s; \omega)) \right\} + \sum_{s: c_\theta(s) > c} q_\theta(s \mid \omega) F_\theta^n(v_\theta(s; \omega)) \equiv T_1^n(\omega) + T_2^n(\omega) + T_3^n(\omega).
\]

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 \( \{F^n\}_\eta \) is vanishing, \( \lim_{\eta \to 0} T_2^n(\omega) + T_3^n(\omega) = 0 \) for all \( \omega < c \) and \( \lim_{\eta \to 0} T_1^n(\omega) + T_2^n(\omega) = \sum_{\theta \in \Theta} \phi(\theta) q_\theta(c_\theta(S_\theta) \leq c \mid \omega) \geq \kappa(\omega; \sigma) \) for all \( \omega > c \). In addition, \( T_3^n(\omega) \leq \kappa(\omega; \sigma) \) and \( T_3^n(\omega) \geq 0 \) for all \( \omega \). Therefore, \( \lim_{\eta \to 0} \bar{\kappa}^n(\omega) \leq \kappa(\omega; \sigma) < \kappa(c; \sigma) = \rho \) for all \( \omega < c \) and \( \lim_{\eta \to 0} \bar{\kappa}^n(\omega) \geq \kappa(\omega; \sigma) > \kappa(c; \sigma) = \rho \) for all \( \omega > c \). Consequently, by continuity of \( \kappa^n(\cdot) \), there exists \( (c^n)_\eta \) such that \( c^n \to c \in (-1, 1) \) and \( \bar{\kappa}^n(c^n) = \rho \) for all sufficiently small \( \eta \). By letting \( \sigma^n_\theta(s) = F_\theta^n(v_\theta(s; c^n)) \) for all \( \theta, s \), it follows that \( \kappa(c^n; \sigma^n) = \bar{\kappa}^n(c^n) = \rho \) for all sufficiently small \( \eta \) and, by Theorem 2, that \( \sigma^n \) is a limit equilibrium and \( c^n \) its corresponding cutoff for all sufficiently small \( \eta \). Finally, it remains to establish that \( \sigma^n \to \sigma \). 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 \( c^n \to c \), it follows that \( v_\theta(s; c^n) > 0 \) for all sufficiently small \( \eta \) and, therefore, because \( \{F^n\}_\eta \) is vanishing, that \( \lim_{\eta \to 0} \sigma^n_\theta(s) = 1 = \sigma_\theta(s) \), where the last equality follows since \( \sigma \) is optimal given \( c \)—see equation (12). A similar argument establishes that \( \lim_{\eta \to 0} \sigma^n_\theta(s) = 0 = \sigma_\theta(s) \) 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 \( \{\tilde{F}_\eta\}_\eta \) with the property that \( \lim_{\eta \to 0} \tilde{F}_\eta^n(v_\theta(s; c^n)) = \sigma_\theta(s) \) 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 \((\kappa(\omega; \sigma) = \rho \text{ for all } \omega)\): Without loss of generality, suppose that \( S_\theta \subset \)
for all $\theta$. Let $\mathcal{T}_L = \{(\theta, s) : v_\theta(s; c) < 0 \text{ or } (v_\theta(s; c) = 0 \& \sigma_\theta(s) = 0)\}$, $\mathcal{T}_R = \{(\theta, s) : v_\theta(s; c) > 0 \text{ or } (v_\theta(s; c) = 0 \& \sigma_\theta(s) = 1)\}$, and $\mathcal{T}_0 = \{(\theta, s) : v_\theta(s; c) = 0 \& \sigma_\theta(s) \in (0, 1)\}$. Note that, since $(\sigma, c)$ is a voting equilibrium, then $\sigma_\theta(s) = 0$ if $(\theta, s) \in \mathcal{T}_L$ and $\sigma_\theta(s) = 1$ if $(\theta, s) \in \mathcal{T}_R$. Define $X_L \equiv \sum_{(\theta, s) \in \mathcal{T}_L} \phi(\theta)q(s \mid c)s \geq 0$, $X_R \equiv \sum_{(\theta, s) \in \mathcal{T}_R} \phi(\theta)q(s \mid c)s \geq 0$, and $X_0 \equiv \sum_{(\theta, s) \in \mathcal{T}_0} \phi(\theta)q(s \mid c) \geq 0$. The proof constructs a specific family of perturbations. For all $\eta$ and all $\theta \in \Theta$ and $s \in S_\theta$ let

$$F_\theta^n(v_\theta(s; c)) = \begin{cases} 
\zeta_L s\eta & \text{if } v_\theta(s; c) < 0 \text{ or } (v_\theta(s; c) = 0 \& \sigma_\theta(s) = 0) \\
\sigma_\theta(s) + \zeta_0 \eta & \text{if } \{v_\theta(s; c) = 0 \& \sigma_\theta(s) \in (0, 1)\} \\
1 - \frac{\zeta_R}{s} \eta & \text{if } v_\theta(s; c) > 0 \text{ or } (v_\theta(s; c) = 0 \& \sigma_\theta(s) = 1) 
\end{cases}$$

By construction, for all $\zeta_j \in (0, \infty), j = R, L$ and $\zeta_0 \in [0, \infty)$, and for all $\eta$ sufficiently low, there exists a vanishing family $\{F^n\}_n$ that satisfies the above restrictions; note that, by MLRP, for each $\theta$ there is at most one signal that satisfies $v_\theta(s; c) = 0$. Then, since $c \in (-1, 1)$,

$$\tilde{k}^n(c) - \rho = \tilde{k}^n(c) - k(c; \sigma) = \sum_{(\theta, s)} \phi(\theta)q(s \mid c) (F^n_\theta(v_\theta(s; c)) - \sigma_\theta(s))$$

$$= \eta (-\zeta_R X_R + \zeta_L X_L + \zeta_0 X_0).$$

It is straightforward to check that we can always pick $\zeta_R, \zeta_L, \zeta_0$ such that $-\zeta_R X_R + \zeta_L X_L + \zeta_0 X_0 = 0$ and, therefore, $\tilde{k}^n(c) = \rho$ for all $\eta$ sufficiently small. As in Case 1, by letting $\sigma^n_\theta(s) = F^n_\theta(v_\theta(s; c))$ for all $(\theta, s)$, it follows that $\sigma^n$ is a limit equilibrium and $c$ its corresponding cutoff for all sufficiently small $\eta$. The proof is completed by noting that, by construction, $\lim_{\eta \to 0} \sigma^n = \sigma$. □