Learning foundation and equilibrium selection in voting environments with private information*

Ignacio Esponda (NYU Stern)

Demian Pouzo (UC Berkeley)

January 25, 2012

Abstract

We use a dynamic learning model to investigate different behavioral assumptions in voting environments with private information. We show that a simple rule, where players learn based on the outcomes of past elections in which they were pivotal but requires no prior knowledge of the payoff structure or of the rules followed by other players, provides a foundation for Nash equilibrium. In contrast, a rule where voters learn from *all* past elections provides a foundation for a new notion of naive voting where players vote sincerely but have endogenously-determined beliefs. Finally, we use the model to select among multiple equilibria in the jury model. We find that the well-known result that elections aggregate information under Nash equilibrium relies on the selection of symmetric equilibria which are unstable. Nevertheless, we show that there exist (possibly asymmetric) Nash equilibria that are asymptotically stable and aggregate information.

*We thank Heski Bar-Isaac, Drew Fudenberg, Johannes Hörner, Navin Kartik, Alessandro Lizzeri, Wolfgang Pesendorfer, Emanuel Vespa and several seminar participants for helpful comments. Esponda: 44 W Fourth Street KMC 7-76, New York, NY 10012, iesponda@stern.nyu.edu; Pouzo: 508-1 Evans Hall #3880, Berkeley, CA 94720, dpouzo@econ.berkeley.edu.

1 Introduction

Games with asymmetric information are often analyzed under the assumption of Nash equilibrium, which requires players to play strategies that are a best response to the equilibrium strategies of other players and the true primitives of the game. What justifies the Nash equilibrium assumption? The literature on learning in games (see Fudenberg and Levine (1998, 2009) for surveys) has convincingly argued that equilibrium can be viewed as the steady-state of a dynamic process, where players learn the stage-game strategies of other players from past play.[1] Most of the literature has focused on games with complete information, where these strategies are likely to be observed ex-post. However, in games with asymmetric information, a stage-game strategy is a mapping from private information to actions. To the extent that private information remains private, then such strategies are never observed, raising the question of whether it is reasonable to expect Nash play in such games.

In this paper, we study a dynamic model of learning for one such game of asymmetric information, a voting game, in order to shed light on the plausibility of different solution concepts that have been applied in the voting literature. We focus on a voting game for three reasons. First, beginning with Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1997, 1998), a large literature has studied the important question of information aggregation in elections under the assumption of Nash equilibrium. However, the view that voters are sophisticated enough to play a Nash equilibrium remains controversial.[2] Second, the literature often selects *symmetric* equilibria, but there is no formal justification for this selection. Third, when strategies are not observed ex-post, there are many learning rules one could study depending on the feedback received by players. By focusing on a particular context, we can better judge the plausibility of our rules and relate them to the existing evidence.

To understand the controversy surrounding Nash equilibrium in voting games, note that a player's vote can only make a difference if it is pivotal, i.e., if it decides the outcome of the election; hence, players should vote for the alternative that is best conditional not only on their private information but also on the information that can be inferred from the event that they are pivotal. There are two steps involved

[1] Models where players explicitly form beliefs about the strategies of other players are known as belief-based models (e.g., Fudenberg and Kreps, 1993).

[2] For different views, see Green and Shapiro (1994), Friedman (1996), and Feddersen (2004).

in such sophisticated voting.[3] First, players must realize that they should use the hypothetical information that their vote is pivotal when making decisions. Second, players should be able to make correct inferences from their signal and the event that they are pivotal. A modeler who computes a best response makes these inferences by using information about the primitives of the game and the (equilibrium) strategies of other players.[4] A natural question arises as to how players would come to have correct beliefs about these primitives and the strategies of other players. A common alternative to Nash equilibrium, known as *sincere voting*, postulates that players vote given their private information alone, therefore ignoring others' information. While players who vote sincerely do not need to learn the strategies of other players, a question still remains as to how they can learn the primitives of the game in order to respond appropriately to their private signals.

We make three main contributions in the context of voting games with private information. First, we show that there is a simple learning rule that justifies the notion of Nash equilibrium. Second, we show that there is another learning rule, perhaps more natural than the first, that justifies an alternative notion of equilibrium (Esponda, 2008) and captures a new notion of naive voting where players vote sincerely but have endogenously determined beliefs. Finally, we use the dynamic model to select among the multiple equilibria that often exist in voting games.

Instead of postulating that players try to learn the strategies of other players and the primitives of the game, we postulate that players simply try to learn how their private information correlates with the benefits observed from the election of the different alternatives.[5] Our assumption that voters learn from the past is motivated by the literature on retrospective voting (Key (1966), Kramer (1971), Fiorina (1978)). That literature finds that voters' propensity to vote for a candidate depends

[3]Esponda and Vespa (2011) emphasize these two steps and show that about half of their experimental subjects are unsophisticated and fail even at the former step.

[4]To illustrate, consider the case of majority voting, where the pivotal event is that the votes are evenly split between the two alternatives. Suppose, for concreteness, that there are two signals and everyone votes for one alternative under the low signal and for the other alternative under the high signal. Then a player should infer that half of the signals are low and half of the signals are high. Together with her own signal and knowledge of the primitives, the player should then compute the posterior distribution over the state of nature, and take expectations to determine the best alternative to vote for.

[5]As discussed in Section 7, our approach is also different from reinforcement-based learning models (e.g., Erev and Roth, 1998), where a player chooses future *strategies* based on their past performance.

on the candidate's (or the party's) past performance during office. A natural question is whether such retrospective voting leads to any of the familiar solution concepts applied by the voting literature.

We analyze a setup where a standard voting stage game is repeated every period. Voters simultaneously decide which of two alternatives to support. The best alternative depends on the state of the world, and votes are cast after observing private signals that are correlated with the state. The outcome of the election is determined by a particular voting rule (for example, majority voting).

We study two plausible learning rules. According to our first rule, voters learn the desirability of an alternative conditional on their private information by observing the payoff in previous elections where that alternative was elected. For example, consider a voter that is only uncertain about the payoff of electing a Democratic candidate. She will form a belief about this payoff by evaluating the performance of similar Democratic candidates that were elected in previous elections. And she will then vote for the candidate that she sincerely believes to provide the highest expected payoff, given the history so far. This learning rule does not account for the fact that the sample of *elected* candidates from which she learns is biased: to the extent that other voters use their private information to make decisions, then she learns from the "good" Democrats that were elected into office but not from the "bad" Democrats that were not elected and whose performance in office was, therefore, never observed. We can think of this rule, which fails to account for sample selection, as the analogue of failing to account for the informational content of other players' actions in static settings.[6],[7] For this reason, we refer to this rule as *non-strategic*.

The second rule that we analyze is strategic in the sense that voters try to account for the information content of other voters' strategies. As before, voters update their beliefs using past feedback, except that now voters use feedback *only* from those elections in which their vote was pivotal.

Our first main result (Theorem 2) links the steady-states of the dynamic voting environment under each type of voting rule to solution concepts of the stage game. Consider first the strategic voting rule. We show that an asymptotic steady-state of

[6]See Eyster and Rabin (2005), Jehiel and Koessler (2008), and Esponda (2008) for equilibrium concepts that capture this mistake, and Aragones et al. (2005), Al-Najjar (2009), Al-Najjar and Pai (2009), and Schwartzstein (2009) for theoretical foundations of related forms of naivete.

[7]See Kagel and Levin (2002) and Charness and Levin (2009) for supporting experimental evidence of the failure to account for the information content of others' strategies in auction-like contexts.

the dynamic environment with a strategic learning rule constitutes a Nash equilibrium of the stage game; further, any Nash equilibrium is an asymptotic steady-state for some strategic rule that meets our assumptions. Consequently, we provide a learning foundation for *all* Nash equilibria of the voting game. In particular, a simple rule that requires players to update their beliefs only on those periods when their votes were pivotal can justify Nash equilibrium. This is true despite the assumption that players cannot observe counterfactual payoffs, and, therefore, cannot directly compare the benefits from each alternative in the same period. In Section 7, we discuss natural extensions of the strategic rule that may be more appropriate whenever the probability of being pivotal is small relative to the frequency of elections.

Next, consider the non-strategic voting rule, where players learn from all periods and do not attempt to correct for potential sample selection bias. We establish an analogous result, but the corresponding solution concept turns out not to be Nash equilibrium, but rather the (naive) behavioral equilibrium introduced by Esponda (2008), which we refer to as a *non-strategic equilibrium*: An asymptotic steady-state of the dynamic environment with a non-strategic learning rule constitutes a non-strategic equilibrium of the stage game; further, any non-strategic equilibrium is an asymptotic steady-state for some non-strategic rule that meets our assumptions.

In particular, an examination of the voting game from a learning perspective leads to a new notion of naive voting that had previously not been applied to voting games. This solution concept is distinct from the notion of Nash equilibrium, therefore suggesting that it may be worthwhile to re-evaluate some of the previous conclusions in the literature under this alternative solution concept. In addition, a non-strategic equilibrium differs from the standard approach to naivete, which is to simply postulate that players follow "sincere" strategies based on their (often unjustified) knowledge of the primitives and unresponsive to important aspects of the game, such as the voting rule or the number of other players. The reason for this difference is that counterfactual payoffs are not observed: Sincere voting would be the outcome if players also observed the hypothetical performance of the losing alternative. Under the new notion of naivete embodied by non-strategic equilibrium, it is also the case that players vote for the alternative that they consider best given their private information alone, thus ignoring others' votes. However, the meaning of players' private information arises endogenously in equilibrium. The reason is that a belief about the payoff from an alternative depends on the event that the alternative is chosen, which is in turn determined by everyone's strategy.[8]

The result in Theorem 2 that, under our rules, attention can be restricted either to Nash or non-strategic equilibrium is uncontroversial. However, the result that *all* such equilibria can be approached as steady-states relies on the construction of a particular rule that requires players to (somehow) have equilibrium beliefs and do not deviate from the corresponding best response unless it is sufficiently beneficial to do so (with the required benefit going to zero as time goes to infinity). In addition, we cannot use this result to select among the multiple equilibria that exist in voting games. Therefore, we next restrict attention to rules that require players to exactly best respond every period. Our second main result, Theorem 3, provides conditions for equilibria to be asymptotically stable (in the sense that the stochastic dynamics converge to it with positive probability) and unstable (i.e., zero-probability convergence).

Finally, we apply Theorem 3 to select among multiple equilibria in a standard voting game with two states of the world.[9] The Nash equilibrium literature restricts attention to symmetric equilibria, and these equilibria are often in mixed-strategies. We establish (Proposition 1) that mixed-strategy Nash equilibria are unstable. To the extent that the flexibility in rules assumed by Theorem 2 may be controversial, this result casts some doubts on the results obtained in the literature (such as information aggregation). However, in the context of the jury model with binary signals, we show (Proposition 2) that there exist (not necessarily symmetric) pure-strategy equilibria that are asymptotically stable. Further, these equilibria do aggregate information, in the sense that the difference between equilibrium and first-best welfare vanishes as the number of players goes to infinity (Proposition 3).

The formal analysis adapts the stochastic fictitious play model which Fudenberg and Kreps (1993) introduced to provide a learning foundation for mixed-strategy equilibrium in strategic-form games. In that model, stage-game strategies are fully observable and, therefore, players learn the strategies being followed by other players.

[8]Esponda and Pouzo (2012) further pursue the comparison between Nash and non-strategic equilibrium. They show that these solution concepts also differ in the limiting case where the number of players goes to infinity and, in particular, information is not necessarily aggregated under non-strategic equilibrium.

[9]When discussing multiplicity, we are not referring to *non-responsive* equilibria where, say, everyone is indifferent to vote for the same alternative because they cannot affect the outcome of the election. Indeed, we follow the literature and focus on *responsive* equilibria. However, there is often multiplicity of responsive equilibria, and the goal is to use our dynamics to select among them.

Fudenberg and Kreps (1994, 1995) extended their setup to extensive-form games, where players observe only the path of play, thus providing a foundation for self-confirming equilibrium (see Section 7 for the relationship between our Nash foundation and the notion of a self-confirming equilibrium). Our analysis (in particular, Theorem 2) adapts the notions of empirical beliefs, asymptotic empirism, and stability of Fudenberg and Kreps (1993, 1994, 1995) to a particular game with private information. A few differences are worth noting. In our case, players no longer attempt to learn the (information-contingent) strategies of other players, since private information of other players is never observed, but instead learn the expected payoffs for each of the two voting alternatives. As a consequence, our players must learn from an endogenously biased sample and whether players take selection into account or not becomes important.

In addition, Fudenberg and Kreps (1993) follow Harsanyi (1973) in introducing independent payoff perturbations in their model. An important role of these perturbations is to make sure that steady-state strategies are independent, as in standard (e.g., Nash) solution concepts. In our context, players also observe private information, and one might conjecture, following the purification literature (see Morris (2008) for a brief review), that a sufficiently rich set of informative signals (e.g., a continuum of signals) would play a similar role. While this is true for a Nash equilibrium, it is no longer true for a non-strategic equilibrium, where beliefs are endogenously affected by one's own strategy and, consequently, if a (perceived) best response is a mixed strategy, then it is not necessarily true that the pure strategies in the support are also best responses. Hence, we also rely on payoff perturbations, uninformative about the state of the world, in order to eliminate potential correlation in steady-state strategies. This point is likely to be relevant in establishing learning foundations for other games with private information where players have biased beliefs.

The payoff perturbations also play two new roles in our setup. First, they guarantee that both alternatives are perpetually chosen in a steady-state. Therefore, in contrast to a self-confirming equilibrium (e.g., Fudenberg and Kreps, 1995), incorrect beliefs under the non-strategic rule are not due to the lack of experimentation, but rather to failure to account for a biased sample. Second, the perturbations explain why the strategic learning rule leads to a Nash equilibrium despite the fact that players never observe the difference in payoffs from both alternatives in a same time period.

Our results on the selection of equilibria (Theorem 3) rely on the literature on stochastic approximation, which provides conditions for a discrete stochastic system to be asymptotically well-approximated by solutions of a system of ordinary differential equations. These techniques have been applied to select equilibria both in microeconomics (e.g., Benaim and Hirsch, 1999) and macroeconomics (e.g., Marcet and Sargent, 1989).

A few papers also study other solution concepts in voting games. Osborne and Rubinstein (2003) propose the steady-state concept of sampling equilibrium for elections between more than two candidates but without private information. The concept informally captures a learning environment where voters use the observed actions of a small number of other voters to infer the distribution of actions in the entire population. In the context of common value elections studied in our paper, two other solution concepts are Eyster and Rabin's (2005) notion of (partially) cursed equilibrium—which captures a convex combination of the sincere and Nash approaches—and Costinot and Kartik's (2007) application of level-k thinking (Stahl and Wilson (1995), Nagel (1995))—which focuses on introspection rather than learning.[10]

In Section 2, we illustrate the non-strategic and strategic learning rules. In Section 3, we present the voting stage game and the notions of Nash and naive (or non-strategic) equilibrium. In Section 4, we provide learning foundations for these two notions of equilibrium. We study selection of equilibrium in Section 5 and apply these results in Section 6. We conclude in Section 7 by discussing additional related topics. Some tedious proofs are collected in the Online Appendix.[11]

2 Illustration of learning rules

A group of n players plays an infinitely repeated voting game. Each period, a new state is drawn and players choose between alternatives A and B. For example, consider the payoffs in Figure 1, where A is best in state ω_A and B is best in state ω_B. Before casting their vote, players observe private signals $s \in \{a, b\}$ that are in-

[10] There is an alternative literature on learning and experimentation by multiple agents (Bolton and Harris (1999), Keller, Rady and Cripps (2005); Strulovici (2010) in a voting context). That literature studies learning in an equilibrium context, while we study learning as a justification for equilibrium.

[11] Available at http://people.stern.nyu.edu/iesponda/Ignacio_Esponda/Research.html.

	ω^A	ω^B
A	1	-1
B	0	0

Figure 1: Payoffs

dependently drawn, conditional on the state. After observing their signals, players simultaneously cast their vote for one of the two alternatives. The outcome of the election is A if and only if the proportion of votes in favor of A is higher than some threshold.

Figure 2 shows a history of past outcomes observed by a particular player after playing the game for 8 periods. The private history includes her signal, her vote, whether her vote was pivotal, the outcome of the election, and her payoff in each period. Suppose that in period 9, the player observes signal a. The behavior that we postulate differs for non-strategic and strategic players.

Suppose that the player is non-strategic. Then she forms beliefs about the expected benefit of outcome A by computing the average *observed* payoff obtained from A when her signal was a, which in this case is $(-1+1+1)/3 = 1/3$.[12] Second, the player votes for the alternative that she believes has the highest expected payoffs: in this case $1/3 > 0$ and, therefore, she votes for A.

Non-strategic players do not take into account two sources of sample selection. The first source is exogenous: Estimates are likely to be biased upwards if alternatives tend to be chosen when they are most likely to be successful—which is to be expected if players use their private information to make decisions. In Figure 2, counterfactual payoffs for A are not observed conditional on signal a in periods 3 and 7, but the fact that A was not chosen makes it likely that counterfactual payoffs would have been lower, on average, than observed payoffs for A. The second source is endogenous: A player's vote affects the sample that she will observe (in a period when she is

[12] Our results hold as long as players consistently estimate observed mean payoffs; in particular, players could have a prior and apply Bayesian updating based on observed payoff outcomes.

time	signal	vote	piv?	election outcome	observed payoff	
1	**a**	**A**	**y**	**A**	**-1**	
2	**a**	**B**	**n**	**A**	**1**	
3	a	A	n	B	0	→ counterfactual not observed
4	b	B	n	A	-1	
5	b	B	y	B	0	
6	**a**	**A**	**n**	**A**	**1**	
7	a	A	n	B	0	→ counterfactual not observed
8	b	B	n	A	1	

Figure 2: History of play after 8 periods

pivotal), and, consequently, her beliefs and behavior. In both the exogenous and endogenous cases, the underlying source of the bias is that other players use their private information to make decisions. Failing to account for selection in a learning environment is, then, analogous to failing to account for the informational content of other players' actions.

Suppose now that the player is strategic and realizes that her vote only matters in the event that it decides the election. Then she forms beliefs about the expected benefit of outcome A by computing the average *observed* payoff obtained from A when the observed signal was a and when her vote was pivotal. In particular, the belief about the benefit of voting A of a strategic player who observes signal a in period 9 is now $-1 < 0$ (because only the first period counts); hence, a strategic player votes for B.

3 Equilibrium model of voting

3.1 Voting stage game

A committee or electorate of $n \geq 3$ players must choose between two alternatives, A and B. A state $\omega \in \Omega$ is first drawn from a finite set according to a full-support probability distribution $p \in \Delta(\Omega)$ and then each player $i = 1, ..., n$ privately observes a signal s_i. Each signal s_i is drawn independently from a finite set S_i, and, conditional on the realized state ω, with probability $q_i(s_i \mid \omega)$; let $s = (s_1, ..., s_n) \in S \equiv \Pi_{i=1}^n S_i$ denote a profile of signals. After observing their signals, players simultaneously cast a vote $x_i \in X_i = \{0, 1\}$ for either alternative, where $x_i = 1$ denotes a vote for A and $x_i = 0$ denotes a vote for B; let $x = (x_1, ..., x_n) \in X \equiv \Pi_{i=1}^n X_i$ denote a profile of votes. Votes are aggregated according to a threshold voting rule: The election outcome $o(x)$ is A if and only if $k \geq 1$ or more players vote for A; otherwise, $o(x) = B$.

Player i's payoff is given by

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

where $v_i \in V_i \subset \mathbb{R}$ is a privately-observed payoff perturbation à la Harsanyi (1973) and $o \in \{A, B\}$ is the election outcome. Utility is bounded: $|u_i(o, \omega)| < K < \infty$ for all $i = 1, ..., n$, $o \in \{A, B\}$, and $\omega \in \Omega$. Moreover, the perturbation v_i is independently drawn from an absolutely continuous distribution F_i that satisfies $F_i(-2K) > 0$ and $F_i(2K) < 1$.

For strategic players (i.e., Nash equilibrium), the event that a player's vote is pivotal plays an important role. We denote the event that player i's vote can change the outcome of the election by

$$piv_i^{NE} = \left\{ x_{-i} \in \Pi_{j \neq i} X_j : \sum_{j \neq i} x_j = k - 1 \right\}.$$

On the other hand, non-strategic voters do not condition on particular profiles of other player's votes. In order to simplify notation and facilitate comparison of strategic and non-strategic behavior, we let $piv_i^N \equiv \Pi_{j \neq i} X_j$ denote the event on which non-strategic players condition, which is the non-informative event consisting of all profiles of others' votes.

Let $\mathrm{Y}_i = \{0, 1\}^{S_i}$ be the set of signal-contingent actions of player i. An *action* *plan* for player i is a function $\phi_i : V_i \to Y_i$ that describes player i's signal-contingent action as a function of her realized payoff perturbation.[13] We restrict attention to weakly undominated strategies, so that every action plan satisfies: $\phi_i(v_i)(s_i) = 0$ for $v_i < -2K$ and $\phi_i(v_i)(s_i) = 1$ for $v_i > 2K$, for all $i = 1, ..., n$ and $s_i \in S_i$. For each action plan ϕ_i, there is an associated (mixed) *strategy* $\alpha_i \in \mathcal{A}_i$, where

$$\mathcal{A}_i = \left\{\alpha_i \in \mathbb{R}^{\# S_i} : F_i(-2K) \leq \alpha_i(s_i) \leq F_i(2K)\ \forall s_i \in S_i\right\}$$

is the set of player i's strategies, and

$$\alpha_i(s_i) = \Pr\left(\{v_i : \phi_i(v_i)(s_i) = 1\}\right)$$

is the probability that player i votes for A after observing signal s_i. Each strategy profile $\alpha = (\alpha_1, ..., \alpha_n) \in \mathcal{A} \equiv \Pi_{i=1}^n \mathcal{A}_i$ induces a distribution over outcomes, $P(\alpha) \in \Delta(Z)$, where $Z \equiv X \times S \times \Omega$ is the set of relevant outcomes from players' point of view, and

$$P(\alpha)(x, s, \omega) = p(\omega) \prod_{i=1}^{n} \alpha_i(s_i)^{x_i} \left(1 - \alpha_i(s_i)\right)^{1-x_i} q_i(s_i \mid \omega). \tag{1}$$

Whenever an expectation E_P has a subscript P, this means that the probabilities are taken with respect to the distribution P.

3.2 The role of payoff perturbations

The independent payoff perturbations play two important roles in this paper. The first is as a refinement criterion in the spirit of Selten (1975) and, in the context of voting, formalizes the standard restriction to *responsive* strategy profiles (see Section 6). This role is described by the following result; the proof is straightforward and, therefore, omitted.

Lemma 1. *Let $\alpha \in \mathcal{A}$. Then, for all i and $s_i \in S_i$:*

(i) $P(\alpha_i', \alpha_{-i})(o = A \mid s_i) - P(\alpha_i'', \alpha_{-i})(o = A \mid s_i) > 0$ *for all* $\alpha_i'(s_i) > \alpha_i''(s_i)$.

(ii) $P(\alpha)(o = A \mid s_i) \in (0, 1)$.

[13] The restriction to *pure* action plans is justified because F is absolutely continuous (Harsanyi, 1973).

Part (i) says that player i's vote affects the outcome of the election; i.e., the probability of being pivotal is strictly greater than zero. Without this result, existence of a Nash equilibrium would be trivial—e.g., everyone voting for the same alternative is an equilibrium if any voting rule other than unanimity is used. Part (ii) says that alternatives A and B are chosen with positive probability. If one of the alternatives were never chosen, then beliefs about the payoffs from that alternative would be arbitrary in our learning environment, hence justifying the decision not to choose the alternative in the first place. The perturbations provide the experimentation necessary to pin down beliefs in the steady-state.

A second role of payoff perturbations is to guarantee that, if behavior and beliefs stabilize in the dynamic version of the voting game, then players' steady-state voting strategies are independent.[14] This result is important for two reasons. First, if steady-state strategies were not independent, then we would need to focus on the (more permissive) notion of a correlated equilibrium. Second, the fact that strategies are asymptotically independent explains how a learning rule that updates beliefs only on periods when a vote is pivotal can lead to Nash equilibrium despite the fact that the payoff difference between the two alternatives cannot be observed in any period.

3.3 Definition of non-strategic and Nash equilibrium

A naive (or, more generally, behavioral) equilibrium (Esponda, 2008) combines the idea of a self-confirming equilibrium (Battigalli (1987), Fudenberg and Levine (1993), Dekel, Fudenberg, and Levine (2004)) with an information-processing bias. Players choose strategies that are optimal, given their beliefs about the consequences of following each possible strategy. In contrast to Nash equilibrium, these beliefs are not necessarily restricted to being correct, but, rather, to being consistent with the information feedback players receive. This information is, in turn, endogenously generated by the equilibrium strategies followed by all players. Our feedback assumption is that players observe only the realized payoff of the alternative that the committee chooses, but not the counterfactual payoff of the other alternative.[15] A naive equilibrium requires beliefs to be *naive-consistent*, meaning that information is not

[14] Harsanyi (1973) introduced independent payoff perturbations to justify mixed-strategy equilibrium in a static context. Fudenberg and Kreps (1993) applied these perturbations to justify mixed-strategy equilibrium in a learning context.

[15] The assumption that counterfactuals are not observed guarantees that players' naive model of the world is consistent with their feedback (see Esponda (2008) for further discussion).

correctly processed by players. In particular, players do not take into account that other players' actions may be correlated with the true state of nature. In this way, we formalize the idea that players do not take into account the informational content of other players' actions, or, equivalently, the sample selection problem. In Section 7, we argue that naivete may be sensible even for sophisticated players that understand selection. Hence, given that the defining characteristic of naive players is that they do not make inferences from the strategies of other players, we refer to a naive equilibrium as a non-strategic equilibrium.

We now define two solution concepts, one for each value of $m \in \{NE, N\}$, corresponding to Nash and non-strategic equilibrium, respectively. Throughout the paper, results that are stated to hold for m should be understood to hold both for the case $m = NE$ and the case $m = N$. To gain some intuition for the solution concepts, suppose that player i repeatedly faces a sequence of stage games where players use strategies α every period and that she wants to learn the value of voting for A and B. If player i learns the value of an alternative by conditioning on her own signal, on the event that the alternative is elected (hence, observed), and on the event that her vote is m-pivotal, then her beliefs about the expected difference between A and B are given by

$$\Delta_i^m(P(\alpha), s_i) \equiv E_{P(\alpha)}\left[u_i(A,\omega) \mid o = A, piv_i^m, s_i\right] - E_{P(\alpha)}\left[u_i(B,\omega) \mid o = B, piv_i^m, s_i\right],$$

which is well-defined by Lemma 1. Player i then votes for A whenever $\Delta_i^m(P(\alpha), s_i) \geq v_i$, where v_i is her realized payoff perturbation, therefore motivating the following definition.

Definition 1. A strategy profile $\alpha \in \mathcal{A}$ is an *m-equilibrium* of the stage game if for every player $i = 1, ..., n$ and for every $s_i \in S_i$,

$$\alpha_i(s_i) = F_i\left(\Delta_i^m(P(\alpha), s_i)\right). \tag{2}$$

We refer to $P(\alpha) \in \Delta(Z)$ as an m-equilibrium distribution.

In an m-equilibrium, players best respond to beliefs that are endogenously determined by both their own strategy and those of other players. The case $m = NE$ corresponds to the definition of a Nash equilibrium. To see this claim, note that the event $\{o = A, piv_i^{NE}, s_i\}$ is identical to the event $\{x_i = 1, piv_i^{NE}, s_i\}$ because player i's vote determines the outcome when she is pivotal. Moreover, conditioning on the latter event is equivalent to conditioning on the event $\{piv_i^{NE}, s_i\}$, since the probability that i votes for A or B depends only on the payoff perturbation and is, therefore, uncorrelated with the state of the world, once we condition for the observed signal s_i. A similar argument is valid for the event $\{o = B, piv_i^{NE}, s_i\}$. Therefore,

$$\Delta_i^{NE}(P, s_i) = E_{P(\alpha)}\left[u_i(A,\omega) - u_i(B,\omega) \mid piv_i^{NE}, s_i\right],$$

which is the standard criterion for making choices in a Nash equilibrium, where players choose the best alternative conditional on the event that their vote is pivotal.

The case $m = N$ corresponds to a naive (or non-strategic) equilibrium (Esponda, 2008). In this case, players' beliefs are consistent with observed equilibrium outcomes but players do not account for the correlation between others' votes and the state of the world (conditional on their own private information). In particular, naive or non-strategic players fail to account for the selection problem.

Theorem 1. *There exists an m-equilibrium of the stage game.*

Proof. Let $\Phi^m : \mathcal{A} \to \mathcal{A}$ be given by $\Phi^m(\alpha) = \left(F_i\left(\Delta_i^m(P(\alpha), s_i)\right)_{s_i \in S_i}\right)_{i=1,\ldots,n}$. First, note that $\Phi^m(\alpha) \in \mathcal{A}$ for all $\alpha \in \mathcal{A}$. Second, $\mathcal{A}$ is a convex and compact subset of a Euclidean space. Third, $P(\cdot)$ is continuous, implying that $\Delta_i^m(P(\cdot), s_i)$ (which is well-defined by Lemma 1) is continuous and, by continuity of F_i, that Φ^m is continuous. Therefore, by Brouwer's fixed-point theorem, there exists a fixed point of Φ^m, which is also an m-equilibrium of the stage game. □

The definition of equilibrium was motivated by a learning story. In Section 4 we study an explicit model of learning and provide a foundation for both Nash ($m = NE$) and non-strategic ($m = N$) equilibrium. In addition, while existence of equilibrium is guaranteed, there may exist multiple equilibria. This multiplicity issue is well-known for Nash equilibrium, where applications often focus on symmetric equilibria and ignore asymmetric equilibria. In Sections 5 and 6, we apply the explicit learning model to select among the set of equilibria.

4 Learning foundation for equilibrium

We present a dynamic framework in order to clarify and justify the Nash and non-strategic equilibrium concepts for the voting game with private information. A dynamic game is a repetition of the voting stage game in which the state, the signals, and the payoff perturbations are drawn independently across time periods from the same distribution. We postulate two learning rules and adapt the notion of empirical beliefs, asymptotic myopia, and stability of Fudenberg and Kreps (1993, 1994, 1995) to show that the steady-states under these rules correspond to equilibria of the voting stage game.

4.1 A model of learning

A group of n players play the stage game described in Section 3 for each discrete time period $t = 1, 2, ...$ At time t, the state is denoted by $\omega_t \in \Omega$, the signals by $s_t = (s_{1t}, ..., s_{nt})$, and the votes by $x_t = (x_{1t}, ..., x_{nt})$. The outcome of the election at time t is determined by a threshold voting rule k and denoted by $o_t \in \{A, B\}$. Player i's utility is

$$u_i(o_t, \omega_t) + 1\{o_t = B\}\, v_{it},$$

where v_{it} is the payoff perturbation drawn independently (across players and time) from F_i. As before, let piv_i^m denote the profiles of others' votes such that player i's vote is m-pivotal, where $m \in \{NE, N\}$.

Let $h^t = (z_1, ..., z_{t-1})$ denote the history of the game up to time $t-1$, where $z_t = (x_t, s_t, \omega_t) \in Z$ is the time-t outcome. Let $\mathcal{H}^t$ denote the set of all time-t histories and let $\mathcal{H}$ be the set of infinite histories. At each round of play, players privately collect feedback about past outcomes. For each player i, $Z_i \equiv X \times S_i \times \mathrm{U}_i$ is the set of outcomes that player i may observe at any given period, where $U_i = U_{Ai} \cup U_{Bi}$ is the union of the ranges of her utility functions, i.e., $U_{oi} = u_i(o, \Omega)$. Let $h_i^t = (z_{i1}, ..., z_{it-1})$ denote player i's private history up to time $t-1$, where $z_{it} = (x_t, s_{it}, u_i(o_t, \omega_t)) \in Z_i$ is the privately observed outcome at time t. Note that payoff perturbations are not part of the history, implicitly assuming that players understand that the perturbations are independent payoff shocks that are unrelated to the learning problem.

We complete the specification of the dynamic game by introducing assessment (i.e., belief-updating) and policy rules. An *assessment rule* for player i is a sequence

$\mu_i = (\mu_{i1}, ..., \mu_{it}, ...)$ such that $\mu_{it} : \mathcal{H} \to \mathbb{R}^{\# S_i}$ is measurable with respect to the player i's time-t private history. The interpretation is that the s_i-coordinate of $\mu_{it}(h)$, $\mu_{it}(h)(s_i)$, is player i's beliefs—given her private $t-1$-period history in h—about the *difference* in expected utility between alternatives A and B conditional on her signal s_i and on the m-pivotal event.

A *policy rule* for player i is a history-dependent sequence of action plans $\phi_i^H = \left(\phi_{i1}^H, ..., \phi_{it}^H, ...\right)$, where $\phi_{it}^H : \mathcal{H} \times V_i \to Y_i \equiv \{0,1\}^{S_i}$ is measurable with respect to player i's time-t private history and her time-t payoff perturbation. The interpretation is that $\phi_{it}^H(h, v_i)(s_i)$ is player i's vote at time t, conditional on observing perturbation v_i, signal s_i, and private history h_i^t (as a consequence of history h).

The measurability restrictions on assessment and policy rules imply that players' decisions may depend on the observed payoff outcomes but not on the (unobserved) state of the world, thus capturing the assumption that players do not observe counterfactual payoffs. In particular, players do not know the relationship between the state space and their payoffs; otherwise, under generic payoffs, they could infer the payoff they would have received from a losing alternative by simply observing the payoff received from the elected alternative.[16]

Given a *policy rule profile* $\phi^H = \left(\phi_1^H, ..., \phi_n^H\right)$, let $\mathbf{P}_t^{\phi^H}(\cdot \mid h^t)$ denote the probability distribution over histories, conditional on history up to time t, $h^t \in \mathcal{H}^t$, and let $\mathbf{P}^{\phi^H}$ denote the unconditional probability distribution over histories, which we can construct by Kolmogorov's extension theorem.

4.2 The selection problem: non-strategic and strategic assessments

We study a family of learning rules that are indexed by m and differ in how players form beliefs about the best alternative to vote for. Let

$$Z_{\text{oui}}^m(s_i) = \left\{(x', s', \omega') \in Z : o(x') = o, \omega' \in u_i^{-1}(o, \mathrm{u}), x'_{-i} \in piv_i^m, s'_i = s_i\right\}$$

[16] For example, suppose that, in Figure 1, the payoff of B in state ω_A were changed from 0 to $1/2$. Then, a player who obtains $1/2$ from the election of B and knows the payoff-state structure would also learn the counterfactual payoff of 1. Alternatively, players in this modified example can believe (incorrectly) in a payoff-state structure that has four, rather than two, states. Also, additional "structural" knowledge about states and payoffs could be incorporated, as long as voters cannot perfectly infer counterfactual payoffs (in which case, as mentioned in the introduction, non-strategic learning yields sincere voting).

denote the event that player i is m-pivotal, observes s_i, and obtains payoff $\mathtt{u}$ from the elected outcome o. For a history h, let[17]

$$\xi^m_{\mathtt{u}oit}(h)(s_i) = \frac{1}{t-1}\sum_{\tau=1}^{t-1} \mathbf{1}_{Z^m_{\mathtt{u}oi}(s_i)}(z_\tau) \tag{3}$$

be the proportion of times, up to $t-1$, that player i has observed the event $Z^m_{\mathtt{u}oi}(s_i)$, where z_τ are the time-τ elements of h.

Definition 2. An assessment rule μ^m_i is *m-empirical* if

$$\mu^m_{it}(h)(s_i) = \mu^m_{Ait}(h)(s_i) - \mu^m_{Bit}(h)(s_i),$$

where $\mu^m_{oit} : \mathcal{H} \to \mathbb{R}^{\#S_i}$ for $o \in \{A, B\}$ are given by

$$\mu^m_{oit}(h)(s_i) = \frac{\sum_{\mathtt{u}\in U_{oi}} \xi^m_{\mathtt{u}oit}(h)(s_i)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{oi}} \xi^m_{\mathtt{u}oit}(h)(s_i)} \tag{4}$$

for every $h \in \mathcal{H}$, $s_i \in S_i$, and $t \geq 2$ such that the denominator is greater than zero. If the denominator is zero, then $\mu^m_{oit}(h)(s_i) \in (-2K, 2K)$.[18]

In words, for non-strategic players ($m = N$), this definition assumes that they believe that the difference in expected payoffs from A and B, conditional on an observed signal, is given by the *observed* empirical average difference in payoffs—the key here is that only the payoff to the chosen alternative is observed. Hence, non-strategic players take the information they see at face value. In particular, they make no attempts to account for the information content of others' votes.

For strategic players ($m = NE$), this definition assumes that they form beliefs in a similar way, except that they only use data from periods in which their vote was pivotal. However, since counterfactual payoffs are not observed, players never observe the *difference* in expected payoffs conditional on their vote being pivotal. Instead, players estimate the payoffs for each alternative by observing each alternative in

[17]Throughout the paper, $\mathbf{1}$ stands for the indicator function, i.e., $\mathbf{1}_A(z) = 1$ if $z \in A$ and $\mathbf{1}_A(z) = 0$ if $z \notin A$.

[18]The assumption on the initial prior guarantees that posteriors always belong to $(-2K, 2K)$; hence, the perturbations guarantee that both alternatives are chosen with positive probability.

different periods, and then combining these estimates to form their estimate of the difference in payoffs.

The learning model is completed by assuming that players' votes asymptotically maximize their current period *perceived* expected utility.[19]

Definition 3. A policy rule $\phi_i^{H,m}$ is *asymptotically myopic relative to an assessment rule* μ_i^m if there exists a sequence $(\epsilon_t)_t$ such that $\epsilon_t \geq 0$ for all t and $\lim_{t\to\infty} \epsilon_t = 0$, and for every $h \in \mathcal{H}$, $s_i \in S_i$, and $t \geq 1$,

$$\phi_{it}^{H,m}(h, v_{it})(s_i) = \begin{cases} 1 & \text{if } \mu_{it}^m(h)(s_i) - v_{it} \geq \epsilon_t \\ 0 & \text{if } \mu_{it}^m(h)(s_i) - v_{it} < -\epsilon_t \end{cases}.$$

The policy rule is called *myopic* if $\epsilon_t = 0$ for all t.

The definition of asymptotic myopia allows players to be forward-looking and to experiment (cf. Fudenberg and Kreps 1993, 1994, 1995). In our environment, the reason why non-strategic ($m = N$) players may have incorrect beliefs is not because of lack of experimentation, but, rather, due to their failure to account for the selection problem. Indeed, even with myopic rules, the payoff perturbations guarantee "perpetual experimentation" in the sense that Lemma 1 holds for all t.

4.3 Stability and equilibrium

Our objective is to relate distributions over outcomes of the dynamic game as $t \to \infty$ to equilibrium distributions over outcomes of the stage game. For $t \geq 2$, define the sequence of random variables $\overline{P}_t : \mathcal{H} \to \Delta(Z)$, where

$$\overline{P}_t(h)(z) = \frac{1}{t-1} \sum_{\tau=1}^{t-1} \mathbf{1}_{\{z\}}(z_\tau)$$

is the frequency distribution over outcomes in the dynamic game. We look at the frequency distribution in order to allow for the possibility that play in the dynamic game is correlated. We focus attention on frequency distributions that eventually stabilize around a steady-state distribution over outcomes. The following definitions

[19]Implicitly, we assume that players believe (correctly) that they can be pivotal with positive probability.

of stability account for the probabilistic nature and possible multiplicity of steady states.

Definition 4. $P^m \in \Delta(Z)$ is a *stable outcome distribution* of the dynamic game under policy rules $\phi^{H,m}$ if for all $\varepsilon > 0$ there exists t_ε such that[20]

$$\mathbf{P}^{\phi^{H,m}}\left(\left\|\overline{P}_t - P^m\right\| < \varepsilon \text{ for all } t \geq t_\varepsilon\right) > 0.$$

The following definition is a stronger version of the previous one.

Definition 5. $P^m \in \Delta(Z)$ is a *strongly stable outcome distribution* of the dynamic game under policy rules $\phi^{H,m}$ if

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty}\left\|\overline{P}_t - P^m\right\| = 0\right) = 1.$$

The definition of stability captures the idea that after a *finite* number of periods, there is a positive probability that the frequency distribution over outcomes $\overline{P}_t$ remains forever close to P^m. The definition of strong stability states that the frequency distribution over outcomes $\overline{P}_t$ converges almost surely (under $\mathbf{P}^{\phi^{H,m}}$) to P^m.

Theorem 2. *P^m is an m-equilibrium distribution of the stage game if and only if P^m is stable (or strongly stable) under some policy rule $\phi^{H,m}$ that is* asymptotically *myopic relative to assessment rules μ^m that are m-empirical.*

Theorem 2 establishes two results. The "if" direction provides a justification for discarding outcomes that do not arise in an m-equilibrium of the voting stage game: Any profile that is not an m-equilibrium generates an outcome distribution that is not even stable, much less strongly stable, in the dynamic game with m-learning rules. In particular, correlated strategy profiles do not generate stable outcome distributions in our environment. As the proof makes clear, this result follows from the assumption that payoff perturbations are independent across players and time.

[20] The norm $||\cdot||$ is defined as $||f|| = \max_{y\in Y}|f(y)|$.

The "only if" direction provides a justification for not discarding any m-equilibrium distribution outcome: Any m-equilibrium distribution outcome is strongly stable under some m-learning rule. This result does rely on allowing for rules that are asymptotically optimal, and, arguably, the proof relies on the construction of a particular type of policy rule (see Fudenberg and Kreps (1993) for a discussion). This result leaves open the possibility of placing additional restrictions on policy rules in order to select among the set of m-equilibria, a topic that we study in Section 5.

4.4 Proof of Theorem 2

The proof of Theorem 2 adapts the arguments by Fudenberg and Kreps (1993, Propositions 6.3 and 7.5) to our asymmetric-information setting. The proof relies on several claims, all of which are proven in the Appendix.

4.4.1 Proof of the "if" part

Throughout the proof, we fix a stable outcome distribution P^m and policy rules $\phi^{H,m}$ that are asymptotically myopic—with a fixed sequence $(\epsilon_t)_t$—relative to assessment rules μ^m which are m-empirical. The proof compares "strategies" in the dynamic game with strategies in the stage game. To define the former, let the vector-valued random variable $\alpha_t^{H,m} = (\alpha_{1t}^{H,m}, ..., \alpha_{nt}^{H,m}) : \mathcal{H} \to \Pi_{i=1}^n \mathcal{A}_i$ denote a time-t strategy profile, where

$$\alpha_{it}^{H,m}(h)(s_i) = \int \mathbf{1}_{\left\{v_i:\phi_{it}^{H,m}(h,v_i)(s_i)=1\right\}} dF_i \tag{5}$$

is the probability that player i votes for A when observing signal s_i, conditional on history h^t.

Finally, let $\alpha^m = (\alpha_1^m, ..., \alpha_n^m) \in \Pi_{i=1}^n \mathcal{A}_i$ be such that

$$\alpha_i^m(s_i) = F_i\left(\Delta_i^m(P^m, s_i)\right) \tag{6}$$

is the probability that player i votes for A if she optimally responds to beliefs $\Delta_i^m(P^m, s_i)$.

The proof of the "if" part of Theorem 2 follows from the following claims.

Claim 2.1 For all $\varepsilon > 0$, there exists H_ε with $\mathbf{P}^{\phi^{H,m}}(H_\varepsilon) > 0$ such that for all $h \in H_\varepsilon$, there exists $t_{\varepsilon,h}$ such for all $t \geq t_{\varepsilon,h}$, $\left\|\alpha_t^{H,m}(h) - \alpha^m\right\| < \varepsilon$ and $\left\|\overline{P}_t(h) - P^m\right\| <$

ε.

Claim 2.2 $||P^m - P(\alpha^m)|| = 0$

Claim 2.1 establishes that stability of P^m implies that beliefs eventually remain close to $\Delta_i^m(P^m, s_i)$, thus implying that time-t strategies $\alpha_t^{H,m}$ eventually remain close to α^m. The key of the proof is that players' payoff perturbations are independently drawn from an atom-less distribution, implying that if beliefs settle down, then strategies must also settle down, not just in an average sense, but actually in a per-period sense. In particular, Claim 2.1 implies that any correlation in players' strategies induced by a common history eventually vanishes. In Claim 2.2, we show that the fact that time-t strategies remain close to α^m implies that $P^m = P(\alpha^m)$, where $P(\cdot)$ was previously defined in (1). Both claims rely on a straightforward generalization of a technical result by Fudenberg and Kreps (1993, Lemma 6.2); this result allows us to apply the law of large numbers in a context where a sequence of random variables is not independently distributed, but where the distributions conditional on past history are eventually very close to some common distribution.

Claim 2.2 and equation (6) imply that, for all i and s_i,

$$\alpha_i^m(s_i) = F_i\left(\Delta_i^m(P(\alpha^m), s_i)\right),$$

so that α^m is an m-equilibrium of the stage game. Therefore, $P^m = P(\alpha^m)$ is an m-equilibrium distribution, thus establishing the "if" part of Theorem 2.

4.4.2 Proof of the "only if" part

The idea is to postulate that players follow the rule associated with the m-equilibrium strategies, unless the m-empirical beliefs are too different from the m-equilibrium beliefs. We then show that under these policy rules, the m-empirical beliefs are indeed close to the m-equilibrium beliefs with high probability.

Throughout the proof, we fix m and let P^m denote an m-equilibrium distribution of the voting stage game. We divide the proof into several steps.

Step 1. Construction of policy rules. For all (i, s_i), h and v_i, let

$$\hat{\phi}_{it}^{H,m}(h, v_i)(s_i) = \begin{cases} 1 & \text{if } \gamma_t^i(h)(s_i) > v_{it} \\ 0 & \text{otherwise} \end{cases},$$

where

$$\gamma_t^i(h)(s_i) \equiv \begin{cases} \Delta_i^m(P^m, s_i) & \text{if } \delta_t^m(h) < \epsilon_t \\ \mu_t^m(h)(s_i) & \text{otherwise} \end{cases},$$

where

$$\delta_t^m(h) \equiv ||\mu_t^m(h) - \Delta^m(P^m, \cdot)||.$$

The particular sequence $(\epsilon_t)_t$ is provided in the next step.

Step 2. Construction of $(\epsilon_t)_t$. First, define the history-independent policy rule

$$\phi_i^{*,m}(h, v_{it})(s_i) = \begin{cases} 1 & \text{if } \Delta_i^m(P^m, s_i) > v_{it} \\ 0 & \text{otherwise} \end{cases}.$$

Claim 2.3 $\lim_{t\to\infty} \delta_t^m = 0$ a.s.-$\mathbf{P}^{\phi^{*,m}}$

Claim 2.3 establishes that, under the probability induced by $\phi^{*,m}$, δ_t goes to zero with probability one. In addition, $\delta_t^m(h) < 4K$ for all t and h. Then, by the second Borel-Cantelli lemma (see Billingsley (1995), Theorem 4.4), it follows that for any $\varrho > 0$,

$$\sum_{t=1}^{\infty} \mathbf{P}^{\phi^{*,m}} \left(\delta_t^m(h) \geq \varrho\right) < \infty.$$

Therefore, there exists a sequence $(\tau(j))_{j=1}^{\infty}$ with $\lim_{j\to\infty} \tau(j) = \infty$ such that

$$\sum_{t \geq \tau(j)} \mathbf{P}^{\phi^{*,m}}(\delta_t^m(h) \geq 1/j) < (3/2)4^{-j}. \tag{7}$$

We now choose the sequence $(\epsilon_t)_t$. For $t \leq t^* \equiv \tau(1) < \infty$, we pick any $\epsilon_t \geq 4K$, while for $t > t^*$, we set $\epsilon_t = 1/N(t)$, where $N(t) \equiv \sum_{j=1}^{\infty} 1\{\tau(j) \leq t\}$. Note that, since $\tau(j) \to \infty$ as $j \to \infty$, then $N(t) \to \infty$ as $t \to \infty$. So $\epsilon_t \geq 0$ for all t and $\lim_{t\to\infty} \epsilon_t = 0$. In particular, given this construction of $(\epsilon_t)_t$, the policy rules $\hat{\phi}^{H,m}$ are asymptotically myopic relative to assessment rules μ^m.

Step 3. In the final step, we verify that P^m is strongly stable under $\hat{\phi}^{H,m}$ with our particular choice of $(\epsilon_t)_t$. Let α^m denote the m-equilibrium strategy profile corresponding to P^m, i.e., $P(\alpha^m) = P^m$. Define $H_o \equiv \{h \in H : \delta_t^m(h) < \epsilon_t, \forall t\}$. By construction, for any $h \in H_o$,

$$\begin{aligned}\alpha_{it}^{H,m}(h)(s_i) &= \int \mathbf{1}_{\{v_i : \hat{\phi}_{it}^{H,m}(h,v_i)(s_i)=0\}} dF_i \\ &= F_i(\Delta_i^m(P^m, s_i)) \\ &= \alpha_i^m(s_i)\end{aligned}$$

for all (i, s_i). Thus, for all $h \in H_o$,

$$\mathbf{P}^{\hat{\phi}^{H,m}}(z_t = z \mid h^t) = P(\alpha_t^{H,m}(h))(z) = P(\alpha^m)(z) = P^m(z) \tag{8}$$

for all $z \in Z$.

Claim 2.4 $\mathbf{P}^{\hat{\phi}^{H,m}}(H_o) = 1$

Claim 2.4, which follows by first showing that $\mathbf{P}^{\hat{\phi}^{H,m}}(H_o) > 0$ and then applying the 0-1 Law, implies that P^m is strongly stable under $\hat{\phi}^{H,m}$.

5 Equilibrium selection

Theorem 2 justifies all m-equilibria and, therefore, provides no guidance for selecting among the multiple equilibria that often exist in voting games. This result holds because asymptotic myopia allows players to make (vanishing) optimization mistakes. In this section, we study stability of equilibria when players are restricted to be exactly myopic (i.e., $\epsilon_t = 0$ for all t in Definition 3).

The learning procedure followed by players in Section 4 yields a stochastic dynamical system in $\xi_t^m = (\xi_{\text{ou}it}^m(s_i))_{\text{ou}is_i}$.[21] Conveniently, (3) can be written recursively as

$$\xi_{\text{ou}i\,t+1}^m(s_i) - \xi_{\text{ou}i\,t}^m(s_i) = \frac{1}{t+1}\left(\mathbf{1}_{Z_{\text{ou}i}^m(s_i)}(z_{t+1}) - \xi_{\text{ou}i\,t}^m(s_i)\right).$$

Associated with any state of beliefs ξ, we define the corresponding (myopic) strategy

[21] Equivalently, we could let μ_t^m denote the state space. We do not follow this approach because updating of μ_t^m is asynchronous, leading to additional technical issues that can be avoided in our setting. See Fudenberg and Takahashi (2011) for a model where updating is inherently asynchronous.

profile $\alpha(\xi) = (\alpha_i(s_i;\xi))_{i,s_i}$, where

$$\alpha_i(s_i;\xi_i) = F_i\left(\frac{\sum_{\mathtt{u}\in U_{Ai}} \xi_{A\mathtt{u}i}(s_i)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Ai}} \xi_{A\mathtt{u}i}(s_i)} - \frac{\sum_{\mathtt{u}\in U_{Bi}} \xi_{B\mathtt{u}i}(s_i)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Bi}} \xi_{B\mathtt{u}i}(s_i)}\right). \tag{9}$$

Letting

$$\zeta^m_{o\mathtt{u}i\,t+1}(s_i) = \mathbf{1}_{Z^m_{o\mathtt{u}i}(s_i)}(z_{t+1}) - E_{P(\alpha(\xi^m_t))}\left(\mathbf{1}_{Z^m_{o\mathtt{u}i}(s_i)}(z_{t+1})\right),$$

then

$$\xi^m_{o\mathtt{u}i\,t+1}(h)(s_i) - \xi^m_{o\mathtt{u}i\,t}(h)(s_i) = \frac{1}{t+1}\left(E_{P(\alpha(\xi^m_t))}\left(\mathbf{1}_{Z^m_{o\mathtt{u}i}(s_i)}(z_{t+1})\right) - \xi^m_{o\mathtt{u}i\,t}(h)(s_i) + \zeta^m_{o\mathtt{u}i\,t+1}(s_i)\right).$$

We can then apply results from stochastic approximation theory to show that the stochastic system can be well-approximated by the following system of ordinary differential equations:

$$\dot{\xi}^m_t = L^m(\xi^m_t), \tag{10}$$

where, dropping the time subscripts,

$$L^m_{o\mathtt{u}i}(\xi^m)(s_i) \equiv E_{P(\alpha(\xi^m))}\left(\mathbf{1}_{Z^m_{o\mathtt{u}i}(s_i)}\right) - \xi^m_{o\mathtt{u}i}(s_i).$$

The *m-stationary points* of (10) are those states satisfying $\xi^m_* = L^m(\xi^m_*)$. We now relate these stationary points to the m-equilibria of the game.

Lemma 2. *If ξ^m is an m-stationary point, then $\alpha(\xi^m)$ is an m-equilibrium of the voting game. Conversely, if α^m is an m-equilibrium, then ξ^m such that $\xi^m_{o\mathtt{u}i}(s_i) = E_{P(\alpha^m)}\left(\mathbf{1}_{Z^m_{o\mathtt{u}i}(s_i)}\right)$ for all $o, \mathtt{u}, i, s_i$ is a stationary point and $\alpha(\xi^m) = \alpha^m$.*

Proof. See the Appendix. □

Definition 6. Let $\phi^{H,m}$ be a policy rule that is myopic relative to assessment rules μ^m that are m-empirical. A strategy profile α is *asymptotically stable* if

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} \alpha(\xi^m_t) = \alpha\right) > 0$$

and it is *unstable* if

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} \alpha(\xi^m_t) = \alpha\right) = 0.$$

As is well-known, the stability of (10) at a particular stationary point ξ_*^m is determined by studying the Jacobian of L^m at the stationary point, i.e., $J^m(\xi_*^m) \equiv \frac{\partial L^m}{\partial \xi^m}\Big|_{\xi_*^m}$. We say that a stationary point ξ_*^m is *linearly stable* if all eigenvalues of the Jacobian of L^m at ξ_*^m have negative real part. We say that ξ_* is *linearly unstable* if some eigenvalue has positive real part.

We use the following technical assumption and results from the theory of stochastic approximation to derive the main result of this section.

A1. For all i, F_i is twice differentiable with continuous first derivative and bounded second derivative.

Theorem 3. *Let α be an m-equilibrium and let ξ^m be such that $\xi_{o\mathtt{u}i}^m(s_i) = E_{P(\alpha^m)}\left(\mathbf{1}_{Z_{o\mathtt{u}i}^m(s_i)}\right)$ for all $o, \mathtt{u}, i, s_i$. If ξ^m is linearly stable, then α is asymptotically stable. If ξ^m is linearly unstable, then α is unstable.*

Proof. See the Appendix. □

In the next section, we show how Theorem 3 provides a natural criterion for selecting among multiple equilibria in voting games.[22]

6 Application: equilibrium selection in the jury model

We apply the results in Section 5 to a *symmetric* voting game with standard monotonicity assumptions. This game has received substantial attention in the literature, and we consider general payoffs that include the status quo game (e.g., Figure 1) and the jury game (e.g., Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998)) as special cases. Let $\Omega = \{\omega_A, \omega_B\}$ with $\omega_B < \omega_A$ and $S = \{s_1, ..., s_H\}$ with $s_1 < ... < s_H$. The assumptions in Section 3 are supplemented with the following standard assumptions.

A2. (Symmetry) $u_i = u$, $q_i = q$, $S_i = S$, and $F_i = F$ for all i.

[22] We are unable to show that myopic rules will converge with probability one to an equilibrium (as we obtained in Section 4 under *asymptotically* myopic rules). Roughly speaking, we can show convergence to a chain transitive invariant set of (10) (see Benaim and Hirsch (1999) for definitions), but we are unable to characterize such set for the voting game. In the game-theory literature, global convergence has only been obtained in special classes of games–e.g. zero-sum, potential, and supermodular games (Hofbauer and Sandholm, 2002).

A3. (A best under ω_A; B best under ω_B) $u(A,\omega_A) > u(B,\omega_A)$ and $u(A,\omega_B) < u(B,\omega_B)$.

A4. (MLRP) For all $\omega' > \omega$ and $s' > s$: $q(s' \mid \omega')q(s \mid \omega) > q(s' \mid \omega)q(s \mid \omega')$.

Most applications refine the set of equilibria by focusing on symmetric Nash equilibria that are *responsive*, in the sense that players are pivotal with positive probability. In our setup, the restriction to responsive equilibria (when they exist) is formalized by having (small) payoff perturbations.[23] As shown by Esponda (2011) for the voting game, we can think of the limit of a sequence of equilibria with vanishing payoff perturbations as the (refined) equilibria of the unperturbed game.

We define α to be a *responsive m-equilibrium of the unperturbed game* if, for all i and s, $\Delta_i^m(P(\alpha), s)$ is well-defined and, in addition, $\alpha_i(s) > 0$ implies $\Delta_i^m(P(\alpha), s) \geq 0$ and $\alpha_i(s) < 1$ implies $\Delta_i^m(P(\alpha), s) \leq 0$. We say that the equilibrium is *strict* if $\Delta_i^m(P(\alpha), s) \neq 0$ for all i, s. We do *not* restrict attention to *symmetric* responsive equilibria, since, as we show next, such equilibria may be unstable.

We fix a sequence $(F^\eta)_\eta$ of payoff perturbations that satisfies the following assumptions. We use $(f^\eta)_\eta$ to denote the sequence of densities (with respect to Lebesgue) corresponding to $(F^\eta)_\eta$.

A5. (Vanishing perturbations) For every $\eta > 0$ and $i = 1, ..., n$, (i) assumption A1 is satisfied; (ii) $\lim_{\eta \to 0} f^\eta(v) = 0$ for all $v \neq 0$; (iii) $\lim_{\eta \to 0} f^\eta(v) f^\eta(0) = 0$ for all $v \neq 0$; (iv) f^η is single-peaked at 0.[24]

Assumption A5(i) is needed to apply Theorem 3 and A5(ii) requires perturbations to vanish as η becomes small. Assumptions A5(iii-iv) are further regularity conditions that are used to prove that mixed-strategy Nash equilibria are unstable. Several family of distributions satisfy A5, such as the family of Normal distributions with zero mean and variance that depends on η and goes to zero as $\eta \to 0$. Subsequently, we index the elements introduced in the previous section by η.

Proposition 1. *Suppose that A2-A5 hold and that players follow a policy rule $\phi^{H,m}$ that is myopic relative to assessment rules μ^m that are m-empirical. Let $(\alpha^{m,\eta})_\eta$ be*

[23] The rationale behind this restriction is to rule out trivial equilibria where everyone votes for the same alternative because no one's vote can then change the outcome of the election.

[24] We say a function is single-peaked at m if there exists an m such that f is increasing (decreasing) for all $x < m$ ($x > m$).

a sequence of m-equilibria such that $\lim_{\eta \to 0} \alpha^{m,\eta} = \alpha^m$, *where* α^m *is a responsive m-equilibrium of the unperturbed game.*

(i) If α^m *is a (not-necessarily-symmetric) strict pure-strategy m-equilibrium, then* $\alpha^{m,\eta}$ *is asymptotically stable for all sufficiently small* η.

(ii) If α^{NE} *is a symmetric mixed-strategy Nash equilibrium, then* $\alpha^{NE,\eta}$ *is unstable for all sufficiently small* η.

Proposition 1 establishes that (strict) pure-strategy equilibria (either Nash or non-strategic) are asymptotically stable, but symmetric mixed-strategy Nash equilibria are unstable. In the Online Appendix, we show that symmetric mixed-strategy non-strategic equilibria may be asymptotically stable or unstable depending on the primitives. The reason why some symmetric mixed-strategy non-strategic equilibria are asymptotically stable is that, under non-strategic voting, a player's own vote also affects her beliefs. As we discuss in Section 7.1, for a fixed profile of others' votes, the higher the probability a player votes for A, then the less optimistic she is that A is best. This feature is absent in a Nash equilibrium.[25]

Since most of the literature restricts attention to symmetric Nash equilibria (which are often mixed-strategy equilibria), Proposition 1 raises the question of whether the main results of the literature continue to hold when attention is restricted to asymptotically stable equilibria. In the rest of this section, we show that the result that information is aggregated under the restriction to symmetric Nash equilibria (Feddersen and Pesendorfer (1998), McLennan (1998)) continues to hold under the restriction to asymptotically stable Nash equilibria.

For any $x, y \in \{0, 1, ..., n\}$, let $D(x, y)$ denote the expectation of $u(A, \omega) - u(B, \omega)$ conditional on observing $x + y$ signals, where x is the number of signals that are s_L and y is the number of signals that are s_H. By A2-A4, D is decreasing in its first component and increasing in its second component.

A6. (Binary signals) $S = \{s_L, s_H\}$

A7. (Generic payoffs) $D(x, y) \neq 0$ for all $x, y \in \{0, 1, ..., n\}$.

Proposition 2. *Fix any* $n \geq 3$ *and suppose that A2-A4 and A6 hold. If there exists a symmetric responsive Nash equilibrium of the unperturbed game, then one of*

[25] For games with many players, Esponda and Pouzo (2012) show that all non-strategic equilibria are essentially outcome-equivalent, and, therefore, equilibrium selection becomes unnecessary.

the following pure-strategy profiles constitute a pure-strategy Nash equilibrium of the unperturbed game:

- r_n *players always vote* A *and* $n - r_n$ *vote sincerely, where* $r_n \in \{0, ..., k-1\}$
- r'_n *players always vote* B *and* $n - r'_n$ *vote sincerely, where* $r'_n \in \{0, ..., n-k\}$

Moreover, the equilibrium is strict if A7 holds.

Proof. By Theorem 1 in Esponda (2011), a symmetric responsive equilibrium exists if and only if

$$D(n-k+1, 0) < 0 < D(0, k). \quad (11)$$

Suppose that the profile where all players vote sincerely is not an equilibrium (otherwise, there is nothing to prove). For a player, the pivotal event is that $n-k$ players voted B and $k-1$ voted A. For the sincere profile, the pivotal event is equivalent to $n-k$ signals s_L and $k-1$ signals s_H. Thus, a player who observes s_L believes $D(n-k+1, k-1)$ and a player who observes s_H believes $D(n-k, k)$. Because the sincere profile is not an equilibrium, it must be the case that either (i) $D(n-k+1, k-1) > 0$ (so the player would like to vote A after s_L) or (ii) $D(n-k, k) < 0$ (so the player would like to vote B after s_H).

First, suppose that (i) holds. Consider the profile where 1 player always votes A and $n-1$ players vote sincerely. The player who always votes A faces all sincere players, and the previous analysis and the assumption that (i) holds implies that this player best responds. Consider a player who plays sincerely and faces one player who always votes A and the remaining sincere players. The pivotal event is equivalent to $n-k$ signals s_L and $k-2$ signals s_H. Voting for A after s_H is a best response because $D(n-k, k-1) > D(n-k+1, k-1) > 0$. Voting for B after s_L is a best response if and only if $D(n-k+1, k-2) \leq 0$. If the last inequality holds, then the proof is done. Suppose that it does not hold, so that $D(n-k+1, k-2) > 0$. Then we consider the profile where 2 players always vote A and $n-2$ players vote sincerely. By the previous analysis, the players who always vote A are best responding. For a player who votes sincerely, the pivotal event is equivalent to $n-k$ signals s_L and $k-3$ signals s_H. Voting for A is a best response because $D(n-k, k-2) > D(n-k+1, k-2) > 0$. Voting for B after s_L is a best response if and only if $D(n-k+1, k-3) \leq 0$. If the last inequality holds, then the proof is done. If it does not hold, then we consider a profile where 3 players always vote A and repeat the argument. As long as we have not found an equilibrium, we continue increasing the number of players who vote always A, until we reach the case where there are $k-1$ players who always vote A and $n-k+1$ who vote sincerely. As usual, the players who always vote A are best responding. For the sincere players, the pivotal event is equivalent to $n-k$ signals s_L and no information about the number of s_H signals. Then, voting for A after s_H is a best response (otherwise, the profile with $k-2$ players who always vote A and the remaining vote sincerely would have been an equilibrium and we would have not reached this stage). In addition, after she observes s_L she believes $D(n-k+1,0)$. It then follows by (11) that voting for B after s_L is also a best response. Therefore, the profile where $k-1$ players always vote A and the remaining players vote sincerely is an equilibrium.

Second, suppose that (ii) holds. Then we can follow the same argument as for case (i), except that now we consider profiles where some players always vote B, rather than A, and the remaining players vote sincerely. As long as we have not found an equilibrium, we continue increasing the number of players who vote always B, until we reach the case where there are $n-k$ players who always vote B and k who vote sincerely. We then use $D(0,k)>0$ from equation (11) to show that this profile is an equilibrium, provided that none of the previous profiles was an equilibrium. Finally, if A7 holds, then players are never indifferent in equilibrium. □

Proposition 3. *Suppose that A2-A4 and A6 hold and that A7 holds for all n. Let $k_n = \rho n$ be the number of votes required to elect A when there are n voters, where $\rho \in (0,1)$. Then there exists n' and a sequence of strict pure-strategy Nash equilibria of the unperturbed game $(\alpha^n)_{n \geq n'}$ such that the probability of electing A in state ω_A and B in state ω_B goes to one as $n \to \infty$.*

Proof. Since $\lim_{n\to\infty} D(n-k_n+1,0) = u(A,\omega_B)-u(B,\omega_B) < 0$ and $\lim_{n\to\infty} D(0,k_n) = u(A,\omega_A) - u(B,\omega_A) > 0$, then, by equation (11), there exists n' such that, for all $n \geq n'$, there exists a symmetric responsive Nash equilibrium of the n-player game. Then, by Proposition 2, there exists a sequence of strict pure-strategy Nash equilibria $(\alpha^n)_{n\geq n'}$. In addition, by continuity, there exist $0 < \rho' < \rho$ and $\rho < \rho'' < 1$ such that

$$\lim_{n\to\infty} D(n-k_n+1, \rho' n) < 0 < \lim_{n\to\infty} D(n(\rho''-\rho), k_n). \tag{12}$$

Then, by following the steps we followed in the proof of Proposition 2, but where (11) is replaced with (12), the equilibrium α^n is characterized by Proposition 2, but where now, for all sufficiently large n, $r_n \in \{0, ..., (\rho-\rho')n-1\}$ and $r'_n \in \{0, ..., (1-\rho'')n\}$. Fix any such sufficiently large n. First, consider the case where α^n is such that r_n players always vote A and $n-r_n$ vote sincerely. Then, it must be the case that sincere voting is a symmetric equilibrium of the game with $n-r_n$ voters where $\hat{k}_n = k_n - r_n$ is the number of players required to elect A. Since $n-r_n \geq (1-(\rho-\rho'))n+1 \to \infty$ as $n \to \infty$ and since $\hat{k}_n/n \in (\rho' - 1/n, \rho) \subset (0,1)$ for sufficiently large n, then it follows from the information aggregation result of Feddersen and Pesendorfer (1998, Proposition 3) for symmetric equilibria and non-anonymous rules that, if n is sufficiently large, then the probability of electing the right candidate is sufficiently close to 1.[26] Finally, consider the case where α^η is such that r'_n players always vote B and $n-r'_n$ vote sincerely. Then, it must be the case that sincere voting is a symmetric equilibrium of the game with $n-r'_n$ voters where k_n is the number of players required to elect A. Since $n-r'_n \geq \rho'' n \to \infty$ as $n \to \infty$ and since $k_n/n = \rho \in (0,1)$, then, once again, it follows from the previous information aggregation result that, if n is sufficiently large, then the probability of electing the right candidate is sufficiently close to 1. □

7 Discussion

7.1 Continuum of signals

In many voting games (e.g., Feddersen and Pesendorfer, 1997), the assumption that there is a finite number of signals is replaced with the assumption that there is a continuum of signals. In this case, Nash equilibrium strategies are essentially pure. However, to the extent that the continuum assumption may be a convenient approximation of a game with a large but finite number of signals, where equilibria are possibly in mixed-strategies, then the instability result in Proposition 1 also casts doubts on the standard selection of symmetric Nash equilibria in games with a continuum of signals.

For the case of non-strategic equilibrium, a continuum of signals need not even purify equilibrium strategies. To illustrate, suppose that player i always votes for A, $\alpha'_i(s) = 1$ for all s, or always votes for B, $\alpha_i(s) = 0$ for all s, so that her beliefs, given

[26]Strictly speaking, Feddersen and Pesendorfer (1998) show this result for the payoffs of the jury model, but their proof extends trivially whenever the more general assumption A3 in our paper holds; in fact, Feddersen and Pesendorfer (1997) show that their result extends not only to payoffs that satisfy A3 but also to payoffs that differ across players.

that others follow strategy α_{-i} and that her signal is s, are given by $\Delta_i^N(P(\alpha_i', \alpha_{-i}), s)$ and $\Delta_i^N(P(\alpha_i, \alpha_{-i}), s)$, respectively. Under standard monotone assumptions on information, the above expressions for beliefs are increasing in the signal s. In addition, if α_{-i} is nondecreasing (but not constant), then $\Delta_i^N(P(\alpha_i', \alpha_{-i}), s) < \Delta_i^N(P(\alpha_i, \alpha_{-i}), s)$ for all s. Intuitively, a player who votes for A learns about A from the event that $k-1$ or more other players voted for A, while a player who votes for B learns about A from the event that k or more players voted for A. The second event is better news about A than the first event, and, therefore, beliefs about A are more optimistic when a player votes for B.[27] Suppose that s'' and s' are the signals at which beliefs cross zero when player i votes for A and when she votes for B, respectively. It then follows that $s' < s''$. Consider any $\hat{s} \in [s', s'']$. If player i votes for A, then $\Delta_i^N(P(\alpha_i', \alpha_{-i}), \hat{s}) < 0$ and she prefers to vote for B. But if player i votes for B, then $\Delta_i^N(P(\alpha_i, \alpha_{-i}), \hat{s}) > 0$ and she prefers to vote for A. Therefore, in equilibrium, a player who receives signal $\hat{s}$ will mix between A and B in a way that her beliefs are exactly zero. In particular, there is now an entire interval of signals under which player i will be mixing. The payoff perturbations in our paper, which, unlike the signals, are independent of the state of the world, allow us to purify these strategies and eliminate correlation in strategies.

7.2 Choice of learning rules

We chose to study learning rules which are realistic (based on the retrospective voting literature) and have a simple structure. The question of which learning rule is used by players is ultimately an empirical question and the answer is likely to depend on the setting and the sophistication of the participants.[28] Nevertheless, there are a few conceptual points that illustrate the subtleties involved in the choice of a learning rule.

The first point is that there are a few reasons why even a sophisticated player who understands selection may not use the strategic learning rule. Of course, one reason

[27] A similar logic applies if the player is also learning about alternative B.

[28] In experimental contexts, Guarnaschelli, McKelvey, and Palfrey (2000), Ali et al. (2008) and Battaglini, Morton, and Palfrey (2010) find that voters sometimes vote against their signals and respond to changing rules, and interpret it as evidence of strategic voting. But this response is also consistent with the non-strategic equilibrium analyzed here. Esponda and Vespa (2011) find evidence of non-sophistication. None of these experiments take place in a learning context and, therefore, cannot provide a direct test of our learning rules.

is that, in some settings, vote shares (and consequently whether a vote is pivotal) may not be observed. A less obvious reason is that pivotal learning is not necessarily an optimal way to learn, since sophisticated players may initially face an endogenous selection bias.[29] *If* a player's vote were random and independent of the state of the world (conditional on her private information), then the subsample where she is pivotal would not be biased—the reason is that whether or not she would observe the payoff of an alternative in those periods in which she is pivotal would depend only on whether or not she votes for the alternative, which would be independent of the state of the world. While independence may hold in steady state, voting behavior is likely to be correlated as players are learning to play the game under a common history. More generally, 'optimal' behavior is straightforward for non-strategic voters that do not account for selection, but it is far from obvious for sophisticated players. The reason is that 'optimality' for sophisticated players must depend on their belief of the learning rules being followed by other players.[30] A further issue with pivotal learning is that the inferences that a player makes, conditional on being pivotal, are not suitable for replication, either when facing a similar decision problem with a different group of people (who may have learned to make decisions in a different way) or when facing a decision on her own.

Finally, a sophisticated player may not be able to follow the strategic rule in large elections given the low probability of being pivotal and, therefore, the lack of data from which to learn. In this case, it may be natural to consider an extension of the strategic rule, whereas voters condition not on the event that their vote is exactly pivotal, but, for example, on the event that their vote is close to pivotal. To formalize this extension, let $\rho \equiv k/n$ denote the proportion of voters required to elect A and let $m \in [0,1]$. We can then generalize our previous definition of piv_i^m by letting

$$piv_i^m(\rho) = \left\{ x_{-i} \in \Pi_{j \neq i} X_j : \left| \frac{\sum_{j \neq i} x_j + 1}{n} - \rho \right| \leq m \right\}.$$

For example, $piv_i^{.05}(1/2)$ denotes the event that the proportion of others' votes is in

[29] This bias disappears as the number of players increases and a player becomes negligible; however, the proportion of the sample that can be used to make inferences also goes to zero.

[30] For example, if a sophisticated player believes that every other player is non-strategic, then she has a well-defined (though nontrivial) dynamic optimization problem to solve. But why not believe, instead, that other players believe that every other player is non-strategic? Or even higher order beliefs?

the range 0.5 ± 0.05. In particular, the case $m = 0$ corresponds to the conditioning event in a Nash equilibrium and $m = 1$ corresponds to a non-strategic equilibrium. Theorems 2 and 3 extend naturally to this more general definition of the pivotal event, thus resulting in a continuum of equilibrium concepts.

7.3 Self-confirming equilibrium and reinforcement learning

It is natural to relate our result that the strategic rule justifies Nash equilibrium to the notion of a self-confirming equilibrium in games with payoff-uncertainty (Dekel, Fudenberg, and Levine, 2004). In a self-confirming equilibrium, players best respond to beliefs that are not necessarily correct but that must be consistent with the feedback obtained. In our environment, consider the case where feedback includes a player's own payoff, the elected alternative, and whether she was pivotal or not. Suppose that there are no payoff perturbations. Then, it is not the case that every self-confirming equilibrium must be a Nash equilibrium. The reason is that players may be stuck electing always the same outcome. This implies that a learning justification for self-confirming equilibrium (e.g., Fudenberg and Kreps, 1995) will not necessarily justify Nash equilibrium. Now consider the opposite case, where payoffs are randomly perturbed and the only feedback is own payoff (not even the elected alternative is observed). In this case, the set of self-confirming equilibria must coincide with the set of Nash equilibria. The reason is that, because a player chooses both actions in equilibrium and observes her payoff, then she must have correct beliefs about the expected equilibrium payoff from choosing each of these actions. Thus, the crucial aspect in justifying Nash equilibrium in our setup is not the feedback (besides, of course, payoff feedback), but rather the randomly perturbed payoffs.

As argued above, given the payoff perturbations, then the amount of feedback (beyond payoff feedback) plays no role in the characterization of self-confirming equilibrium. But, in an explicit learning model, such as our paper, the role of feedback does play two important roles in justifying Nash equilibrium. First, the actual stochastic dynamics, applied in Sections 5 and 6 to select among equilibria, do rely on the particular learning rules. Second, the plausibility of these learning rules in turn depend on being able to observe certain feedback.

Relatedly, there is another obvious learning rule that would also justify Nash equilibrium in the voting game but does not even require feedback (beyond payoff feedback). Suppose that players don't have a good understanding of the problem at hand. All they know is that they must vote for either A or B every period, and have no understanding of how their vote affects outcomes, though they do keep track of their payoffs as a function of their own votes (but not of the election outcome). Suppose they compare past payoffs from voting for each alternative and eventually choose the action with the highest payoffs.[31] Since players will learn the correct payoffs from voting for either alternative, this *reinforcement learning* rule also justifies Nash equilibrium. For the voting game, however, it seems unrealistic that voters would not seek to learn the expected payoffs of the alternatives, which is the reason why we choose to focus on different types of rules. More generally, while the notion of self-confirming equilibrium can help us understand when feedback will be sufficient to restrict attention to Nash equilibrium, there are different ways in which players may have correct beliefs and, therefore, an explicit learning model must still take a stand on the choice of a learning rule, which must often be justified in the context of a particular application.

8 Appendix

8.1 Proof of Claims 2.1-2.4.

Proof of Claim 2.1. By continuity of F_i, it suffices to show that for all $\varepsilon > 0$, there exist $\gamma(\varepsilon) > 0$ with $\lim_{\varepsilon \to 0} \gamma(\varepsilon) = 0$ and H_ε with $\mathbf{P}^{\phi^{H,m}}(H_\varepsilon) > 0$ such that for all $h \in H_\varepsilon$, there exists $t_{\varepsilon,h}$ such for all $t \geq t_{\varepsilon,h}$, all i, and all s_i,

$$F_i\left(\Delta_i(P^m, s_i) - \gamma(\varepsilon) - \epsilon_t\right) \leq \alpha_{it}^{H,m}(h)(s_i) \leq F_i\left(\Delta_i(P^m, s_i) + \gamma(\varepsilon) + \epsilon_t\right) \qquad (13)$$

and

$$\left|\overline{P}_t(h)(z) - P^m(z)\right| < \varepsilon \qquad (14)$$

for all $z \in Z$.

With a slight abuse of notation, let $Z_{oi}^m(s_i) \equiv \cup_{\mathrm{u} \in U_{oi}} Z_{o\mathrm{u}i}^m(s_i)$. Note that $\mu_{oit}^m(h)(s_i)$

[31] For example, if a player votes for A but B is the election outcome, then she will attribute the payoff she obtains from B to her choice of A.

in equation (4) can be written as

$$\mu_{oit}^m(h)(s_i) = \frac{\sum_{\mathtt{u}\in U_{oi}} \overline{P}_t(h)(Z_{o\mathtt{u}i}^m(s_i))\mathtt{u}}{\overline{P}_t(h)(Z_{oi}^m(s_i))} \tag{15}$$

provided that $\overline{P}_t(h)(Z_{oi}^m(s_i)) > 0$.

Because P^m is stable, for all $\varepsilon > 0$, there exists t_ε and H_ε^* with $\mathbf{P}^{\phi^{H,m}}(H_\varepsilon^*) > 0$ such that for all $h \in H_\varepsilon^*$ and $t \geq t_\varepsilon^*$, equation (14) holds for all $z \in Z$. In addition, (14) implies that

$$\left|\overline{P}_t(h)(Z') - P^m(Z')\right| < \varepsilon \times \#Z \tag{16}$$

for all $Z' \subset Z$. Next, let $\varsigma = \min\left\{\mathbf{P}^{\phi^{H,m}}(H_\varepsilon^*), .5K_p\right\} > 0$, where $K_p > 0$ is defined by equation (35) in the Online Appendix. By Lemma 4 in the Online Appendix, for all $h \in H \setminus H^o$ (where H^o has zero measure) there exists $t_{\varsigma,h}$ such that for all $t \geq t_{\varsigma,h}$ and $o \in \{A, B\}$

$$\overline{P}_t(h)(Z_{ois_i}^m) > K_p - \varsigma \geq .5K_p > 0. \tag{17}$$

Let $H_\varepsilon = H_\varepsilon^* \cap H \setminus H^o$, and note that by our choice of H_ε^*, $\mathbf{P}^{\phi^{H,m}}(H_\varepsilon) > 0$. Therefore, for all $\varepsilon > 0$, there exists H_ε with $\mathbf{P}^{\phi^{H,m}}(H_\varepsilon) > 0$ such that for all $h \in H_\varepsilon$ and $t \geq t_{\varepsilon,h} \equiv \max\{t_\varepsilon^*, t_{\varsigma,h}\}$, all i, s_i

$$|\mu_{it}^m(h)(s_i) - \Delta_i^m(P^m, s_i)| \leq \gamma(\varepsilon) \equiv \frac{(\varepsilon \times \#Z) \times (0.5K_p)^2}{2K(1 + \#\Omega) + 0.5(\varepsilon \times \#Z)K_p} \underset{\varepsilon\to 0}{\longrightarrow} 0 \tag{18}$$

where the inequality follows from (15), (16), (17), the facts that $|\mathtt{u}| < K$ and $\#\Omega < \infty$, and simple algebra that uses the fact that

$$\Delta_i^m(P^m, s_i) = \frac{\sum_{\mathtt{u}\in U_{Ai}} P^m(Z_{A\mathtt{u}i}^m(s_i))\mathtt{u}}{P^m(Z_{Ai}^m(s_i))} - \frac{\sum_{\mathtt{u}\in U_{Bi}} P^m(Z_{B\mathtt{u}i}^m(s_i))\mathtt{u}}{P^m(Z_{Bi}^m(s_i))}.$$

Then, equation (18) and the definition of the policy rules imply that

$$\phi_{it}^{H,m}(h, v_i)(s_i) = \begin{cases} 1 & \text{if } v_i \leq \Delta_i^m(P^m, s_i) - \gamma(\varepsilon) - \epsilon_t \\ 0 & \text{if } v_i > \Delta_i^m(P^m, s_i) + \gamma(\varepsilon) + \epsilon_t \end{cases},$$

so that (13) holds by (5). □

Proof of Claim 2.2. Note that for each $z \in Z$,

$$\mathbf{P}^{\phi^{H,m}}(z_t = z \mid h^t) = P(\alpha_t^{H,m}(h))(z).$$

Then, Claim 2.1 and the fact that $P(\cdot)$ is continuous imply that for all $\varepsilon > 0$, there exists H_ε with $\mathbf{P}^{\phi^{H,m}}(H_\varepsilon) > 0$ such that for all $h \in H_\varepsilon$, there exists $\hat{t}_{\varepsilon,h}$ such that for all $t \geq \hat{t}_{\varepsilon,h}$,

$$\left|\mathbf{P}^{\phi^{H,m}}(z_t = z \mid h^t) - P(\alpha^m)(z)\right| < \varepsilon \tag{19}$$

and

$$\left|\overline{P}_t(h)(z) - P^m(z)\right| < \varepsilon \tag{20}$$

for all $z \in Z$.

Then, by equation (19) and Lemma 3 in the Online Appendix applied to all the singleton sets of Z,

$$\limsup_{t\to\infty} \overline{P}_t(h)(z) \leq P(\alpha^m)(z) + \varepsilon \quad \text{ and } \quad \liminf_{t\to\infty} \overline{P}_t(h)(z) \geq P(\alpha^m)(z) - \varepsilon \tag{21}$$

for all $z \in Z$, almost surely on H_ε.

By the triangle inequality, for any t,

$$||P^m - P(\alpha^m)|| \leq ||P^m - \overline{P}_t(h)|| + ||\overline{P}_t(h) - P(\alpha^m)|| \tag{22}$$

for any $h \in H_\varepsilon$; we pick one $h \in H_\varepsilon$ (outside the measure zero set). By equation (21), the second summand in the RHS of (22) is less than ε for all t sufficiently large; by equation (20), the first summand of the RHS is also less than ε for all t sufficiently large. Hence, $||P^m - P(\alpha^m)|| \leq \varepsilon$; since this holds for all $\varepsilon > 0$, then we obtain the desired result by taking $\varepsilon \to 0$. $\square$

Proof of Claim 2.3. Since

$$\mathbf{P}^{\phi^{*,m}}(z_t = z \mid h^t) = P(\alpha^m)(z),\ \forall (t, h^t, m) \text{ and } z \in Z,$$

it follows by the SLLN (see Billingsley, 1995, Theorem 6.1) and the fact that $\#Z < \infty$ that

$$\lim_{t\to\infty} ||\overline{P}_t(h) - P^m|| = 0,\ \text{a.s.-}\mathbf{P}^{\phi^{*,m}}. \tag{23}$$

Also, since $P^m(Z^m_{ois_i}) > 0$ (by Lemma 1), equation (23) implies that, for all sufficiently large t, $\overline{P}_t(h)(Z^m_{oi}(s_i)) > 0$ and, for all (i, s_i),

$$\begin{aligned}
|\mu^m_{it}(h)(s_i) - \Delta^m_i(P^m, s_i)| &\leq \max_{o\in\{A,B\}} \left| \frac{\sum_{\mathtt{u}\in U_{oi}} \overline{P}_t(h)(Z^m_{o\mathtt{u}i}(s_i))\mathtt{u}}{\overline{P}_t(h)(Z^m_{oi}(s_i))} - \frac{\sum_{\mathtt{u}\in U_{oi}} P^m(Z^m_{o\mathtt{u}i}(s_i))\mathtt{u}}{P^m(Z^m_{oi}(s_i))} \right| \\
&\leq \max_{o\in\{A,B\}} \left| \frac{\sum_{\mathtt{u}\in U_{oi}} \{\overline{P}_t(h)(Z^m_{o\mathtt{u}i}(s_i)) - P^m(Z^m_{o\mathtt{u}i}(s_i))\}\mathtt{u}}{\overline{P}_t(h)(Z^m_{oi}(s_i))} \right| \\
&\quad + \max_{o\in\{A,B\}} \left| \frac{\sum_{\mathtt{u}\in U_{oi}} P^m(Z^m_{o\mathtt{u}i}(s_i))\mathtt{u}}{\overline{P}_t(h)(Z^m_{oi}(s_i))P^m(Z^m_{oi}(s_i))} \left(\overline{P}_t(h)(Z^m_{oi}(s_i)) - P^m(Z^m_{oi}(s_i))\right) \right|
\end{aligned}$$

a.s.-$\mathbf{P}^{\phi^*,m}$. Therefore, the facts that $|\mathtt{u}| \leq K$, $\#Z < \infty$, and, once again, equation (23) imply that $\delta^m_t \to 0$ as $t \to \infty$, a.s.-$\mathbf{P}^{\phi^*,m}$. $\square$

Proof of Claim 2.4. We first show that $\mathbf{P}^{\hat{\phi}^{H,m}}(H_o) > 0$ and then invoke the 0-1 Law to obtain the desired result. By definition

$$\begin{aligned}
\mathbf{P}^{\hat{\phi}^{H,m}}(H_o) &= \mathbf{P}^{\hat{\phi}^{H,m}}(\delta^m_t(h) < \epsilon_t, \forall t) \\
&= \prod_{t=t^*}^{\infty} \mathbf{P}^{\hat{\phi}^{H,m}}\left(\delta^m_{t+1}(h) < \epsilon_{t+1} \mid \delta^m_\tau(h) < \epsilon_\tau, \forall \tau \leq t]\right) \\
&= \prod_{t=t^*}^{\infty} P^m(\delta^m_{t+1}(h) < \epsilon_{t+1}) \\
&= \mathbf{P}^{\phi^*,m}(H_o), \qquad (24)
\end{aligned}$$

where the second line omits the term $\mathbf{P}^{\hat{\phi}^{H,m}}(\delta^m_t(h) < \epsilon_t \; \forall t \leq t^*)$, which equals 1 because $\epsilon_t \geq 4K$ for all $t \leq t^*$, and the third line follows from equation (8) and omits the term $\mathbf{P}^{\phi^*,m}(\delta^m_t(h) < \epsilon_t \; \forall t \leq t^*)$, which again equals 1 given the restriction on $(\epsilon_t)_{t\leq t^*}$. Moreover,

$$
\begin{aligned}
\mathbf{P}^{\phi^*,m}(H_o) &= \mathbf{P}^{\phi^*,m}\left(\delta_t^m(h) < \epsilon_t, \forall t\right) \\
&= \mathbf{P}^{\phi^*,m}\left(\cap_{t>t^*}\{\delta_t^m(h) < \epsilon_t\}\right) \\
&= \mathbf{P}^{\phi^*,m}\left(\cap_{n\in\{1,2,...\}} \cap_{\{t>t^*:N(t)=n\}} \{\delta_t^m(h) < n^{-1}\}\right) \\
&= 1 - \mathbf{P}^{\phi^*,m}\left(\cup_{n\in\{1,2,...\}} \cup_{\{t>t^*:N(t)=n\}} \{\delta_t^m(h) \geq n^{-1}\}\right) \\
&\geq 1 - \sum_{n=1}^{\infty} \sum_{\{t:N(t)=n\}} \mathbf{P}^{\phi^*,m}\left(\delta_t^m(h) \geq n^{-1}\right).
\end{aligned}
$$

In addition, by definition, if $N(t) = n$ then $t \geq \tau(n)$. Therefore,

$$
\begin{aligned}
\mathbf{P}^{\phi^*,m}(H_o) &\geq 1 - \sum_{n=1}^{\infty} \sum_{t\geq\tau(n)} \mathbf{P}^{\phi^*,m}\left(\delta_t^m(h) \geq n^{-1}\right) \\
&\geq 1 - \sum_{n=1}^{\infty} (3/2)4^{-n} = \frac{1}{2}, \qquad (25)
\end{aligned}
$$

where the last line follows from equation (7).

In addition,

$$
\begin{aligned}
\mathbf{P}^{\phi^*,m}\left(h \in H_o : \lim_{t\to\infty} ||\overline{P}_t(h) - P^m|| = 0\right) &= \mathbf{P}^{\phi^*,m}(H_0) \\
&\geq \frac{1}{2} > 0,
\end{aligned}
$$

where the first line follows from the SLLN and the second line from equation (25). The event $\{h \in H_o : \lim_{t\to\infty} ||\overline{P}_t(h) - P^m|| = 0\}$ is measurable with respect to the sigma algebra $\cap_t \sigma(z_t, z_{t+1}, ...)$ where $\sigma(z_t, z_{t+1}, ...)$ is the σ-algebra generated by $(z_t, z_{t+1}, ...)$. Since the events z_t are independent across t under the probability measure $\mathbf{P}^{\phi^*,m}$ and since $\mathbf{P}^{\phi^*,m}(H_0) > 0$, it then follows from the 0-1 Law (see Billingsley, 1995, Theorem 4.5) that $\mathbf{P}^{\phi^*,m}(H_0) = 1$. Hence, by equation (24), $\mathbf{P}^{\hat{\phi}^{H},m}(H_o) = 1$. $\square$

8.2 Proof of Lemma 2

If ξ^m is an m-stationary point, then $\xi^m_{\mathtt{ou}i}(s_i) = E_{P(\xi^m)}\left(\mathbf{1}_{Z^m_{\mathtt{ou}i}(s_i)}\right) = P(\alpha(\xi^m))(Z^m_{\mathtt{ou}i}(s_i))$. Hence, for each i, s_i

$$\alpha_i(s_i;\xi^m_i) = F_i\left(\frac{\sum_{\mathtt{u}\in U_{Ai}} P(\alpha(\xi^m))(Z^m_{A\mathtt{u}i}(s_i))\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Ai}} P(\alpha(\xi^m))(Z^m_{A\mathtt{u}i}(s_i))} - \frac{\sum_{\mathtt{u}\in U_{Bi}} P(\alpha(\xi^m))(Z^m_{B\mathtt{u}i}(s_i))\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Bi}} P(\alpha(\xi^m))(Z^m_{B\mathtt{u}i}(s_i))}\right).$$

The RHS equals $F_i(\Delta_i(P(\alpha(\xi^m)), s_i)$ and, thus, $\alpha(\xi^m)$ is an m-equilibrium.

If α^m is an m-equilibrium, then, for each i, s_i

$$\alpha^m_i(s_i) = F_i\left(\frac{\sum_{\mathtt{u}\in U_{Ai}} P(\alpha)(Z^m_{A\mathtt{u}i}(s_i))\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Ai}} P(\alpha)(Z^m_{A\mathtt{u}i}(s_i))} - \frac{\sum_{\mathtt{u}\in U_{Bi}} P(\alpha)(Z^m_{B\mathtt{u}i}(s_i))\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Bi}} P(\alpha)(Z^m_{B\mathtt{u}i}(s_i))}\right).$$

Since ξ^m is such that $\xi^m_{\mathtt{ou}i}(s_i) = E_{P(\alpha^m)}\left(\mathbf{1}_{Z^m_{\mathtt{ou}i}(s_i)}\right)$ then the RHS equals

$$F_i\left(\frac{\sum_{\mathtt{u}\in U_{Ai}} \xi^m_{Aui}(s_i)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Ai}} \xi^m_{Aui}(s_i)} - \frac{\sum_{\mathtt{u}\in U_{Bi}} \xi^m_{Bui}(s_i)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Bi}} \xi^m_{Bui}(s_i)}\right) = \alpha_i(s_i;\xi^m_i),$$

which implies that $\alpha^m = \alpha(\xi^m)$. Moreover,

$$\xi^m_{\mathtt{ou}i}(s_i) = E_{P(\alpha^m)}\left(\mathbf{1}_{Z^m_{\mathtt{ou}i}(s_i)}\right) = E_{P(\alpha(\xi^m))}\left(\mathbf{1}_{Z^m_{\mathtt{ou}i}(s_i)}\right). \square$$

8.3 Proof of Theorem 3

The proof relies on the following stochastic approximation result.

Theorem 4. *(Benaim (1996), Pemantle (1990)) Consider the following dynamical system*

$$\xi_{t+1} - \xi_t = \frac{1}{t}\left(L(\xi(t)) + \zeta_{t+1}\right)$$

with $\xi_t \in [0,1]^d$, and the ODE

$$\dot{\xi} = L(\xi)$$

with stationary point ξ_*. Suppose that the following conditions hold:

1. L is twice differentiable with continuous first derivative and bounded second derivative.

2. $(\zeta_{t+1})_t$ is a Martingale with respect to $\sigma(z_1, ..., z_{t-1})$ with $E_t[\zeta_{t+1}] = 0$ and $||\zeta_{t+1}|| \le C < \infty$.

3. There exists a $c > 0$ and a neighborhood $\mathcal{U}$ of ξ_* such that, for any unit vector Θ,

$$E\left(\max\{0, \langle \zeta_{t+1}, \Theta\rangle\}|\xi_t \in \mathcal{U}\right) \geq c.$$

Then:

(a) If ξ_* is a linearly stable stationary point of the ODE,

$$P(\lim_{t\to\infty} ||\xi_t - \xi_*|| = 0) > 0.$$

(c) If ξ_* is a linearly unstable stationary point of the ODE,

$$P(\lim_{t\to\infty} ||\xi_t - \xi_*|| = 0) = 0.$$

The proof of part (a) of Theorem 4 only requires conditions 1 and 2 to hold and follows from Benaim (1996). The proof of part (b) also requires condition 3 and follows from Pemantle (1990).

Proof of Theorem 3. We omit the superscript m to ease the notational burden. The proof has two steps. In the first step, we verify all the conditions for Theorem 4, which establishes results about the asymptotic behavior of beliefs. In the second step, we use continuity of strategies as a function of beliefs to establish the desired result.

Step 1. We first note that, by our assumptions over $(F_i)_i$, $\xi_t \geq c > 0$ for all t, and $\xi_t \leq 1$. Hence, we can focus on $\xi_t \in [c, 1]$. Condition 1 of Theorem 4 follows from simple but tedious algebra and is established in the Online Appendix. By definition, ζ_t (defined as a vector with components $\zeta_{o\mathbf{u}i\,t}(s_i)$) is a Martingale with respect to the sigma algebra generated by $(z_\tau)_{\tau \leq t-1}$. By construction, $||\zeta_t|| \leq 2$ a.s.. Also, by construction $||\xi_t|| \leq 1$ a.s.. So condition 2 of Theorem 4 is satisfied. Finally, we verify condition 3, i.e., there exists a $c > 0$ and a neighborhood $\mathcal{U}$ such that, for any unit vector Θ,

$$E\left(\max\{0, \langle \zeta_{t+1}, \Theta\rangle\}|\xi_t \in \mathcal{U}\right) \geq c.$$

We follow the arguments in Benaim and Hirsch (1999) pp. 30-31. First, we define $\mathbf{z} = (\mathbf{z}^1, ..., \mathbf{z}^n)$ with $\mathbf{z}^i$ belonging to the set of vertices of Δ^{d_i}, where $d_i = \#S_i \times \#U_{iA} \times \#U_{iB}$. So $\mathbf{z}$ is a $1 \times \sum_{i=1}^n d_i$ vector. Also, for each i, let $\mathbb{Z}_{oi} \equiv \cup_{\mathbf{u}\in U_{oi}} \cup_{s\in S_i} Z_{o\mathbf{u}i}(s)$, so $\mathbb{Z}_{oi}$ is the set of all possible values of $Z_{o\mathbf{u}i}(s)$ for a given (o, i); we also use z_{oi} to denote a typical element of $\mathbb{Z}_{oi}$. Finally, let $z \in \mathbb{Z} \equiv \cup_{o\in\{A,B\}} \times_{i=1}^{n} \mathbb{Z}_{oi}$. To every element z in $\mathbb{Z}$, we can assign a unique element $\mathbf{z}$, and we use $\mathbf{z}^{i}_{[o,\mathrm{u},s]}$ to denote the coordinate of $\mathbf{z}$ of the corresponding to $Z_{o\mathrm{u}i}(s)$. Let $\nu(\alpha(\xi))$ be the joint probability over $\mathbb{Z}$, i.e.,

$$\nu(\alpha(\xi))(z) = P_{\alpha(\xi)}\left((x, S, \omega) \colon o(x) = o,\ \forall i : x_{-i} \in piv_i, \omega \in u^{-1}(o, \mathrm{u}_i), s_i\right).$$

Note that,

$$\sum_{z_{-i}} \nu(\alpha(\xi))(z) = P(\alpha(\xi))(Z_{o\mathrm{u}_i i}(s_i)).$$

This expression can be casted as

$$P(\alpha(\xi))(Z_{o\mathrm{u}_i i}(s_i)) = \sum_{z\in\mathbb{Z}} \nu(\alpha(\xi))(z)\mathbf{z}^{i}_{[o,\mathrm{u},s]}.$$

In vector notation,

$$E_{P(\alpha(\xi))}\left(Z \mid \xi\right) = \sum_{z\in\mathbb{Z}} \nu(\alpha(\xi))(z)\mathbf{z}.$$

Finally, note that $\nu(\alpha(\xi))(z) > 0$ for each $z \in \mathbb{Z}$. This follows from the fact that, by assumption of F_i, $\alpha_i(s, \xi_i) \in (0, 1)$ for each (s, ξ_i) and $q_i(s|\omega) \in (0, 1)$ and $p(\omega) \in (0, 1)$ for each (s, ω); that is, our game is *diffuse* (see Benaim and Hirsch (1999) p. 30). Therefore, the same arguments in Benaim and Hirsch (1999) p. 31 can be applied to show

$$E\left(\max\{0, \langle\zeta_{t+1}, \Theta\rangle\}|\xi_t = \xi_*\right) \geq c.$$

Condition 3 follows by continuity of $E\left(\max\{0, \langle\zeta_{t+1}, \Theta\rangle\}|\xi_t = \cdot\right)$.

Step 2. In step 1, we verified that all the conditions for Theorem 4 are satisfied. Now, let α be an m-equilibrium and let $\xi_{o\mathrm{u}i}(s_i) = E_{P(\alpha^m)}\left(\mathbf{1}_{Z^m_{o\mathrm{u}i}(s_i)}\right)$ for all o, u, i, s_i. By Lemma 2, ξ is a stationary point. First, suppose that ξ is linearly stable. By Theorem 4,

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} ||\xi_t - \xi|| = 0\right) > 0.$$

By assumption A1 and the fact that $\xi_t \in [c, 1]$ (see Step 1), $\alpha_i(s_i; \cdot)$ is continuous, for all (i, s_i). Hence,

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} ||\alpha(\xi_t) - \alpha(\xi)|| = 0\right) > 0.$$

Moreover, by Lemma 2, $\alpha(\xi) = \alpha$.

Second, suppose that ξ is linearly unstable. By Theorem 4,

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} ||\xi_t - \xi|| = 0\right) = 0.$$

To show the desired result, we proceed by contradiction. Suppose that

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} ||\alpha(\xi_t) - \alpha|| = 0\right) > 0,$$

and note that, by Lemma 2, $\alpha(\xi) = \alpha$. Since $P(\cdot)(z)$ is continuous for any z (since z lives in a finite set, this holds uniformly),

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} \sup_z |P(\alpha(\xi_t))(z) - P(\alpha)(z)| = 0\right) > 0.$$

Now consider $h \in H_\epsilon \equiv \{h\colon \lim_{t\to\infty} \sup_z |P(\alpha(\xi_t))(z) - P(\alpha)(z)| < \epsilon\}$. By Lemma 3 in the Online Appendix, for all $h \in H_\epsilon \setminus \mathcal{Z}_\epsilon$ (where $\mathcal{Z}_\epsilon$ is a set of measure zero) and all $i, \mathtt{u}, o, s$:

$$\liminf_{t\to\infty} \frac{1}{t-1} \sum_{\tau=1}^{t-1} 1_{Z^m_{\mathrm{ou}i}(s)}(z_\tau) \geq P(\alpha)(Z^m_{\mathrm{ou}i}(s)) - \epsilon \tag{26}$$

and

$$\limsup_{t\to\infty} \frac{1}{t-1} \sum_{\tau=1}^{t-1} 1_{Z^m_{\mathrm{ou}i}(s)}(z_\tau) \leq P(\alpha)(Z^m_{\mathrm{ou}i}(s)) + \epsilon. \tag{27}$$

Taking ϵ in the rationals (this can be done WLOG), it follows that $\cup_{\epsilon>0}\mathcal{Z}_\epsilon$ also has measure zero and thus

$$\mathbf{P}^{\phi^{H,m}}\left(\cap_{\epsilon>0} H_\epsilon \setminus Z_\epsilon\right) \geq \mathbf{P}^{\phi^{H,m}}\left(\cap_{\epsilon>0} H_\epsilon\right) - \mathbf{P}^{\phi^{H,m}}\left(\cup_{\epsilon>0} Z_\epsilon\right) = \mathbf{P}^{\phi^{H,m}}\left(\cap_{\epsilon>0} H_\epsilon\right).$$

The RHS is positive since $\cap_{\epsilon>0} H_\epsilon = \lim_{t\to\infty} |P(\alpha(\xi_t))(Z^m_{\mathrm{ou}i}(s)) - P(\alpha)(Z^m_{\mathrm{ou}i}(s))| = 0$, and thus this implies that

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty} \left| \frac{1}{t-1} \sum_{\tau=1}^{t-1} 1\{z_\tau \in Z_{\mathrm{ou}i}(s)\} - P(\alpha)(Z_{\mathrm{ou}i}(s)) \right| = 0\right) > 0.$$

By definition of $Z_{\mathrm{ou}i}(s)$, $\frac{1}{t}\sum_{\tau=1}^{t-1} \mathbf{1}_{Z_{\mathrm{ou}i}(s)}(z_\tau) = \xi_{\mathrm{ou}i,t}(s)$ and

$P(\alpha)(Z_{\mathrm{ou}i}(s)) = E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z_{\mathrm{ou}i}(s)}\right) = \xi_{\mathrm{ou}i}(s)$ (the last equality follows from the fact that ξ is a stationary point). Therefore,

$$\mathbf{P}^{\phi^{H,m}}\left(\lim_{t\to\infty}|\xi_{o\mathrm{u}i,t}(s)-\xi_{o\mathrm{u}i}(s)|=0\right)>0$$

for all $o\mathrm{u}i$ and s, which contradicts that ξ is linearly unstable. □

8.4 Proof of Proposition 1

Step 1. Compute $E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z^{m,\eta}_{o\mathrm{u}i}(s_i)}\right)$: Define

$$\kappa_i^\eta(\omega;\xi_i)=\sum_{s_i\in S}q(s_i\mid\omega)\alpha_i^\eta(s_i;\xi_i)$$

to be the probability that i votes for A conditional on ω and beliefs ξ_i. Let $K_i^\eta(\omega;\xi_i)$ denote a random variable that is 1 with probability $\kappa_i^\eta(\omega;\xi_i)$ and 0 otherwise, and, for any set $N'\subset\{1,...,n\}$ let $K^\eta_{-N'}(\omega;\xi)=\sum_{i\notin N'}K_i^\eta(\omega;\xi_i)$. Also, let

$$\bar{\alpha}^\eta_{oi}(s_i;\xi_i)=\begin{cases}\alpha_i^\eta(s_i;\xi_i) & \text{if}\, o=A\\ 1-\alpha_i^\eta(s_i;\xi_i) & \text{if}\, o=B\end{cases}$$

Then $E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z^{m,\eta}_{o\mathrm{u}i}(s_i)}\right)=\sum_{\omega\in u_i^{-1}(o,\mathrm{u})}\beta^{m,\eta}_{oi}(\omega,s_i;\xi)\Pr(\omega,s_i)$, where, for Nash equilibrium,

$$\beta^{NE,\eta}_{oi}(\omega,s_i;\xi)=\bar{\alpha}^\eta_{oi}(s_i;\xi_i)\Pr\left(K^\eta_{-i}(\omega;\xi)=k-1\right)$$

and, for non-strategic equilibrium,

$$\begin{aligned}\beta^{N,\eta}_{Ai}(\omega,s_i;\xi)&=\alpha_i^\eta(s_i;\xi_i)\Pr\left(K^\eta_{-i}(\omega;\xi)\geq k-1\right)+(1-\alpha_i^\eta(s_i;\xi_i))\Pr\left(K^\eta_{-i}(\omega;\xi)\geq k\right)\\&=1-\beta^{N,\eta}_{Bi}(\omega,s_i;\xi).\end{aligned}$$

Step 2. Compute derivatives of $L^{m,\eta}_{o\mathrm{u}i}(\xi)(s_i)\equiv E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z^{m,\eta}_{o\mathrm{u}i}(s_i)}\right)-\xi_{o\mathrm{u}i}(s_i)$: Fix any η and any o,u,i,s_i and $\bar{o},\bar{\mathrm{u}},j,s_j,s_i'$ such that $j\neq i$, $s_i'\neq s_i$. Then, $\partial L^{m,\eta}_{o\mathrm{u}i}(\xi)(s_i)/\partial\xi_{\bar{o}\bar{\mathrm{u}}i}(s_i')=0$,

$$\frac{\partial L^{m,\eta}_{o\mathrm{u}i}(\xi)(s_i)}{\partial \xi_{\bar{o}\bar{\mathrm{u}}i}(s_i)} = (-1)^{1+1\{o=A\}} \frac{\partial \alpha^{\eta}_i(s_i;\xi_i)}{\partial \xi_{\bar{o}\bar{\mathrm{u}}i}(s_i)} \sum_{\omega \in u_i^{-1}(o,\mathrm{u})} \Pr\left(K^{\eta}_{-i}(\omega;\xi) = k-1\right) \Pr(\omega, s_i)$$

$$-1\{o=\bar{o}, \mathrm{u}=\bar{\mathrm{u}}\}, \tag{28}$$

$$\frac{\partial L^{NE,\eta}_{oui}(\xi)(s_i)}{\partial \xi_{\bar{o}\bar{\mathrm{u}}j}(s_j)} = \bar{\alpha}^{\eta}_{oi}(s_i;\xi_i) \frac{\partial \alpha^{\eta}_j(s_j;\xi_i)}{\partial \xi_{\bar{o}\bar{\mathrm{u}}j}(s_j)} \times$$
$$\times \sum_{\omega \in u_i^{-1}(o,\mathrm{u})} \left(\Pr\left(K^{\eta}_{-ij}(\omega;\xi) = k-2\right) - \Pr\left(K^{\eta}_{-ij}(\omega;\xi) = k-1\right)\right) q(s_j \mid \omega) \Pr(\omega, s_i),$$

and

$$\frac{\partial L^{N,\eta}_{o\mathrm{u}i}(\xi)(s_i)}{\partial \xi_{\bar{o}\bar{\mathrm{u}}j}(s_j)} = (-1)^{1+1\{o=A\}} \frac{\partial \alpha^{\eta}_j(s_j;\xi_i)}{\partial \xi_{\bar{o}\bar{\mathrm{u}}j}(s_j)} \sum_{\omega \in u_i^{-1}(o,\mathrm{u})} \Big\{ \alpha^{\eta}_i(s_i;\xi_i) \Pr\left(K^{\eta}_{-ij}(\omega;\xi) = k-2\right) +$$

$$+ (1-\alpha^{\eta}_i(s_i;\xi_i)) \Pr\left(K^{\eta}_{-ij}(\omega;\xi) = k-1\right) \Big\} q(s_j \mid \omega) \Pr(\omega, s_i). \tag{29}$$

In addition, $\frac{\partial \alpha^{\eta}_i(s_i;\xi_i)}{\partial \xi_{oui}(s_i)} = f^{\eta}\left(\mu_i(s_i;\xi_i)\right) \frac{\partial \mu_i(s_i;\xi_i)}{\partial \xi_{oui}(s_i)}$. Note that the entries of the matrix $J^m(\xi) = \frac{\partial L^m}{\partial \xi}$ are continuous in ξ.

Step 3. Limit beliefs: Let $(\xi^{m,\eta})_{\eta}$ be the sequence of beliefs corresponding to $(\alpha^{m,\eta})_{\eta}$, i.e. $\xi^{m,\eta}_{o\mathrm{u}i}(s_i) = E_{P(\alpha^{m,\eta})}\left(\mathbf{1}_{Z^m_{oui}(s_i)}\right)$, and denote its limit by ξ^m. Since $\alpha^m = \lim_{\eta\to 0} \alpha^{m,\eta}$ is a responsive equilibrium, then $\Delta^m_i(P(\alpha^m), s_i)$ is well-defined and, consequently, $\xi^m_{o\mathrm{u}i}(s_i) > 0$. Then $\lim_{\eta\to 0} \mu_i(s_i;\xi^{m,\eta}_i) = \Delta^m_i(P(\alpha^m), s)$ for all i, s. Consider first the case where α^m is a strict pure-strategy equilibrium, so that $\Delta^m_i(P(\alpha^m), s_i) \neq 0$. Then $\lim_{\eta\to 0} \mu_i(s_i;\xi^{m,\eta}_i) \neq 0$, and, by A5 (ii) and (iv), $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s_i;\xi^{m,\eta}_i)\right) = 0$. Finally, consider the case where α^{NE} is a symmetric mixed-strategy Nash equilibrium. Then there exists $s^* \in S$ such that $\alpha^{NE}(s^*) \in (0,1)$, $\Delta^{NE}_i(P(\alpha^{NE}), s^*) = 0$, and $\Delta^{NE}_i(P(\alpha^{NE}), s) \neq 0$ for all $s \neq s^*$. By previous arguments, $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s;\xi^{NE,\eta}_i)\right) = 0$ for all $s \neq s^*$. In addition, by A5(iv), $f^{\eta}_i\left(\mu_i(s^*;\xi^{NE,\eta}_i)\right) \leq f^{\eta}_i(0)$, and thus by A5(iii),

$$\lim_{\eta\to 0} f^{\eta}_i\left(\mu_i(s^*;\xi^{NE,\eta}_i)\right) f^{\eta}_i\left(\mu_i(s;\xi^{NE,\eta}_i)\right) = 0$$

for all $s \neq s^*$. We conclude this step by showing that $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s^*;\xi^{NE,\eta}_i)\right) = \infty$. Suppose not, so that $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s^*;\xi^{NE,\eta}_i)\right) = C < \infty$. Fix any i and any $\varepsilon > 0$. Fix a subsequence such that $\mu_i(s^*;\xi^{NE,\eta}_i) > 0$ for all η (a similar proof holds for negative subsequences). Then

$$1 - F^{\eta}\left(\mu_i(s^*;\xi_i^{NE,\eta})\right) = \int_{\mu_i(s^*;\xi_i^{NE,\eta})}^{\mu_i(s^*;\xi_i^{NE,\eta})+\varepsilon} f^{\eta}(x)dx + \int_{\mu_i(s^*;\xi_i^{NE,\eta})+\varepsilon}^{\infty} f^{\eta}(x)dx.$$

In addition, let $\epsilon^{\eta} \in [0,\varepsilon]$ be such that

$$\int_{\mu_i(s^*;\xi_i^{NE,\eta})}^{\mu_i(s^*;\xi_i^{NE,\eta})+\varepsilon} f^{\eta}(x)dx \leq \varepsilon f^{\eta}(\mu_i(s^*;\xi_i^{NE,\eta}) + \epsilon^{\eta})$$

for all η. By A5(iv), $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s^*;\xi_i^{NE,\eta}) + \epsilon^{\eta}\right) \leq \lim_{\eta\to 0} f^{\eta}\left(\mu_i(s^*;\xi_i^{NE,\eta})\right) = C$. Then, there exists η_{ε} such that, for all $\eta \leq \eta_{\varepsilon}$, $\int_{\mu_i(s^*;\xi_i^{NE,\eta})}^{\mu_i(s^*;\xi_i^{NE,\eta})+\varepsilon} f^{\eta}(x)dx \leq \varepsilon(C+\varepsilon)$ and, by A5(ii), $\int_{\mu_i(s^*;\xi_i^{NE,\eta})+\varepsilon}^{\infty} f^{\eta}(x)dx \leq \varepsilon$. Therefore, for all $\eta \leq \eta_{\varepsilon}$,

$$1 - F^{\eta}\left(\mu_i(s^*;\xi_i^{NE,\eta})\right) \leq \varepsilon(C+\varepsilon) + \varepsilon.$$

Since the above holds for any $\varepsilon > 0$, then $\alpha^{NE}(s^*) = \lim_{\eta\to 0} F^{\eta}\left(\mu_i(s^*;\xi_i^{NE,\eta})\right) = 1$, which contradicts the assumption that α^{NE} is a mixed-strategy profile.

Proof of part (i). Fix any $o, \mathtt{u}, i, s_i$. Since α^m is strict, Step 3 implies that $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s_i;\xi_i^{m,\eta})\right) = 0$. Therefore, $\lim_{\eta\to 0} J^{m,\eta}(\xi^{m,\eta})$ is a diagonal matrix with all 1's in the diagonal. Recall that the eigenvalues of a matrix J are the solutions $\{\lambda\}$ to $\det(J - \lambda I) = 0$, where I is the identity matrix. Then, the eigenvalues of $\lim_{\eta\to 0} J^{m,\eta}(\xi^{m,\eta})$ are all equal to -1. By continuity of the determinant function, it follows that the eigenvalues of $J^{m,\eta}(\xi^{m,\eta})$ are all strictly negative for all sufficiently small η.

Proof of part (ii). Let $U_o = \{u_o^1, ..., u_o^{r_o}\}$ be the utility range and $d \equiv n \times r_A \times r_B$. Then ξ is a vector of dimension dH, where H is the number of signals, and $J^{NE,\eta}(\xi)$ is an $dH \times dH$ matrix. Since α^{NE} is a symmetric mixed-strategy Nash equilibrium, Step 3 implies that $\lim_{\eta\to 0} f^{\eta}\left(\mu_i(s;\xi_i^{NE,\eta})\right) = 0$ for all $s \neq s^*$. Then, for all η sufficiently small, by continuity the eigenvalues of $J^{NE,\eta}(\xi^{NE,\eta})$ are sufficiently close to the eigenvalues of a matrix $J_*^{NE,\eta}$ defined by setting $f^{\eta}\left(\mu_i(s;\xi_i^{NE,\eta})\right) = 0$ and $f^{\eta}\left(\mu_i(s^*;\xi_i^{NE,\eta})\right) f^{\eta}\left(\mu_i(s;\xi_i^{NE,\eta})\right) = 0$ for all $s \neq s^*$ and replacing $\xi_i^{NE,\eta}$ with its limit ξ^{NE}, replacing $F^{\eta}\left(\mu_i(s;\xi_i^{NE,\eta})\right)$ with its limit $\alpha^{NE}(s)$, and replacing

$f^\eta\left(\mu_i(s^*;\xi_i^{NE,\eta})\right)$ with $f^\eta\left(\mu_i(s^*;\xi^{NE})\right)$, for all i, s. We now show that some eigenvalues of $J_*^{NE,\eta}$ are positive for all sufficiently small η, which implies that the eigenvalues of $J^{NE,\eta}(\xi^{NE,\eta})$ are positive for all sufficiently small η, and which, by Theorem 3, implies the desired result. Note that $J_*^{NE,\eta}$ depends on η only through f^η, since the derivatives are now evaluated at the limits ξ^{NE} and α^{NE}.

The matrix $J_*^{NE,\eta}$ can be decomposed into H^2 matrices of size $d \times d$, i.e., $M_{ss'}^\eta \equiv \frac{\partial L^{NE,\eta}(s)}{\partial \xi(s')}$ evaluated at ξ^{NE} and replacing $f^\eta\left(\mu_i(s;\xi_i^{NE,\eta})\right) = 0$ for all $s \neq s^*$. Note that $M_{ss'}^\eta = \mathbf{0}^{d\times d}$ is a zero matrix for all $s \neq s'$ and $s' \neq s^*$. In addition, M_{ss}^η is diagonal matrix with common element -1 in the diagonal for all $s \neq s^*$ and $M_{ss^*}^\eta$ has all zeroes in its diagonal for all $s \neq s^*$. Therefore, using the fact that the entries of the matrix do not depend on i, it is straightforward to check that $\det\left(J_*^{NE,\eta} - \lambda I^{dH\times dH}\right) = 0$ if and only if either $\lambda = -1$ or

$$\det\left(M_{s^*s^*}^\eta - \lambda I^{d\times d}\right) = 0. \tag{30}$$

We conclude by showing that there exist solutions $\{\lambda\}$ to (30) that go to ∞ as $\eta \to 0$. Let $M^\eta(\lambda) \equiv M_{s^*s^*}^\eta - \lambda I^{d\times d}$. We decompose M^η into $r_A \times r_B$ matrices of size $n \times n$, i.e., $M_{\mathtt{u_o}\bar{\mathtt{u}}_{\bar{\mathtt{o}}}}^\eta \equiv \left.\frac{\partial L_{o\mathtt{u_o}}^{NE,\eta}(s^*)}{\partial \xi_{\bar{o}\bar{u}_{\bar{o}}}(s^*)}\right|_{\xi^{NE}}$. We can write M^η as a block matrix

$$M^\eta(\lambda) = \begin{bmatrix} \bar{M}_A^\eta & \bar{M}_{AB}^\eta \\ \bar{M}_{BA}^\eta & \bar{M}_B^\eta \end{bmatrix}, \tag{31}$$

where $\bar{M}_o^\eta = (M_{\mathtt{u_o}\bar{\mathtt{u}}_\mathtt{o}}^\eta)_{u_o,\bar{u}_o\in U_A} - \lambda I^{nr_o\times nr_o}$ is of size $nr_o \times nr_o$, $\bar{M}_{AB}^\eta = (M_{\mathtt{u_A}\bar{\mathtt{u}}_\mathtt{B}}^\eta)_{u_A\in U_A,\bar{u}_B\in U_B}$ is of size $nr_A \times nr_B$, and $\bar{M}_{BA}^\eta = (M_{\mathtt{u_B}\bar{\mathtt{u}}_\mathtt{A}}^\eta)_{u_B\in U_B,\bar{u}_A\in U_A}$ is of size $nr_B \times nr_A$. By symmetry of ξ^{NE}, each matrix $M_{\mathtt{u_o}\bar{\mathtt{u}}_{\bar{\mathtt{o}}}}^\eta$ has a common diagonal and a common off-diagonal element; it is then straightforward to check that the matrices commute, i.e., $M_{\mathtt{u_o}\bar{\mathtt{u}}_{\bar{\mathtt{o}}}}^\eta = M_{\mathtt{u}'_o\bar{\mathtt{u}}'_{\bar{o}}}^\eta$ for all $\mathtt{u}_o, \bar{\mathtt{u}}_{\bar{o}}, \mathtt{u}'_o\bar{\mathtt{u}}'_{\bar{o}}$.

From this point onwards, the proof relies on algebra and exploits the fact that the state space is binary. By A3, there are two cases to consider: In case (1), none of the alternatives yields a state-independent payoff; in case (2), one of the alternatives, say B, yields a state-independent payoff. In the Online Appendix, we establish that, for the case $\lambda \neq -1$, $\det M^\eta(\lambda) = 0$ if and only if

$$\left(1 + \lambda + (n-1)a_j^\eta\right)\left(1 + \lambda - a_j^\eta\right)^{n-1} = 0, \tag{32}$$

where $j = 1, 2$ represents the case (1) or (2), respectively, and $a_j^\eta = f^\eta\left(\mu_i(s^*; \xi^{NE})\right) a_j$, where $a_j \neq 0$. By Step 3, $f^\eta\left(\mu_i(s^*; \xi^{NE})\right)$ is arbitrarily large for sufficiently small η; therefore, since $a_j \neq 0$, there exists a solution to (32) that is positive for all sufficiently small η. $\square$

References

Al-Najjar, N., "Decision Makers as Statisticians: Diversity, Ambiguity and Learning," *Econometrica*, 2009, *77* (5), 1371–1401.

_ **and M. Pai**, "Coarse Decision Making," 2009.

Ali, S.N., J.K. Goeree, N. Kartik, and T.R. Palfrey, "Information aggregation in standing and ad hoc committees," *The American Economic Review*, 2008, *98* (2), 181–186.

Aragones, E., I. Gilboa, A. Postlewaite, and D. Schmeidler, "Fact-Free Learning," *American Economic Review*, 2005, *95* (5), 1355–1368.

Austen-Smith, D. and J.S. Banks, "Information aggregation, rationality, and the Condorcet jury theorem," *American Political Science Review*, 1996, pp. 34–45.

Battaglini, M., R.B. Morton, and T.R. Palfrey, "The Swing Voter's Curse in the Laboratory," *Review of Economic Studies*, 2010, *77* (1), 61–89.

Battigalli, P., *Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali*, Vol. 100, Universita Bocconi, Milano, 1987.

Benaim, M. and M.W. Hirsch, "Mixed equilibria and dynamical systems arising from fictitious play in perturbed games," *Games and Economic Behavior*, 1999, *29* (1-2), 36–72.

Billingsley, P., *Probability and Measure*, Wiley, 1995.

Bolton, P. and C. Harris, "Strategic experimentation," *Econometrica*, 1999, *67* (2), 349–374.

Charness, G. and D. Levin, "The origin of the winner's curse: a laboratory study," *American Economic Journal: Microeconomics*, 2009, *1* (1), 207–236.

Costinot, A. and N. Kartik, "On Optimal Voting Rules under Homogeneous Preferences," 2007.

Dekel, E., D. Fudenberg, and D.K. Levine, "Learning to play Bayesian games," *Games and Economic Behavior*, 2004, *46* (2), 282–303.

Erev, I. and A.E. Roth, "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria," *American economic review*, 1998, pp. 848–881.

Esponda, I., "Behavioral equilibrium in economies with adverse selection," *The American Economic Review*, 2008, *98* (4), 1269–1291.

_ , "Equilibrium refinement in elections with private information," *working paper*, 2011.

_ **and D. Pouzo**, "Conditional retrospective voting in large elections," *working paper*, 2012.

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

Eyster, E. and M. Rabin, "Cursed equilibrium," *Econometrica*, 2005, *73* (5), 1623–1672.

Feddersen, T. and W. Pesendorfer, "Voting behavior and information aggregation in elections with private information," *Econometrica*, 1997, pp. 1029–1058.

_ **and** _ , "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting," *American Political Science Review*, 1998, pp. 23–35.

Feddersen, T.J., "Rational choice theory and the paradox of not voting," *Journal of Economic Perspectives*, 2004, pp. 99–112.

Fiorina, M.P., "Economic retrospective voting in American national elections: A micro-analysis," *American Journal of Political Science*, 1978, pp. 426–443.

Friedman, J., *The rational choice controversy: Economic models of politics reconsidered* number 1-2, Yale Univ Pr, 1996.

Fudenberg, D. and D.K. Levine, "Self-confirming equilibrium," *Econometrica*, 1993, pp. 523–545.

_ **and** _ , *The theory of learning in games*, Vol. 2, The MIT press, 1998.

_ **and** _ , "Learning and Equilibrium," *Annual Review of Economics*, 2009, *1*, 385–420.

_ **and D.M. Kreps**, "Learning in extensive form games, II: Experimentation and Nash equilibrium," Technical Report, mimeo 1994.

_ **and** _ , "Learning in extensive-form games I. Self-confirming equilibria," *Games and Economic Behavior*, 1995, *8* (1), 20–55.

_ **and S. Takahashi**, "Heterogeneous beliefs and local information in stochastic fictitious play," *Games and Economic Behavior*, 2011, *71* (1), 100–120.

Green, D.P. and I. Shapiro, *Pathologies of rational choice theory: A critique of applications in political science*, Cambridge Univ Press, 1994.

Guarnaschelli, S., R.D. McKelvey, and T.R. Palfrey, "An experimental study of jury decision rules," *American Political Science Review*, 2000, pp. 407–423.

Harsanyi, J.C., "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points," *International Journal of Game Theory*, 1973, *2* (1), 1–23.

Jehiel, P. and F. Koessler, "Revisiting games of incomplete information with analogy-based expectations," *Games and Economic Behavior*, 2008, *62* (2), 533–557.

Kagel, J.H. and D. Levin, *Common value auctions and the winner's curse*, Princeton Univ Pr, 2002.

Keller, G., S. Rady, and M. Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, 2005, *73* (1), 39–68.

Key, V.O., *The responsible electorate*, Vintage Books New York, 1966.

Kramer, G.H., "Short-term fluctuations in US voting behavior, 1896-1964," *The American Political Science Review*, 1971, pp. 131–143.

Marcet, A. and T.J. Sargent, "Convergence of least squares learning mechanisms in self-referential linear stochastic models," *Journal of Economic theory*, 1989, *48* (2), 337–368.

McLennan, A., "Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents," *American Political Science Review*, 1998, pp. 413–418.

Morris, S., "Purification," *The New Palgrave Dictionary of Economics*, 2008, pp. 779–782.

Nagel, R., "Unraveling in guessing games: An experimental study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Osborne, M.J. and A. Rubinstein, "Sampling equilibrium, with an application to strategic voting," *Games and Economic Behavior*, 2003, *45*, 434–441.

Schwartzstein, J., "Selective Attention and Learning," *Unpublished Manuscript, Harvard University*, 2009.

Selten, R., "Reexamination of the perfectness concept for equilibrium points in extensive games," *International journal of game theory*, 1975, *4* (1), 25–55.

Stahl, D.O. and P.W. Wilson, "On Players? Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 1995, *10* (1), 218–254.

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

1 Online Appendix

This appendix contains additional proofs for "Learning foundation and equilibrium selection in voting environments with private information," by Ignacio Esponda and Demian Pouzo.

1.1 Lemmas 3 and 4

The statement and proof of the next lemma are straightforward adaptations of a result by Fudenberg and Kreps (1993).

Lemma 3. *(cf. Fudenberg and Kreps, Lemma 6.2, 1993) Let $(z_t)_t$ be a sequence of random variables with range on a finite set Z. Fix a set-function $\pi : 2^Z \to [0,1]$ (not necessarily a probability measure) and fix $\varepsilon \in \mathbb{R}$. Let H_ε be a subset of infinite histories such that for all $h \in H_\varepsilon$ there exists $t_{\varepsilon,h}$ such that for all $t \geq t_{\varepsilon,h}$, the distribution of each z_t conditional on $h^t = (z_1, ..., z_{t-1})$, denoted $\pi_t(\cdot \mid h^t)$, satisfies*

$$\max_{Z' \in \mathcal{Z}} \pi_t(Z') - \pi(Z') > -\varepsilon, \tag{33}$$

where $\mathcal{Z} \subset 2^Z$ is a set of subsets of Z.[32]

Then

$$\liminf_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} \mathbf{1}_{Z'}(z_\tau) \geq \pi(Z') - \varepsilon \tag{34}$$

for all $Z' \in \mathcal{Z}$, almost surely on H_ε. Moreover, if (33) is replaced by $\max_{Z' \in \mathcal{Z}} \pi_t(Z') - \pi(Z') < \varepsilon$, then the conclusion in (34) is replaced by $\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} \mathbf{1}_{Z'}(z_\tau) \leq \pi(Z') + \varepsilon$.

Proof. First note that $\#Z < \infty$ and thus any subset of $\mathcal{Z} \subset 2^Z$ has also finitely many elements. Therefore, it suffices to show the result for any (arbitrary) subset $Z' \in \mathcal{Z}$ since there are only finitely many of them (roughly speaking, Z' is what a is in Fudenberg and Kreps, 1993). Since Z is finite we can order the elements as $(z_1, ...z_{\#Z})$, and WLOG we set the first $\#Z'$ to be the elements of Z'. Just as FK 93, let $(\omega_t)_t$ be an independent sequence of uniform random variables and let $y_t : \Omega \to Z$ be a new random variable.

[32] If H_ε has zero probability, the lemma is taken to be vacuous.

As in FK 93, we construct $(y_t(\omega_t))_t$ as follows. For $t = 1$, $y_1(\omega_1) = z_m$ iff $\sum_{n=1}^{m-1} \pi_1(z_n) \leq \omega_1 < \sum_{n=1}^{m} \pi_1(z_n)$. For $t = \tau$, let $y_\tau(\omega_\tau) = z_m$ iff $\sum_{n=1}^{m-1} \pi_\tau(z_n|y_1, ..., y_{\tau-1}) \leq \omega_\tau < \sum_{n=1}^{m} \pi_\tau(z_n|y_1, ..., y_{\tau-1})$. Moreover, by construction the probability over h^t coincides with the probability over $(\omega_\tau)_{\tau\leq t}$; we thus can use both interchangeably. In particular, the set of ω for which $y_t(\omega_t) \in Z'$ is the set of $\{\omega : \omega_t \leq \sum_{n=1}^{\#Z'} \pi_t(z_n|y_1, ..., y_{t-1}) = \pi_t(Z'|y_1, ..., y_{t-1})\}$ (recall that Z' consists of the first $\#Z'$ elements in Z).

Under equation 33 the latter set includes the set $\{\omega : \omega_t \leq \pi(Z') - \varepsilon\}$; thus $\mathbf{1}_{\{\omega:\omega_t\leq\pi(Z')-\varepsilon\}} \leq \mathbf{1}_{\{\omega:\omega_t\leq\pi_t(Z'|y_1,...,y_{t-1})\}} = \mathbf{1}_{\{\omega:y_t(\omega_t)\in Z'\}} = \mathbf{1}_{\{z_\tau\in Z'\}}$.

Let $\nu_t(r, \omega)$ be the number of times $\omega_t \leq r$. Then $\nu_t(\pi(Z') - \varepsilon) \leq \sum_{\tau=1}^{t} \mathbf{1}_{Z'}(z_\tau)$. By the strong law of large numbers, $\lim_{t\to\infty} \nu_t(\pi(Z')-\varepsilon) = \pi(Z')-\varepsilon$ $\mathbf{P}^{\phi^{H,m}}$-a.s. on H_ε (this is under the measure $(\omega_t)_t$, which by construction is equal to $\mathbf{P}^{\phi^{H,m}}$). Therefore it must follow that $\liminf_{t\to\infty} t^{-1} \sum_{\tau=1}^{t} \mathbf{1}_{Z'}(z_\tau) \geq \pi(Z') - \varepsilon$.

Similarly, under equation (33), the set $\{\omega : \omega_t \leq \pi_t(Z'|y_1, ..., y_{t-1})\}$ is included in the set $\{\omega : \omega_t \leq \pi(Z')+\varepsilon\}$. By a similar argument as before, $\limsup_{t\to\infty} t^{-1} \sum_{\tau=1}^{t} \mathbf{1}_{Z'}(z_\tau) \leq \pi(Z') + \varepsilon$. □

Lemma 4. *Fix $m \in \{N, NE\}$. There exists H' with $\mathbf{P}^{\phi^{H,m}}(H') = 1$, such that for all $\varrho > 0$ and for all $h \in H'$ there exists $t_{\varrho,h}$ such that for all $t \geq t_{\varrho,h}$ and all $o \in \{A, B\}$, $\overline{P}_t(h)(Z^m_{oi}(s_i)) > K_p - \varrho$, where*

$$K_p \equiv \frac{1}{2} \min_{i,s_i} \left\{ \Psi_i \times \sum_{\omega\in\Omega} q_i(s_i \mid \omega)G(\omega) \right\} > 0, \tag{35}$$

and, for all i, $\Psi_i \equiv min \Big\{ \prod_{j=1}^{n} F_j(-2K), \prod_{j=1}^{n} (1 - F_j(2K)), \prod_{l\in\widetilde{N}_i} (1 - F_l(2K)) \cdot \cdot \prod_{j\in N_i\cup\{i\}} F_j(-2K), \prod_{l\in\widetilde{N}_i} F_l(-2K) \cdot \prod_{j\in N_i\cup\{i\}} (1 - F_j(2K)) \Big\} > 0$, where $N_i \cup \widetilde{N}_i \cup \{i\} = \{1, ..., n\}$, $\#N_i = k - 1$, and $\#\widetilde{N}_i = n - k$.

Proof. By definition of assessment rules, $\mu^m_{it}(h)(s_i) \in [-K, K]$, and, therefore, $\alpha^H_{it}(h)(s_i) \in [F_i(-2K), F_i(2K)]$ for all i, s_i, for all h, and for all t. Hence, for all i, s_i, for all h, and for all t,

$$\mathbf{P}^{\phi^{H,N}}_t \left(z_t \in Z^N_{Ai}(s_i) \mid h^t\right) \geq \prod_{j=1}^{n} F_j(-2K) \sum_{\omega\in\Omega} q_i(s_i \mid \omega)G(\omega) > K_p,$$

and, similarly, $\mathbf{P}^{\phi^{H,N}}_t \left(z_t \in Z^N_{Bi}(s_i) \mid h^t\right) > K_p$.

Similarly, for all i, s_i, for all h, and for all t,

$$\mathbf{P}_t^{\phi^{H,NE}}\left(z_t \in Z_{Ai}^{NE}(s_i) \mid h^t\right) \geq \prod_{l \in \widetilde{N}_i} (1 - F_l(2K)) \cdot \prod_{j \in N_i \cup \{i\}} F_j(-2K) \sum_{\omega \in \Omega} q_i(s_i \mid \omega) G(\omega) > K_p,$$

and, similarly, $\mathbf{P}_t^{\phi^{H,NE}}\left(z_t \in Z_{Bi}^{NE}(s_i) \mid h^t\right) > K_p$.

An application of Lemma 3 (by setting $\pi(Z_{Ai}^m(s_i)) = K_p$—the case of $Z_{Bi}^m(s_i)$ is analogous and thus omitted—$\varepsilon = 0$, and $H_\varepsilon = H$) implies that $\liminf_{t\to\infty} \overline{P}_t(h)(Z_{oi}^m(s_i)) \geq K_p$ $\mathbf{P}^{\phi^{H,m}}$- a.s. on H. Therefore, there exists a $H' \subseteq H$ with $\mathbf{P}^{\phi^H}(H') = 1$ such that for all $\varrho > 0$ and all $h \in H'$, there exists a $t_{\varrho,h}$ such that for all $t \geq t_{\varrho,h}$, $\overline{P}_t(h)(Z_{oi}^m(s_i)) > K_p - \varrho$. □

1.2 Proof of equation (32)

We drop the player subscripts and the superscript NE from ξ^{NE} and α^{NE} for simplicity. First, consider case (1), where payoffs are state-dependent for both alternatives. Given that payoffs are uniquely identified by the outcome and state, we use the simpler subscript notation $o\omega$ rather than ou_o. Let $z_o \equiv \sum_{s\in S} q(s \mid \omega_o)\alpha(s)$ and $piv_o \equiv \binom{n-1}{k-1} z_o^{k-1}(1-z_o)^{n-k} \Pr(\omega_o, s^*)$. Since ξ is a steady-state, then

$$\begin{aligned} \alpha(s^*) piv_o &= \xi_{A\omega_o} \\ (1-\alpha(s^*)) piv_o &= \xi_{B\omega_o}, \end{aligned} \tag{36}$$

and, therefore,

$$\frac{piv_B}{piv_A} = \frac{\xi_{A\omega_B}}{\xi_{A\omega_A}} = \frac{\xi_{B\omega_B}}{\xi_{B\omega_A}}.$$

Let $\rho \equiv \frac{\xi_{A\omega_B}}{\xi_{A\omega_A}} = \frac{\xi_{B\omega_B}}{\xi_{B\omega_A}}$, and note that $\rho > 0$. Then $piv_A = (1/\rho) piv_B$. In addition, $\frac{\partial \mu(s^*;\xi)}{\partial \xi_{A\omega_A}(s^*)} = -\rho \frac{\partial \mu(s^*;\xi)}{\partial \xi_{A\omega_B}(s^*)}$ and $\frac{\partial \mu(s^*;\xi)}{\partial \xi_{B\omega_A}(s^*)} = -\rho \frac{\partial \mu(s^*;\xi)}{\partial \xi_{B\omega_B}(s^*)}$, where

$$\begin{aligned} \frac{\partial \mu(s^*;\xi)}{\partial \xi_{A\omega_B}(s^*)} &= -\frac{\xi_{A\omega_A}\left(u(A,\omega_A) - u(A,\omega_B)\right)}{\left(\xi_{A\omega_A} + \xi_{A\omega_B}\right)^2} \\ \frac{\partial \mu(s^*;\xi)}{\partial \xi_{B\omega_B}(s^*)} &= -\frac{\xi_{B\omega_A}\left(u(B,\omega_B) - u(B,\omega_A)\right)}{\left(\xi_{B\omega_A} + \xi_{B\omega_B}\right)^2}. \end{aligned} \tag{37}$$

Let

$$r_o^\eta \equiv f^\eta\left(\mu(s^*;\xi)\right)\frac{\partial\mu(s^*;\xi)}{\partial\xi_{o\omega_B}(s^*)}piv_o$$

and

$$y_o \equiv q(s^* \mid \omega_o)\left(\frac{k-1}{n-1}\frac{1}{z_o}-\frac{k-1}{n-1}\frac{1}{1-z_o}\right).$$

Finally, let I be the $n\times n$ identity matrix and let I^* be the $n\times n$ matrix with all zeroes in the diagonal and all ones outside the diagonal. Then, using the above notation to simplify the derivatives in Step 2 of Proposition 1, it follows that the matrix in equation (31) can be decomposed as follows:

$$\bar{M}_A^\eta = \begin{bmatrix} M^\eta_{A\omega_A A\omega_A} & M^\eta_{A\omega_A A\omega_B} \\ M^\eta_{A\omega_B A\omega_A} & M^\eta_{A\omega_B A\omega_B}\end{bmatrix} \quad \bar{M}_{AB}^\eta = \begin{bmatrix} M^\eta_{A\omega_A B\omega_A} & M^\eta_{A\omega_A B\omega_B} \\ M^\eta_{A\omega_B B\omega_A} & M^\eta_{A\omega_B B\omega_B}\end{bmatrix}$$

$$\bar{M}_{BA}^\eta = \begin{bmatrix} M^\eta_{B\omega_A A\omega_A} & M^\eta_{B\omega_A A\omega_B} \\ M^\eta_{B\omega_B A\omega_A} & M^\eta_{B\omega_B A\omega_B}\end{bmatrix} \quad \bar{M}_B^\eta = \begin{bmatrix} M^\eta_{B\omega_A B\omega_A} & M^\eta_{B\omega_A B\omega_B} \\ M^\eta_{B\omega_B B\omega_A} & M^\eta_{B\omega_B B\omega_B}\end{bmatrix},$$

where

$$
\begin{aligned}
M^\eta_{A\omega_A A\omega_A} &= -\left(r^\eta_A + 1 + \lambda\right) I - \alpha(s^*) r^\eta_A y_A I^* \\
M^\eta_{A\omega_A A\omega_B} &= (1/\rho) r^\eta_A I + \alpha(s^*)(1/\rho) r^\eta_A y_A I^* \\
M^\eta_{A\omega_B A\omega_A} &= -\rho r^\eta_A I - \alpha(s^*)\rho r^\eta_A y_B I^* \\
M^\eta_{A\omega_B A\omega_B} &= (r^\eta_A - (1+\lambda)) I - \alpha(s^*) r^\eta_A y_B I^* \\
M^\eta_{A\omega_A B\omega_A} &= -r^\eta_B I - \alpha(s^*) r^\eta_B y_A I^* \\
M^\eta_{A\omega_A B\omega_B} &= (1/\rho) r^\eta_B I + \alpha(s^*)(1/\rho) r^\eta_B y_A I^* \\
M^\eta_{A\omega_B B\omega_A} &= -\rho r^\eta_B I - \alpha(s^*)\rho r^\eta_B y_B I^* \\
M^\eta_{A\omega_B B\omega_B} &= r^\eta_B I + \alpha(s^*) r^\eta_B y_B I^* \\
M^\eta_{B\omega_A A\omega_A} &= r^\eta_A I - (1-\alpha(s^*)) r^\eta_A y_A I^* \\
M^\eta_{B\omega_A A\omega_B} &= -(1/\rho) r^\eta_A I + (1/\rho)(1-\alpha(s^*)) r^\eta_A y_A I^* \\
M^\eta_{B\omega_B A\omega_A} &= \rho r^\eta_A I - \rho(1-\alpha(s^*)) r^\eta_A y_B I^* \\
M^\eta_{B\omega_B A\omega_B} &= -r^\eta_A I + (1-\alpha(s^*)) r^\eta_A y_B I^* \\
M^\eta_{B\omega_A B\omega_A} &= (r^\eta_B - (1+\lambda)) I - (1-\alpha(s^*)) r^\eta_B y_A I^* \\
M^\eta_{B\omega_A B\omega_B} &= -(1/\rho) r^\eta_B I + (1/\rho)(1-\alpha(s^*)) r^\eta_B y_A I^* \\
M^\eta_{B\omega_B B\omega_A} &= \rho r^\eta_B I - \rho(1-\alpha(s^*)) r^\eta_B y_B I^* \\
M^\eta_{B\omega_B B\omega_B} &= -(r^\eta_B + 1 + \lambda) I + (1-\alpha(s^*)) r^\eta_B y_B I^*.
\end{aligned}
$$

It is easy to check that all of the above matrices commute. Therefore, $\det M^\eta = \det\left(\bar{M}^\eta_A \bar{M}^\eta_B - \bar{M}^\eta_{AB}\bar{M}^\eta_{BA}\right)$. By simple algebra,

$$
\bar{M}^\eta_A \bar{M}^\eta_B - \bar{M}^\eta_{AB}\bar{M}^\eta_{BA} = \begin{bmatrix} M^\eta_1 & M^\eta_2 \\ M^\eta_3 & M^\eta_4 \end{bmatrix},
$$

where

$$
\begin{aligned}
M^\eta_1 &= (1+\lambda)\left(r^\eta_A - r^\eta_B + 1 + \lambda\right) I + z_A I^* \\
M^\eta_2 &= (1/\rho)(1+\lambda)\left(r^\eta_A - r^\eta_B\right) I - (1/\rho) z_A I^* \\
M^\eta_3 &= \rho(1+\lambda)\left(r^\eta_A - r^\eta_B\right) I + \rho z_B I^* \\
M^\eta_4 &= (1+\lambda)\left(r^\eta_B - r^\eta_A + 1 + \lambda\right) I - z_B I^*
\end{aligned}
$$

and

$$z_o = (1+\lambda)y_o\left(\alpha(s^*)r_A^\eta + (1-\alpha(s^*))r_B^\eta\right).$$

Once again, the above matrices commute. Therefore, $\det\left(\bar{M}_A^\eta \bar{M}_B^\eta - \bar{M}_{AB}^\eta \bar{M}_{BA}^\eta\right) = \det\left(M_1 M_4 - M_2 M_3\right)$. By simple algebra,

$$M_1^\eta M_4^\eta - M_2^\eta M_3^\eta = (1+\lambda)^3 \times \left((1+\lambda)I + \left(\alpha(s^*)r_A^\eta + (1-\alpha(s^*))r_B^\eta\right)(y_A - y_B)\, I^*.\right. \tag{38}$$

To compute the determinant of (38), we perform the following operations on that matrix, which do not affect the determinant: First, we add columns 2, ..., n to column 1. Second, we subtract row 1 from each of the rows 2,...,n. The result is a triangular matrix, and, therefore, its determinant is obtained by multiplying the terms in the diagonal:

$$\det\left(M_1^\eta M_4^\eta - M_2^\eta M_3^\eta\right) = (1+\lambda)^{3n} \times \left((1+\lambda) + (n-1)a_1^\eta\right)\left(1+\lambda - a_1^\eta\right)^{n-1},$$

where

$$\begin{aligned} a_1^\eta &= \left(\alpha(s^*)r_A + (1-\alpha(s^*))r_B\right)(y_A - y_B) \\ &= f^\eta\left(\mu(s^*;\xi)\right) piv_B\left(\alpha(s^*)\frac{\partial\mu(s^*;\xi)}{\partial\xi_{A\omega_B}(s^*)} + (1-\alpha(s^*))\frac{\partial\mu(s^*;\xi)}{\partial\xi_{B\omega_B}(s^*)}\right)(y_A - y_B) \\ &\equiv f^\eta\left(\mu(s^*;\xi)\right) \times a_1. \end{aligned}$$

The proof of case (1) concludes by showing that $a_1 \neq 0$. Clearly, $piv_B \neq 0$ because the equilibrium is responsive. In addition, using (36) and (37), $\alpha(s^*)\frac{\partial\mu(s^*;\xi)}{\partial\xi_{A\omega_B}(s^*)} + (1-\alpha(s^*))\frac{\partial\mu(s^*;\xi)}{\partial\xi_{B\omega_B}(s^*)}$ equals

$$-\frac{(u(A,\omega_A) - u(B,\omega_A) + u(B,\omega_B) - u(A,\omega_B))}{(1+\rho)^2 piv_A} < 0,$$

where the inequality follows by A3. Finally, A4 (MLRP) implies that

$$\begin{aligned} \frac{\Pr\left(S > s^* \mid \omega_A\right)}{q(s^* \mid \omega_A)} &\geq \frac{\Pr\left(S > s^* \mid \omega_B\right)}{q(s^* \mid \omega_B)} \\ \frac{\Pr\left(S < s^* \mid \omega_A\right)}{q(s^* \mid \omega_A)} &\leq \frac{\Pr\left(S < s^* \mid \omega_B\right)}{q(s^* \mid \omega_B)}, \end{aligned}$$

with at least one of the above inequalities being strict. Then

$$\frac{q(s^* \mid \omega_A)}{z_A} \leq \frac{q(s^* \mid \omega_B)}{z_B}$$
$$\frac{q(s^* \mid \omega_A)}{1 - z_A} \geq \frac{q(s^* \mid \omega_B)}{1 - z_B},$$

with at least one of the above inequalities being strict. But then $y_A - y_B < 0$ is different from zero.

For case (2), one can ignore learning about outcome B and the relevant determinant becomes the determinant of $\bar{M}_A^\eta$. By simple algebra, $M^\eta_{A\omega_A A\omega_A} M^\eta_{A\omega_B A\omega_B} - M^\eta_{A\omega_A A\omega_B} M^\eta_{A\omega_B A\omega_A}$ equals

$$(1+\lambda) \times \left((1+\lambda)I - \alpha(s^*) r_A^\eta \left(y_A - y_B\right) I^*\right),$$

and, following the same steps as for case (1),

$$\det \bar{M}_A^\eta = f^\eta\left(\mu(s^*;\xi)\right) \times a_2,$$

where

$$\begin{aligned} a_2 &= -piv_B \alpha(s^*) \frac{\partial \mu(s^*;\xi)}{\partial \xi_{A\omega_B}(s^*)} \left(y_A - y_B\right) \\ &= piv_B \alpha(s^*) \frac{\xi_{A\omega_A}\left(u(A,\omega_A) - u(A,\omega_B)\right)}{\left(\xi_{A\omega_A} + \xi_{A\omega_B}\right)^2} \left(y_A - y_B\right). \end{aligned}$$

The fact that $a_2 \neq 0$ follows from previous arguments and the facts that $\alpha(s^*) \in (0,1)$, $\xi_{A\omega} > 0$ (because the equilibrium is responsive and ξ is determined by (36)), and $u(A,\omega_A) > u(A,\omega_B)$ (which is implied by A3 given that, in this case, $u(B,\omega_A) = u(B,\omega_B)$). □

1.3 Examples: stability of mixed-strategy non-strategic equilibrium

In this section we document that mixed-strategy non-strategic equilibrium can be asymptotically stable or unstable depending on the primitives. We consider two examples, that differ only in the payoffs and satisfy A2-A5 . Let $n = 3, k = 2$;

$p(\omega_A) = p(\omega_B) = .5$; $q(s_H \mid \omega_A) = q(s_L \mid \omega_B) = 2/3$; $F^{\eta}(v) = \frac{1}{1+e^{-\frac{1}{\eta}v}}$.

Example I. $u(A, \omega_A) = 1 > 0 = u(B, \omega_A); u(A, \omega_B) = -1 < 0 = u(B, \omega_B)$

Example II. $u(A, \omega_A) = 1 > .75 = u(B, \omega_A); u(A, \omega_B) = 0 < .25 = u(B, \omega_B)$

It is straightforward to check that $\alpha^I(s_L) = .19, \alpha^I(s_H) = 1$ and $\alpha^{II}(s_L) = .77, \alpha^{II}(s_H) = 1$ are symmetric responsive non-strategic equilibria of the unperturbed game for examples I and II, respectively (i.e., $\Delta_i^N(P(\alpha^j), s_H) > 0 = \Delta_i^N(P(\alpha^j), s_L)$ for $j = I, II$). Stability is determined by applying Theorem 3 and computing the corresponding Jacobian. We use the steps in the proof of Proposition 1 (in particular, we replace the NE-derivatives with the N-derivatives provided in Step 1 and use the approximations described in part (ii)) to compute the Jacobian for a small value of η (we use $\eta = .0000001$). By numerical computation, we find that the analog of the matrix $M_{s^*s^*}$ that we constructed for the case of Nash equilibrium is given by

$$M^I_{LL} = \begin{pmatrix} 815916 & 442841 & 442841 & -816078 & -442928 & -442928 & 0 & 0 & 0 & 0 & 0 & 0 \\ 442841 & 815916 & 442841 & -442928 & -816078 & -442928 & 0 & 0 & 0 & 0 & 0 & 0 \\ 442841 & 442841 & 815916 & -442928 & -442928 & -816078 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2050425 & 1304273 & 1304273 & -2050831 & -1304531 & -1304531 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1304273 & 2050425 & 1304273 & -1304531 & -2050831 & -1304531 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1304273 & 1304273 & 2050425 & -1304531 & -1304531 & -2050831 & 0 & 0 & 0 & 0 & 0 & 0 \\ -815917 & -442841 & -442841 & 816078 & 442928 & 442928 & -1 & 0 & 0 & 0 & 0 & 0 \\ -442841 & -815917 & -442841 & 442928 & 816078 & 442928 & 0 & -1 & 0 & 0 & 0 & 0 \\ -442841 & -442841 & -815917 & 442928 & 442928 & 816078 & 0 & 0 & -1 & 0 & 0 & 0 \\ -2050425 & -1304273 & -1304273 & 2050830 & 1304531 & 1304531 & 0 & 0 & 0 & -1 & 0 & 0 \\ -1304273 & -2050425 & -1304273 & 1304531 & 2050830 & 1304531 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1304273 & -1304273 & -2050425 & 1304531 & 1304531 & 2050830 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

and

$$M^{II}_{LL} = \begin{pmatrix} 81959 & 52403 & 52403 & -42930 & -27448 & -27448 & -710711 & -454414 & -454414 & 163998 & 104857 & 104857 \\ 52403 & 81959 & 52403 & -27448 & -42930 & -27448 & -454414 & -710711 & -454414 & 104857 & 163998 & 104857 \\ 52403 & 52403 & 81959 & -27448 & -27448 & -42930 & -454414 & -454414 & -710711 & 104857 & 104857 & 163998 \\ 301002 & 241889 & 241889 & -157665 & -126701 & -126701 & -2610129 & -2097535 & -2097535 & 602295 & 484012 & 484012 \\ 241889 & 301002 & 241889 & -126701 & -157665 & -126701 & -2097535 & -2610129 & -2097535 & 484012 & 602295 & 484012 \\ 241889 & 241889 & 301002 & -126701 & -126701 & -157665 & -2097535 & -2097535 & -2610129 & 484012 & 484012 & 602295 \\ -81960 & -52403 & -52403 & 42930 & 27448 & 27448 & 710710 & 454414 & 454414 & -163998 & -104857 & -104857 \\ -52403 & -81960 & -52403 & 27448 & 42930 & 27448 & 454414 & 710710 & 454414 & -104857 & -163998 & -104857 \\ -52403 & -52403 & -81960 & 27448 & 27448 & 42930 & 454414 & 454414 & 710710 & -104857 & -104857 & -163998 \\ -301002 & -241889 & -241889 & 157664 & 126701 & 126701 & 2610129 & 2097535 & 2097535 & -602296 & -484012 & -484012 \\ -241889 & -301002 & -241889 & 126701 & 157664 & 126701 & 2097535 & 2610129 & 2097535 & -484012 & -602296 & -484012 \\ -241889 & -241889 & -301002 & 126701 & 126701 & 157664 & 2097535 & 2097535 & 2610129 & -484012 & -484012 & -602296 \end{pmatrix}$$

for examples I and II respectively. The set of eigenvalues Λ^j are given by

$$\Lambda^I = \{-1, -2958294, -373224\}$$

and

$$\Lambda^{II} = \{-1, -175080, 136606\}.$$

By Theorem 3, α^I is asymptotically stable and α^{II} is unstable.[33]

1.4 Verification of Condition 1 in Theorem 4

Here, we show that $L \in C^2$ (i.e., is twice differentiable with continuous first derivative and bounded second derivative). By inspection, it suffices to show that, $E_{P(\alpha(\cdot))}\left(\mathbf{1}_{Z^m_{oui}(s)}\right) \in C^2$. By definition of ξ, for $o = A$ (for $o = B$ the expression is similar and, therefore, omitted)

$$\begin{aligned}
E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z^m_{Aui}(s)}\right) =& P(\alpha(\xi))\left\{o = A, \omega \in u_i^{-1}(A,\mathtt{u}), x_{-i} \in piv_i^m, s\right\} \\
=& \Pr\left\{o = A, x_{-i} \in piv_i^m \mid \omega \in u_i^{-1}(A,\mathtt{u}), s\right\} \Pr\left\{\omega \in u_i^{-1}(A,\mathtt{u}), s\right\} \\
=& \alpha_i(s;\xi_i) \Pr\left\{x_{-i} \in E_i^m(\omega;\xi) \mid \omega \in u_i^{-1}(A,\mathtt{u}), s\right\} \Pr\left\{\omega \in u_i^{-1}(A,\mathtt{u}), s\right\} \\
&+ (1-\alpha_i(s;\xi_i)) \Pr\left\{x_{-i} \in G_i^m(\omega;\xi) \mid \omega \in u_i^{-1}(A,\mathtt{u}), s\right\} \Pr\left\{\omega \in u_i^{-1}(A,\mathtt{u}), s\right\}
\end{aligned}$$

where the event $E_i^m(\omega;\xi) = \{x_{-i} \in piv_i^m \cap \{\cup_{k\geq K-1}\{K_{-i}(\omega;\xi) = k\}\}\}$ and $G_i^m(\omega;\xi) = \{x_{-i} \in piv_i^m \cap \{\cup_{k\geq K}\{K_{-i}(\omega;\xi) = k\}\}\}$. Where, for any set $N' \subset \{1,2,...,n\}$, $K_{-N'}(\omega;\xi) \equiv \sum_{i\notin N'} K_i(\omega;\xi_i)$ and $K_i(\omega;\xi_i)$ takes value 1 with probability $\kappa_i(\omega;\xi_i) \equiv \sum_{s\in S_i} \alpha_i(s;\xi_i) q(s|\omega)$. The term $\Pr\left\{\omega \in u_i^{-1}(A,\mathtt{u}), s\right\}$ does not depend on ξ^m, and is bounded above by one; thus, it can be ignored.

For the case $m = NE$, $G_i^m(\omega;\xi)$ is the empty set, and thus can be ignored; $\Pr\left\{x_{-i} \in E_i^{NE}(\omega;\xi) \mid \omega \in u_i^{-1}(A,\mathtt{u})\right\} = \Pr(K_{-i}(\omega;\xi) = k-1 \mid \omega \in u_i^{-1}(A,\mathtt{u}), s)$ which in turn equals

$$\kappa_j(\omega;\xi_j) \Pr\left\{K_{-ij}(\omega;\xi) = k-2\right\} + (1-\kappa_j(\omega;\xi_j)) \Pr\left\{K_{-ij}(\omega;\xi) = k-1\right\}.$$

So, $E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z^{NE}_{Aui}(s)}\right)$ is a function which contains algebraic operations (multiplication and sums) of $\alpha_i(s;\xi_i)$ for all i, s. In turn, under assumption A1,

$$\frac{d\alpha_j(s;\xi_j)}{d\xi_{Auj}(s)} = f_j\left(\frac{\sum_{\mathtt{u}\in U_{Aj}} \xi_{A\mathtt{u}j}(s)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Aj}} \xi_{A\mathtt{u}j}(s)} - \frac{\sum_{\mathtt{u}\in U_{Bj}} \xi_{B\mathtt{u}j}(s)\cdot \mathtt{u}}{\sum_{\mathtt{u}\in U_{Bj}} \xi_{B\mathtt{u}j}(s)}\right)\left\{\frac{\sum_{\bar{\mathtt{u}}\in U_{Aj}} \xi_{A\mathtt{u}j}(s)(\mathtt{u}-\bar{\mathtt{u}})}{\left(\sum_{\mathtt{u}\in U_{Aj}} \xi_{A\mathtt{u}j}(s)\right)^2}\right\}.$$

Since $\xi_{oui}(s) \in [c,1]$ for all $o, \mathtt{u}, i, s$, it follows that $\frac{d\alpha_j(s;\cdot)}{d\xi_{Auj}(s)}$ is continuous (for all j, u', s).

[33] We have also verified that these examples are robust to small changes in the primitives and in the value of η.

This automatically implies that $Z_i(\omega)$ is differentiable and continuous and thus so are $\Pr\left\{x_{-i} \in E_i^{NE}(\omega;\xi) \mid \omega \in u_i^{-1}(A,\mathtt{u})\right\}$ and $E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z_{\mathtt{Au}i}^{NE}(s)}\right)$. Hence we established that for $m = NE$, $E_{P(\alpha(\xi))}\left(\mathbf{1}_{Z_{\mathtt{Au}i}^{NE}(s)}\right)$ is differentiable with continuous first derivative. Further algebra and the fact that f_i is differentiable (assumption A1) imply that the second derivative exists and is bounded. For the case $m = N$ we note that $E_i^N(\omega;\xi) = \{x_{-i}\colon \ \cup_{k\geq K-1}\{K_{-i}(\omega;\xi)=k\}\}$ and $G_i^N(\omega;\xi) = \{x_{-i}\colon \ \cup_{k\geq K}\{K_{-i}(\omega;\xi)=k\}\}$. By simple algebra, $\Pr\left\{x_{-i} \in E_i^N(\omega;\xi) \mid \omega \in u_i^{-1}(A,\mathtt{u})\right\} =$

$= \kappa_j(\omega;\xi_j)\Pr\{K_{-ij}(\omega;\xi) \geq k-2\} + (1-\kappa_j(\omega;\xi_j))\Pr\{K_{-ij}(\omega;\xi)\geq k-1\}$,

and $\Pr\left\{x_{-i} \in G_i^N(\omega;\xi) \mid \omega \in u_i^{-1}(A,\mathtt{u})\right\} =$

$= \kappa_j(\omega;\xi_j)\Pr\{K_{-ij}(\omega;\xi) \geq k-1\} + (1-\kappa_j(\omega;\xi_j))\Pr\{K_{-ij}(\omega;\xi)\geq k\}$. These expressions are analogous to the previous expressions we derived for the case $m = N$; hence it is easy to see that for $m = N$, $E_{P(\alpha(\cdot))}\left(\mathbf{1}_{Z_{\mathtt{Au}i}^{N}(s)}\right) \in C^2$ for all $(i,\mathtt{u},s)$. Thus, condition 1 in Theorem 4 is verified. □