Learning What Matters

Itai Arieli[*], Manuel Mueller-Frank[†]

July 14, 2014

Abstract

This paper considers a strict sequential learning model with a compact metrizable state and action space, thus capturing rich environments which have not been analyzed so far. We study both the case where the utility function is common to all decision maker and the case where the utility function of each agent is an iid draw from the space of continuous utility functions. Each agent can only identify whether or not the utility function of any given predecessor is close to his own. We establish asymptotic learning result as a general equilibrium property of sequential social learning if private signals are unbounded.

1 Introduction

The seminal papers on sequential social learning by Banerjee [3] and Bikhchandani, Hirshleifer and Welch [4] provide a profound formal explanation for herding, i.e. the tendency to act as others. In the standard model, a countable set of fully rational privately informed agents make an irreversible choice in a predetermined sequence under uncertainty while observing the choices of all predecessors. Each agent progressively learns by making inferences regarding the private information

[*]Faculty of Industrial Engineering, Technion- Israel Institute of Technology, iarieli@tx.technion.ac.il

[†]IESE Business School, University of Navarra, mmuellerfrank@iese.edu.

of her predecessors based on their actions. Banerjee [3] and Bikhchandani, Hirshleifer and Welch [4] show that observing the choices of some early movers can entice an infinite set of individuals to disregard their own private information and as a result lead to herding. Such herds (formally information cascades) are prone to lead to suboptimal choices as they are based on a finite number of random signals which fail to convey the true state with certainty.

The main objective of the literature thereafter has been to characterize necessary and sufficient conditions on the environment, like the signal structure and/or the observational structure such that asymptotic learning, where actions converge to the correct action in probability, occurs. For examples see Smith and Sorensen [10] and Acemoglu, Dahleh, Lobel and Ozdaglar [1].

So far, the literature focused primarily on environments where the set of states and actions are binary, and every agent would want to match his action with the true state. Surprisingly, this simple environment generates deep phenomena such as information cascades and asymptotic learning while being expressive enough to encompass many real life application. For some cases, however, the binary environment fails to adequately represent the complexity of decision making.

As an example, consider dietary and lifestyle choices. Here the set of possible choices is certainly rich. Among others one has to decide how much and what to eat, how much and how to exercise, how much to sleep, et cetera. The state space for this example summarizes the possible outcomes of each lifestyle choice, like for example the life span, types of illness avoided and suffered, etc. Moreover, in the context of lifestyle choices the assumption of a common utility function is too restrictive. Instead, some agents might knowingly trade-off future health for current benefits, for example smokers, while others do not.

For the finite case Smith and Sorensen [10] establish a striking asymptotic learning characterization using properties of the signals. They define private signals to be unbounded if the support of the conditional private beliefs contains 0 and 1, and bounded if the support includes neither 0 nor 1. Their main result shows that asymptotic learning holds if the private signals are unbounded and fails if private signals are bounded.

However, there is no understanding of the condition that assure asymptotic learning in general environments. The objective of this paper is to close this gap in the literature. We consider the case where both the set of states of the world and the action space are general compact metrizable spaces, and where any continuous utility function on the state and action space is admissible. First, we focus on the case of a general common continuous utility function and establish a sufficient condition on the private signals such that asymptotic learning occurs. Second, we allow for heterogeneity in preferences and provide an approximate asymptotic learning result for the case where every decision maker has some limited knowledge in regards to his predecessors preferences.

1.1 Common Utility Function

We shall first focus on the case where all agents share a common utility function. The characterization provided by Smith and Sorensen for the finite case asserts that having unbounded signals is a sufficient condition for learning. One may interpret Smith and Sorensen's result using *cascading regions*, i.e. a range of public beliefs where the optimal action of an agent is independent of his private signal realization. That is, if the public belief at stage t enters a cascading regions then from this stage on all agents play the same action. Note that in the finite environment unbounded signals are equivalent to having no cascading regions. If signals are unbounded then no matter how certain the decision maker is in regards to the true state, as long as he assigns positive probability to both states, his signal with positive probability might be strong enough to make him revert his decision. In this context one can rephrase Smith and Sorensen result as follows: If there are no cascading regions then asymptotic learning holds.

Naturally the question arises whether the absence of cascading regions is also a sufficient condition for learning in our general environment. We provide an example of a signalling structure with no cascading regions in which asymptotic learning fails. The example illustrates two related points. First, it is not clear how to generalize the unboundedness property to rich state spaces. Second, it is not a straight-forward exercise to provide a sufficient condition for asymptotic learning in the general environment.

We introduce the following general notion of unbounded signals. An event E is *uncertain* if it has prior probability strictly between zero and one. Say that beliefs are *unbounded with respect to* E if the support of the posterior beliefs of E after receiving the signal contains both zero and one. Private signals are defined as unbounded if beliefs are unbounded with respect to *any* uncertain event E. We note that in the finite case this condition coincides with Smith and Sorensen's definition of unbounded signals.

Our first main result (Theorem 1) shows that unbounded private signals are a sufficient condition for asymptotic learning. More precisely, we show that when signals are unbounded asymptotic learning holds for *any* common continuous utility in every equilibrium. Our result thus strengthens the result of Smith and Sorensen and establishes the robustness of asymptotic learning in the general case as it holds independent of the utility function.

1.2 Heterogeneous Preferences

Whenever agents have heterogeneous preferences and are uninformed in regards to the preferences of their predecessors, learning can fail regardless of the properties of the signals. We consider the case where agents do have some limited knowledge in regards to their predecessors preferences. We say that agents are *culturally close* if their preferences are close. We analyze an environment where each agent observes the full history of actions but can only identify those agents which are culturally close to him. The logic behind this assumption is that agent mostly interact with similar types and can identify them as such. As a result, agents can only interpret those decision taken by agents which satisfy this cultural proximity assumption.

Our notion of cultural proximity bears some resemblance to notion of homophily, the tendency of individuals of similar traits to associate with each other, which is common in the network literature. The main feature of our model is that agents observe the actions of all predecessors but can only identify types that are similar to themselves, as opposed to observing only similar types. For all other, non-similar types, a given agent has no information on their preferences and as such can hardly interpret the actions they are taking.

More precisely, we consider a given distribution $\mathbf{Q}$ over the space of all continuous utility functions which we identify as types. The type distribution represents the preference heterogeneity within the population. At every time period t an agent of a certain type is drawn in accordance with $\mathbf{Q}$ and then chooses an action. The given agent observes the full history of actions taken by the preceding agents, but he only has a partial knowledge with respect to their type. We think of this knowledge as derived from a finite partition over the set of all types.

The knowledge of the current decision maker is derived from the partition as follows. The realized utility function of a given agent is unknown by others. However, an agent can identify whether or not any given predecessor has preferences close to his own, i.e. whether they lie within the same cell of partition of types as his own. Our second main result (Theorem 2) shows that for every small ϵ there exists a partition which is based on cultural proximity such that for every large enough time t the payoff to agent t will be ϵ close to the optimal payoff, given the true state of the world, with a probability larger than $1 - \epsilon$.

2 Related Literature

Following the seminal paper by Bikhchandani, Hirshleifer and Welch [4] the social learning literature focused on understanding the phenomena of information cascades and learning in more general environments. Smith and Sorensen [10] generalize the signal structure from finite to metric spaces and introduce the concept of asymptotic learning. Thereafter, the literature primarily focused on more general observation structures as for example in Acemoglu, Dahleh, Lobel and Ozdaglar [1], Lobel and Sadler [6], and Arieli and Mueller-Frank [2].

This paper differs from the existing literature predominately in two dimensions. First, we extend the finite environment with a particular utility function to a general, compact metrizable environment while considering the space of continuous utility functions. Our main contribution here is to establish that asymptotic learning is indeed a general equilibrium property under unbounded private signals, as it holds for all continuous utility functions in every equilibrium.

Second, we generalize the standard model to allow for heterogeneous types. The most closely related papers in this context are Smith and Sorensen [10] and Lobel and Sadler [7]. Smith and Sorensen consider finite multiple types that are drawn from a commonly known distribution and mainly focus on confounded learning where no inference from actions can be drawn. Lobel and Sadler [7] consider a sequential social learning model where the observation structure of agents and their preferences are randomly drawn. The distribution is commonly known but the observation structure and preference of any given agent is his private information. They consider linear preferences that contain a common and private value component and show that if preferences are sufficiently diverse then asymptotic learning fails in sparse networks. On the other hand, if each agent can identify a neighbor that is arbitrarily close to his own preference then asymptotic learning holds.

The main distinction of our analysis to the related papers is that we allow for arbitrary distribution over the space of continuous utility functions and in our model every agent can identify whether or not any given predecessor's utility function lies with a fixed distance or not. As opposed to Lobel and Sadler [7] we focus on the case where the history of actions is commonly known. Our main contribution in this context is to establish that the more precisely one can identify whether or not a utility function is similar, the closer is the realized utility of agents to the optimal utility under knowledge of the realized state of the world with increasing probability.

3 A General Model of Sequential Social Learning with Homogeneous Types

A countably infinite set of agents $\{1,2,3,...\}$ sequentially select an irreversible action. The actions are taken under uncertainty which is represented by a pair (Ω,μ) where Ω is a compact metrizable set of states of the word and $\mu\in\Delta(\Omega)$ a common prior. The set of actions A is a compact metrizable space. The utility of each agent t depends on the realized state and his own action a_t. All agents share the same continuous utility function $u : A\times\Omega\to\mathbb{R}$.

At time $t=0$ the state of the world is drawn according to μ. At any later time $t\in\mathbb{N}$, prior to selecting his action, agent t observes the history of actions of his predecessors, $h^t\in A^{t-1}$, and receives a conditionally iid signal $s_t\in S$, where the signal space S is standard Borel. The signal s_t is drawn according to a state dependent probability measure $F(\omega)$. We assume that for every $\omega,\omega'\in\Omega$ the probability measures $F(\omega)$ and $F(\omega')$ are absolutely continuous with respect to each other.

A strategy $\sigma_t : A^{t-1}\times S\to A$ of agent t is a measurable mapping that assigns an action to each possible information set, i.e. for every possible pair of observed history and private signal. As common in the literature we solve the game for its pure strategy Perfect Bayesian equilibria, the set of strategy profiles $\langle\sigma_t\rangle_{t\in\mathbb{N}}$ such that each σ_t maximizes the expected utility of agent t given the strategies of all other agents.

3.1 The Private Signal Structure

The standard model features a binary state space, $\Omega=\{0,1\}$. Let q_1 be the random variable that represents the posterior probability agent 1 assigns to state 1. Smith and Sorensen [10] define unbounded private signals as follows.

Definition 1. Let $\underline{\beta},\overline{\beta}$ be defined as follows

$$\begin{aligned}\underline{\beta} &= \inf\{r\in[0,1] : \Pr(q_1\le r)>0\}\,,\text{ and}\\ \overline{\beta} &= \sup\{r\in[0,1] : \Pr(q_1\le r)<1\}\,.\end{aligned}$$

The information structure (F, μ) generates *bounded private signals* if $0 < \underline{\beta} < \overline{\beta} < 1$ and *unbounded private signals* if $\underline{\beta} = 1 - \overline{\beta} = 0$.

Smith and Sorensen [10] show that asymptotic learning occurs in any equilibrium if private signals are unbounded, and fails in any equilibrium if private signals are bounded.

It is not clear at this stage how to define unbounded private signals in our general model. In the standard model unbounded signals are equivalent to the absence of *cascading regions*. A cascading region consists of an interval of beliefs such that if at a certain time period the public prior belief lies in this region then the optimal action of each subsequent agent is independent of her private signal. Once the public belief enters the cascading region all subsequent actions remain fixed with probability one. We note that signals are unbounded if and only if there exists no cascading region. Therefore, one interpretation of Smith and Sorensen's [10] result is that in the binary case a sufficient condition for asymptotic learning is the absence of cascading regions.

The role of the following example is to demonstrate that in our case the absence of cascading regions is not a sufficient condition for asymptotic learning.

Example 1. Let $\Omega = [0, 1]$ and $A = \{0, 1\}$. Define the payoff function u as follows,

$$u(a, \omega) = (1 - a)\left(\frac{1}{2} - \omega\right) + a\left(\omega - \frac{1}{2}\right).$$

That is, when the state ω is above $\frac{1}{2}$ players would like to play action 1 and when ω is below $\frac{1}{2}$ players would like to play action 0. Let μ_1 be the uniform measure over $\Omega = [0, 1]$ and let μ_2 be any measure that assigns probability one to the set of rational numbers that are greater than 0 and smaller than 1, and assigns positive probability to every such rational number. The prior equals $\mu = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$. Let ν be any distribution over the natural numbers $\mathbb{N}$ that assigns a positive probability to every positive natural number. Define $F(\omega)$ as follows. First a natural number n is drawn according to ν, then if ω is rational, a coin with a parameter ω for head is drawn n times, and if ω is irrational a coin with a parameter $\frac{1}{\pi}$ for head is drawn n times. Consider any public belief μ_t which is absolutely continuous with respect to μ. Note first that p_t, the posterior belief of agent t, is arbitrarily close to the point belief $\mathbf{1}_x$ for any $x \in [0,1]$ with positive probability. To see this, note that for any state $\omega \in (0,1)$ there exists a positive probability that n is very large and the proportion of heads observed is close to x. This yields a posterior belief p_t that assigns high probability to a neighborhood of rational numbers around x. Hence there is no range of public beliefs where agents are cascading on a certain action.

Nevertheless, if the true state ω is irrational, which happens with probability $\frac{1}{2}$, then, since the signal distribution is the same for all the irrationals, no information can be drawn with respect to the identity of the state. To see this note that if ω is irrational even observing the full history of signals is not sufficient to determine the true state, or the optimal action that corresponds to it.

We now generalize the concept of unbounded private signals to our framework. For every measurable event $E \subset \Omega$ and for the given prior μ let $q_E = p(E|s_1) \in [0,1]$ be the random variable that represents agent 1's posterior probability of event E, conditional on his private signal $s_1 \in S$. Let $\mathcal{E} \subset \mathcal{B}$ be the set of all measurable events in Ω such that their prior probability lies strictly between zero and one,

$$\mathcal{E} = \{E \in \mathcal{B} : \mu(E) \in (0,1)\}.$$

Unbounded private signals are now defined as follows.

Definition 2. Let $\underline{\beta}_E, \overline{\beta}_E$ be defined as

$$\begin{aligned} \underline{\beta}_E &= \inf\{r \in [0,1] : \Pr(q_E \leq r) > 0\}, \text{ and} \\ \overline{\beta}_E &= \sup\{r \in [0,1] : \Pr(q_E \leq r) < 1\}. \end{aligned}$$

The information structure (F, μ) generates *unbounded private signals* if $\underline{\beta}_E = 1 - \overline{\beta}_E = 0$ for all $E \in \mathcal{E}$.

In words, if the support of the posterior probability for every uncertain, positive probability event contains zero and one, then private signals are unbounded. Note that in the binary state space setting our definition is equivalent to the standard notion of bounded and unbounded signals of Definition 2 since for binary states the collection $\mathcal{E}$ consists only of $\omega = 0$ and $\omega = 1$. The private signals in our example above are not unbounded, since the posterior probability of the set of irrationals smaller than 0.5 is bounded away from 1.

3.2 Asymptotic Learning

The central question in the social learning literature concerns the learning properties of equilibria. The benchmark is asymptotic learning where the probability with which agents select the optimal action converges to one along the sequence, as defined in Smith and Sorensen [10]. Note that in the SSLM asymptotic learning is equivalent to convergence of beliefs, i.e. where the posterior probability distribution converges to the realized Dirac measure in probability, and convergence of payoffs, where the payoff of agents converges to the optimal payoff in probability.

In our general environment with a compact metrizable state and action space and a general utility function belief convergence might be impossible. A utility function that is constant on the domain may serve as a trivial example. Additionally, since the optimal action might not be unique even under knowledge of the realized state, we define asymptotic learning in terms of convergence of the payoff to the optimal payoff given the realized state of the world.

Consider a strategy profile σ and let $\mathcal{H}_t$ be the sigma algebra on $\Omega \times H^\infty \times S^\infty$ generated by the history h_t. Let $\mathcal{F}_t$ be the sigma algebra on $\Omega \times H^\infty \times S^\infty$ generated by the history h_t and agent t's private signal.

Let a_t be the random variable representing the realized action of agent t for a given strategy profile σ, and let $y_t = u(a_t, \omega)$ be it's realized payoff. The expected utility of agent t conditional on his realized information is denoted by r_t,

$$r_t = E_\sigma[u(a_t, \omega)|\mathcal{F}_t] = E_\sigma[y_t|\mathcal{F}_t].$$

In addition let v_t be the maximal expected utility conditional on $\mathcal{H}_t$. That is,

$$v_t = \max_{a \in A} E_\sigma[u(a, \omega)|\mathcal{H}_t].$$

Alternatively we can have b_t as a random action, measurable with respect to $\mathcal{H}_t$, that maximizes expected utility given the public history h_t. Then,

$$v_t = E_\sigma[u(b_t, \omega)|\mathcal{H}_t].$$

Let O be the random variable that assigns the maximal utility to every state ω,

$$O(\omega) = \max_{a \in A} u(a, \omega).$$

Asymptotic learning is now defined as follows.

Definition 3. Asymptotic learning holds if for every Bayesian equilibrium σ the payoff y_t converges to the optimal payoff $O(\omega)$ in probability.

In other words, asymptotic learning requires the realized payoff of agents to converge in probability to the optimal payoff under knowledge of the true state of the world.

4 A General Sufficient Condition for Asymptotic Learning

The main question this paper addresses concerns the possibility of asymptotic learning. The following theorem provides our main result for the case of a common utility function.

Theorem 1. *If (F, μ) generates unbounded signals, then asymptotic learning holds for every continuous utility function $u \in C(A \times \Omega)$ in any Perfect Bayesian equilibrium.*

The theorem establishes that asymptotic learning is a robust outcome of sequential observational learning if private signals are unbounded. So far, it was known that unbounded signals imply asymptotic learning in finite environments with a particular utility function. Theorem 1 shows that under unbounded signals, asymptotic learning holds for any compact metrizable state and action space and for every continuous utility function in any equilibrium of the sequential game.

4.1 The Proof of Theorem 1

The proof of Theorem 1 relies on two propositions. The following proposition establishes a general result on observational learning which is of independent interest.

Proposition 1. *The difference between the expected utility conditional on public and private information respectively, $r_t - v_t$, converges to zero almost surely.*

According to the proposition, the value of private information goes to zero almost surely, in any sequential social learning model with full observability of the history and a common continuous utility function on a compact metrizable state and action space. Applied to the binary setting, the value of private information goes to zero either because the history reveals the true state of the world with increasing certainty or because one might approach or has entered a cascading region. One nice side effect of Proposition 1 is that enables a very short and straight-forward alternative proof of Smith and Sorensen's [10] characterization of asymptotic learning.

The main difficulty in proving Theorem 1 lies in establishing that given the true state of the world ω the posterior belief of agent t gets arbitrarily close to the point belief, $\mathbf{1}_\omega$ infinitely often with probability 1. Formally, $\mathbf{1}_\omega$ is the random measure that assigns probability one to the true state of the world. The random posterior probability measure of agent t is denoted by p_t which takes values in $\Delta(\Omega)$,

$$p_t = \Pr\left[\mathrm{d}\omega \,|\mathcal{F}_t\right].$$

Let d_w denote the Prokhorov metric on the space of probability measures over Ω. The following proposition constitutes the core of the proof of Theorem 1.

Proposition 2. *For every $\epsilon > 0$, we have $d_w(\mathbf{1}_\omega, p_t) \leq \epsilon$ infinitely often with probability one.*

Proposition 2 is the mathematical most challenging part of both of our main theorems. Here we briefly outline its proof. Proposition 2 is established in two steps. First, for every ϵ we designate a range of public beliefs M_δ^ϵ such that if the public belief μ_t at any time t lies in this set then with some fixed probability $\delta > 0$ the private belief p_t of agent t satisfies $d_w(\mathbf{1}_\omega, p_t) \leq \epsilon$. Then we show that the public beliefs μ_t lies in M_δ^ϵ infinitely often with probability that approaches one when δ decreases. This shows that for every ϵ, $d_w(\mathbf{1}_\omega, p_t) \leq \epsilon$ infinitely often with a probability which is arbitrarily close to 1.

We now prove Theorem 1 given Proposition 1 and Proposition 2.

Proof. Theorem 1 is established in three steps. The function $\bar{u} : A \times \Delta(\Omega) \to \mathbb{R}$ is generated from u as follows

$$\bar{u}(a, \mu) = \int u(a, \omega) \mathrm{d}\mu.$$

1. Proposition 2 together with compactness of A implies that with probability one there exists a subsequence (a_t, p_t) that converges to a limit $(a^*, \mathbf{1}_\omega)$ where a_t satisfies

   $$u(a_t, p_t) \geq u(a, p_t)$$

   for all $a \in A$.

2. Continuity of u implies continuity of the function $\bar{u} : A \times \Delta(\Omega) \to \mathbb{R}$. Since σ is an equilibrium strategy the realized action a_t at time t is almost surely a best reply to p_t. Hence for all $a \in A$,

   $$r_t = \bar{u}(a_t, p_t) \geq \bar{u}(a, p_t).$$

   Since the best reply correspondence is a closed mapping, it follows that for every $a \in A$,

   $$\bar{u}(a^*, 1_\omega) \geq \bar{u}(a, 1_\omega).$$

3. Step 2 establishes that with probability one there exists a subsequence of r_t that converges to O. By Proposition 1 r_t converges almost surely. Hence r_t converges to O almost surely. Let $E_\sigma[O - y_t]$ be the deviation of the actual payoff of agent t from the maximal payoff. We will show first that

$\lim_{t\to\infty} E_\sigma[O - y_t] = 0$. To see this note that,

$$E_\sigma[O - y_t] = E_\sigma[O - E_\sigma[y_t|\mathcal{F}_t]] \quad (1)$$

$$= E_\sigma[O - r_t]. \quad (2)$$

(1) follows from the law of iterated expectation and (2) from the definition of r_t. Since the payoff function u is bounded and since r_t converges to O it follows from the dominated convergence theorem that, $E_\sigma[O - r_t]$ converges to zero.

4. Since $O(\omega) \geq u(a_t, \omega) = y_t$ it follows that for every ϵ,

$$\mathbf{P}_\sigma(O - y_t > \epsilon) \to_{t\to\infty} 0.$$

Hence y_t converges to O in probability.■

□

5 Asymptotic Learning with Heterogeneous Types

In this section we relax the assumption of all agents sharing a commonly known utility function. To be more precise, we allow for random, privately observed utilities. Let $\mathbf{Q}$ be a probability distribution over $C(A\times\Omega)$, the space of continuous functions on $A \times \Omega$. The utility function u_t of each agent t is independently drawn from $C(A\times\Omega)$ according to the probability distribution $\mathbf{Q}$ and is privately observed by agent t. No restrictions are imposed on the measure $\mathbf{Q}$.

In general, asymptotic learning might fail, if agents have no information on their predecessors' realized utility functions. Instead, we assume that each agent is able to identify whether any given predecessor has a similar utility function or not. To be more precise, we partition the space of continuous utility functions $C(A \times \Omega)$ according to a commonly known, finite partition $\mathcal{P}$. Beyond knowing his own utility function u_t that belongs to some partition cell $P \in \mathcal{P}$, agent t knows if the utility function of each predecessor u_{t-k} lies in his cell P or not. The idea behind this approach is that individuals with similar preferences can identify each other even though they are not being able to identify the exact preference of the close types.

Intuitively, our objective is to establish an approximate learning result, where agents take almost optimal actions in the long run, despite the very coarse knowledge regarding the types of other agents. Let σ_t be any strategy and let y_t be the corresponding payoff of agent t,

$$y_t = u_t(a_t, \omega).$$

The maximal payoff player t can achieve under his realized utility function and under knowledge of the true state of the world ω is denoted by O_t. Our main theorem states the following:

Theorem 2. *Let (F, μ) generate unbounded signals. For every $\epsilon > 0$ and distribution of types $\mathbf{Q}$ there exists a finite partition $\mathcal{P}_\epsilon$ of $C(A \times \Omega)$ such that for every equilibrium strategy σ there exists a time t_0 such that for every $t > t_0$,*

$$P_\sigma \left(|O_t - y_t| \leq \epsilon \right) \geq 1 - \epsilon.$$

The theorem establishes that asymptotic learning can be approximated with finite partitions of the space of continuous utility functions. Despite each agent being only able to identify for each predecessor to which uncountable set his utility function belongs, the long run actions are optimal up to an error of ϵ optimal with probability at least $1 - \epsilon$. The maximal deviation from the optimal utility can be made arbitrarily small by increasing the cardinality of the finite partition on the space of continuous utility functions.

The main challenge of Theorem 2 is to construct the finite partition $\mathcal{P}_\epsilon$. We first define two payoff functions as ϵ close if every best reply action with respect to any belief in one function is an ϵ-best reply with respect to the same belief in the other. That is, the payoffs u and v are ϵ close if for every $\mu \in \Delta(\Omega)$, every action a that is a best reply to μ for the function u, a yields a payoff that differs at most by ϵ from the optimal payoff for the belief μ with respect to the function v, and vice versa. We show that one can find a finite partition such that $\mathcal{P}_\epsilon$ such that with probability at least $1-\epsilon$ the chosen element $P \in \mathcal{P}_\epsilon$ has the property that any two members of P are ϵ close. We then apply similar considerations as in Theorem 1 to show that the members of every such partition element learn to approximately best reply to the true state.

6 Appendix

6.1 Proof of Proposition 1

Lemma 1. *v_t is a convergent submartingale. Additionally, for every t the following holds almost surely*

$$v_t \leq r_t, \tag{3}$$

$$r_t = v_{t+1}. \tag{4}$$

Begin Proof: To see that v_t is a submartingale we need to show that

$$v_t \leq E\left[v_{t+1} \left| \mathcal{H}_t\right.\right]$$

for all $t \in \mathbb{N}$. By the law of iterated expectations we have

$$v_t = E_\sigma\left[E_\sigma[u(b_t,\omega)|\mathcal{H}_{t+1}] \left| \mathcal{H}_t\right.\right]$$

By definition of b_{t+1} it holds that,

$$E_\sigma[u(b_t,\omega)|\mathcal{H}_{t+1}] \leq E_\sigma[u(b_{t+1},\omega)|\mathcal{H}_{t+1}].$$

Therefore,

$$v_t = E\left[E_\sigma[u(b_t,\omega)|\mathcal{H}_{t+1}] \left| \mathcal{H}_t\right.\right] \leq E\left[E_\sigma[u(b_{t+1},\omega)|\mathcal{H}_{t+1}] \left| \mathcal{H}_t\right.\right] = E\left[v_{t+1} \left| \mathcal{H}_t\right.\right].$$

establishing that v_t is indeed a submartingale. The utility function u is bounded due to compactness of $A \times \times \Omega$ and continuity, and hence v_t converges by the Martingale Convergence Theorem. Equation (3) follows since $\mathcal{H}_t \subset \mathcal{F}_t$. To show Equation (4) note first that the action a_{t+1} is measurable with respect $\mathcal{H}_{t+1}$, hence by observing the history h_{t+1} one can can imitate the action of agent t and achieve an expected payoff which is not lower than his. That is,

$$v_{t+1} \geq r_t \tag{5}$$

To see the reverse inequality note that since $\mathcal{H}_{t+1} \subset \mathcal{F}_t$ we have that

$$v_{t+1} \leq r_t.$$

■

The proof of Proposition 1 is now straight-forward.

Beginn Proof: By the Martingale Convergence Theorem v_t converges to v^* almost surely. By Lemma 1 we have $r_t = v_{t+1}$ which implies almost sure convergence of r_t to v^*.■

6.2 Proof of Proposition 2

Lemma 2. *Assume that (F, μ) generate unbounded beliefs and let $\nu \in \Delta(\Omega)$ be absolutely continuous with respect to μ such that for μ almost every $\omega \in \Omega$,*

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega) \leq M,$$

for some $M > 0$. Then (F, ν) generate unbounded beliefs.

Proof. For every $\mu \in \Delta(\Omega)$ let $\mu^* \in \Delta(\Omega \times S)$ be the measure that induced by μ and F over $\Omega \times S$. For convenience, with a slight abuse of notations, we denote by μ^* also the measure that is obtained on the signal space S from μ and F.

By definition (ν, F) generate unbounded beliefs iff for every event E with $\nu(E) > 0$ and for every δ there exists a set of signals S_δ with $\nu^*(S_\delta) > 0$ such that,

$$\nu^*(E|S_\delta) \geq 1 - \delta.$$

Since $\nu(E) > 0$ and since $\nu \ll \mu$ we can find an event $E' \subset E$ such that $\nu(E') > 0$ and $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega) \geq \alpha$, for some $\alpha > 0$. It clearly holds that $\mu(E') > 0$. Fix $\delta > 0$ we shall show that there exists a set of signals T_δ with $\nu^*(T_\delta) > 0$ such that $\nu^*(E'|T_\delta) \geq 1 - \delta$.

By the above since $\mu(E') > 0$, for every ϵ there exists a set S_ϵ with $\mu^*(S_\epsilon) > 0$ such that,

$$\mu^*(E'|S_\epsilon) \geq 1 - \epsilon. \tag{6}$$

Note that for every ϵ it holds by Bayes rule,

$$\begin{aligned} \nu^*(E' \times S_\epsilon) &= \int_{E'} F(\omega)(S_\epsilon)\mathrm{d}\nu(\omega) \qquad (7) \\ &= \int_{E'} F(\omega)(S_\epsilon)\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\mathrm{d}\mu(\omega) \\ &\geq \alpha\int_{E'} F(\omega)(S_\epsilon)\mathrm{d}\mu(\omega) = \alpha\mu^*(E' \times S_\epsilon), \end{aligned}$$

The last inequality of (7) follows from the definition of E'.

Similarly, it holds that

$$\nu^*((E')^c \times S_\epsilon) \leq M\mu^*((E')^c \times S_\epsilon) \tag{8}$$

A simple application of Bayes rule together with equation (6) implies that.

$$\frac{\mu^*((E')^c \times S_\epsilon)}{\mu^*(E' \times S_\epsilon)} \leq \frac{\epsilon}{1-\epsilon}.$$

It then follows from equations (7) and (8) that, for every $\epsilon > 0$,

$$\frac{\nu^*((E')^c \times S_\epsilon)}{\nu^*(E' \times S_\epsilon)} \leq \frac{M\epsilon}{\alpha(1-\epsilon)}.$$

We can therefore choose $\epsilon > 0$ to be small enough such that,

$$\frac{\nu^*((E')^c \times S_\epsilon)}{\nu^*(E' \times S_\epsilon)} \leq \frac{\delta}{1-\delta}. \tag{9}$$

Simple application of Bayes rule yields,

$$\nu^*(E'|S_\epsilon) = \frac{1}{1 + \frac{\nu^*((E')^c \times S_\epsilon)}{\nu^*(E' \times S_\epsilon)}}.$$

Hence Equation (9) implies that,

$$\nu^*(E'|S_\epsilon) \geq 1 - \delta.$$

This concludes the proof of the Lemma. □

For every t let, $\mu_t = \mu_\sigma(\ |\mathcal{H}_t)$ be the public belief at time t. Let $\mu_\infty = \mu_\sigma(\ |\mathcal{H}_\infty)$ be the belief at infinity. By the martingale converges theorem we know that for every event $E \subset \Omega$,

$$\lim_t \mu_t(E) = \mu_\infty(E).$$

The following lemma demonstrates that with probability 1, no fully wrong belief can occur.

Let $\mathbf{P}_\sigma$ be the probability measure that μ and σ induce on $\Omega \times H_\infty \times S^\infty$. For every measurable set E let $\mathbf{P}_\sigma^E$ be the probability that is induced by $\mathbf{P}_\sigma$ conditional on the the true state lies in E.

Lemma 3. *Let $E \subset \Omega$ be a measurable set with $\mu(E) > 0$. The following hold with probability 1 under $\mathbf{P}_\sigma^E$.*

$$\lim_t \mu_t(E) = \mu_\infty(E) > 0. \tag{10}$$

That is, conditional on the true state of the world lies in E the public belief assigns a positive probability bounded away from 0 to E for every time t with probability 1.

Proof. Assume by contradiction that the limit in expression (10) is zero with positive probability under $\mathbf{P}_\sigma^E$. Let H_∞ be the set of all infinite histories. Let $H_0 \subset H_\infty$ be the subset of all histories h for which

$$\lim_t \mu_t(E)(h_t) = \mu_\infty(E)(h) = 0. \tag{11}$$

By the contradiction assumption we get that $\mathbf{P}_\sigma(E \times H_0) > 0$. Let ν be the marginal of $\mathbf{P}_\sigma$ over H_∞, by the law of total probability,

$$\mathbf{P}_\sigma(E \times H_0) = \int_{H_0} \mu_\infty(E)(h) \mathrm{d}\nu(h) = 0.$$

The last equality follows from (11). This stands in contradiction to $\mathbf{P}_\sigma(E \times H_0) > 0$. Hence $\mathbf{P}_\sigma^E(H_0) = 0$. □

Note that for every $t \in \mathbb{N} \cup \{\infty\}$ the conditional probability μ_t may be seen as a random measure that obtains values in $\Delta(\Omega)$ for every realization of history

h_t. Let $\mathbf{P}_t \in \Delta(\Delta(\Omega))$ be the measure that is induced by μ_t over $\Delta(\Omega)$. For every measure $\mathbf{P} \in \Delta\Delta(\Omega)$ and a measurable set E let,

$$\tilde{\mathbf{P}}(E) = \int_{\Delta(\Omega)} \lambda(E) \mathrm{d}\mathbf{P}(\lambda). \tag{12}$$

Note that for every $t \in \mathbb{N} \cup \{\infty\}$ it holds that $\tilde{\mathbf{P}}_t(E) = \mu(E)$. Fix a ball $\mathrm{B} \subset \Omega$ with $\mu(\mathrm{B}) > 0$, and let $\mathbf{P}_t^{\mathrm{B}} \in \Delta(\Delta(\Omega))$ be the measure that obtained from $\mathbf{P}_t$ conditional on the true state lies in the ball B.

Note first that for every $t \in \mathbb{N} \cup \{\infty\}$ whenever $\mu(B) > 0$, the measure $\tilde{\mathbf{P}}_t^{\mathrm{B}}$ is absolutely continuous with respect to μ. To see this note that $\mu_t(E)$ is a martingale for every measurable event $E \subset \Omega$, hence by (12) $\tilde{\mathbf{P}}_t(E) = \mu(E)$. Therefore, by Bayes rule

$$\mu(E) = \mu(\mathrm{B}) \tilde{\mathbf{P}}_t^{\mathrm{B}}(E) + \mu((\mathrm{B})^c) \tilde{\mathbf{P}}_t^{(\mathrm{B})^c}(E). \tag{13}$$

Hence $\tilde{\mathbf{P}}_t^{\mathrm{B}}(E) = 0$ whenever $\mu(E) = 0$. More generally we have the following lemma.

Lemma 4. *Let $\mathbf{P} \in \Delta(\Delta(\Omega))$ and let $K \subset \Delta(\Omega)$ be a measurable subset such that $\mathbf{P}(K) > 0$, define $\mathbf{Q} = \mathbf{P}(\ |K)$. The measure $\tilde{\mathbf{Q}}$ is absolutely continuous with respect to the measure $\tilde{\mathbf{P}}$ and the Radon Nykodim derivative $\frac{\mathrm{d}\tilde{\mathbf{Q}}}{\mathrm{d}\tilde{\mathbf{P}}}$ is $\tilde{\mathbf{P}}$ almost surely bounded by $\frac{1}{\mathbf{P}(K)}$.*

Proof. The lemma follows by applying similar considerations as above. Let $\mathbf{M} = \mathbf{P}(\ |(K)^c)$. For every measurable set $E \subset \Delta(\Omega)$ we have that,

$$\tilde{\mathbf{P}}(E) = \mathbf{P}(K) \tilde{\mathbf{Q}}(E) + \mathbf{P}((K)^c) \tilde{\mathbf{M}}(E). \tag{14}$$

Hence $\tilde{\mathbf{Q}}(E) > 0$ implies that $\tilde{\mathbf{P}}(E) > 0$. Moreover, by equation (14) we have for $\tilde{\mathbf{P}}$ a.e. $\omega \in \Omega$ that,

$$1 = \mathbf{P}(K) \frac{\mathrm{d}\tilde{\mathbf{Q}}}{\mathrm{d}\tilde{\mathbf{P}}}(\omega) + \mathbf{P}((K)^c) \frac{\mathrm{d}\tilde{\mathbf{M}}}{\mathrm{d}\tilde{\mathbf{P}}}(\omega).$$

Since both derivatives are non-negative we have that for $\tilde{\mathbf{P}}$ a.e. $\omega \in \Omega$,

$$\frac{\mathrm{d}\tilde{\mathbf{Q}}}{\mathrm{d}\tilde{\mathbf{P}}}(\omega) \leq \frac{1}{\mathbf{P}(K)}.$$

□

Lemma 5. *Let* B *be any ball such that* $\mu(\mathrm{B}) > 0$*. For every* $\epsilon > 0$ *and* η *there exists a* δ *such that if* $\mu(E) \geq 1 - \delta$ *then for every* t *it holds that*

$$\mathbf{P}_t^{\mathrm{B}}(\{\lambda : \lambda(E) \geq 1 - \eta) \geq 1 - \epsilon.$$

Proof. The proof follows readily from equations (12) and (13). □

Fix a ball B with $\mu(\mathrm{B}) > 0$. For every $0 < \alpha < 1$ let $M_\delta^\alpha \subset \Delta(\Omega)$ be the set of all measures λ for which there exists a measurable set of signals $T \subset S$ such that the following condition satisfied,

$$\lambda^*(\mathrm{B} \times T) > (1 - \alpha)\lambda^*(T) > \delta. \tag{15}$$

Let $M^\alpha = \bigcup_{\delta > 0} M_\delta^\alpha$.

Lemma 6. $\mathbf{P}_\infty^{\mathrm{B}}$ *assigns probability one to* M^α *for every* $\alpha > 0$*.*

Proof. Assume by contradiction that $\mathbf{P}_\infty^{\mathrm{B}}(M^\alpha) < 1$ for some $\alpha > 0$. Hence $\mathbf{P}_\infty^{\mathrm{B}}((M^\alpha)^c) > 0$. Let $\mathbf{Q}$ be the conditional probability of $\mathbf{P}_\infty^{\mathrm{B}}$ given the complement $(M^\alpha)^c$.

By Lemma 4 $\tilde{\mathbf{Q}}$ is absolutely continuous with respect to $\tilde{\mathbf{P}}_\infty^{\mathrm{B}}$ with a bounded Radon-Nykodim derivatives. Similarly, the measure $\tilde{\mathbf{P}}_\infty^{\mathrm{B}}$ is absolutely continuous with respect to μ with a bounded Radon Nykodim derivative. By Lemma 3 the measure $\mathbf{P}_\infty^{\mathrm{B}}$ assigns probability one to the set of all measures $\lambda \in \Delta(\Omega)$ for which $\lambda(\mathrm{B}) > 0$. Hence, $\tilde{\mathbf{Q}}(\mathrm{B}) > 0$. By Lemma 2 the pair $(\tilde{\mathbf{Q}}, F)$ forms an unbounded beliefs. Hence by the unboundedness condition there exists a set S_α such that $(\tilde{\mathbf{Q}})^*(S_\alpha) > 0$ and,

$$(\tilde{\mathbf{Q}})^*(\mathrm{B}|s) > (1 - \alpha)\ \forall s \in S_\alpha.$$

As a result,

$$(\tilde{\mathbf{Q}})^*(\mathrm{B} \times S_\alpha) = \int_{S_\alpha} (\tilde{\mathbf{Q}})^*(\mathrm{B}|s) \mathrm{d}(\tilde{\mathbf{Q}})^*(s) > (1 - \alpha)(\tilde{\mathbf{P}})^*(S_\alpha).$$

Therefore there exists a measurable set $K \subset \Delta(\Omega)$ and $\delta > 0$ such that $\mathbf{Q}(K) > 0$, and for every $\lambda \in K$,

$$\lambda^*(\mathrm{B} \times S_\alpha) > (1 - \alpha)\lambda^*(S_\alpha) > \delta.$$

Therefore we have established that $\mathbf{Q}(M^\alpha) > 0$. This stands in contradiction to the definition of $\mathbf{Q}$. □

Lemma 7. *For every $\alpha > 0$ and $\epsilon > 0$ there exists a δ and a time t_0 such that*

$$\mathbf{P}^{\mathrm{B}}_{\sigma}(\{\mu_t \in M^{\alpha}_{\delta}\ \forall t > t_0\}) \geq 1 - \epsilon.$$

Proof. Fix $\alpha > 0$. Since $\mathbf{P}^{\mathrm{B}}_{\infty}(M^{\alpha}) = 1$, and since

$$\lim_{\delta \to 0} \mathbf{P}^{\mathrm{B}}_{\infty}(M^{\alpha}_{\delta}) = \mathbf{P}^{\mathrm{B}}_{\infty}(M^{\alpha}),$$

for every ϵ there exists a $\delta > 0$ such that $\mathbf{P}^{\mathrm{B}}_{\infty}(M^{\alpha}_{\delta}) > 1 - \epsilon$.

Moreover, since S is a standard Borel space, its sigma algebra is countably generated. Therefore, there exists a countable family of measurable subsets $\{S_k\}_{k \in \mathbb{N}}$ such that condition (15) is satisfied with respect to some measurable subset $T \subset S$ iff it satisfied with respect to S_k for some $k > 0$. That is, for every $\lambda \in M^{\alpha}_{\delta}$ there exists a k such that,

$$\lambda^*(\mathrm{B} \times S_k) > (1 - \alpha)\lambda^*(S_k).$$

Therefore for every ϵ we can find a finite set of indices $\{1, \ldots, m\}$ such that the following condition hold with $\mathbf{P}^{B}_{\infty}$ probability at least $1 - \frac{\epsilon}{2}$.

- There exists a $k \in \{1, \ldots, m\}$ such that,

$$\lambda^*(\mathrm{B} \times S_k) > (1 - \alpha)\lambda^*(S_k) > \delta.$$

Note that by definition $\mu^*_t(\mathrm{B} \times T) = E_{\sigma}(\mathbf{1}_{\mathrm{B} \times T} | \mathcal{H}_t)$ for every measurable $T \subset S$. Hence for every k $\mu^*_t(\mathrm{B} \times S_k)$ is a martingale. By the martingale convergence theorem it holds that

$$\lim_{t \to \infty} \mu_t(\mathrm{B} \times S_k) = \mu^*_{\infty}(\mathrm{B} \times S_k),$$

for every k. Therefore we can find a large enough t_0 such that the following event holds with probability at least $1 - \frac{\epsilon}{2}$ with respect to $\mathbf{P}^{\mathrm{B}}_{\sigma}$,

- For every $t > t_0$ there exists a $k \in \{1, \ldots, m\}$ such that,

$$\mu^*_t(\mathrm{B} \times S_k) > (1 - \alpha)(\mu^*_t(S)) > \delta.$$

Hence in particular $\mathbf{P}^{\mathrm{B}}_{\sigma}(\{\mu_t \in M^{\alpha}_{\delta}\ \forall t > t_0\}) \geq 1 - \epsilon$. □

Lemma 8. *For every $\epsilon > 0$ and a constant $0 < \beta \leq 1$ there exists a constant L and a measurable set $K_\beta^\epsilon \subset \Omega$ such that $\mu(K_\beta^\epsilon) \geq 1 - \epsilon$ and for every $\omega, \omega'^\epsilon_\beta$ and a measurable subset $T \subset S$ it holds that $F(\omega')(T) \geq \frac{1}{L}$ whenever $F(\omega)(T) \geq \beta$.*

Proof. For every large $R > 1$ and $0 < \beta < 1$ let G_β^R be the set of all states $\omega \in \Omega$ that satisfy the following two conditions,

1. For every measurable subset $T \subseteq S$ with $\mu^*(T) \geq \beta$,

$$\frac{1}{R} \leq F(\omega)(T).$$

2. For every measurable subset $T \subseteq S$ for which $F(\omega)(T) \geq \beta$,

$$\frac{1}{R} \leq \mu^*(T).$$

We claim that for every $\omega \in \Omega$ and for every $0 < \beta < 1$ there exists R such that $\omega \in G_\beta^R$. To see this note that $F(\omega)$ is absolutely continuous with respect to μ^*. We can therefore let

$$\phi(s) = \frac{\mathrm{d}F(\omega)}{\mathrm{d}\mu^*}(s)$$

be the Radon-Nykodim derivative of $F(\omega)$ with respect to μ^*. For every constant D let $T_D = \{s : \phi(s) > \frac{1}{D}\}$. Since $\lim_{D\to\infty} \mu^*(T_D) = 1$, there exists a large enough constant $D > 0$ such that,

$$\mu^*(T_D) \geq 1 - \frac{\beta}{2}.$$

Hence for every set T for which $\mu^*(T) \geq \beta$ it holds that,

$$\begin{aligned} F(\omega)(T) &= \int_{(T_D \cap T)} \phi(s)\mathrm{d}\mu^*(s) + \int_{(T_D \cap (T)^c)} \phi(s)\mathrm{d}\mu^*(s) \\ &\geq \int_{(T_D \cap T)} \phi(s)\mathrm{d}\mu^*(s) > \frac{\beta}{2D}. \end{aligned} \tag{16}$$

Where (16) follows since $\mu^*(T \cap T_D) \geq \frac{\beta}{2}$. Therefore we can set $R_1 = \frac{2D}{\beta}$ and get that $F(\omega)(T) \geq \frac{1}{R_1}$ for every measurable T such that $\mu^*(T) \geq \beta$. By similar considerations we can find a positive constant R_2 such that $\mu^*(T) \geq \frac{1}{R_2}$ for every measurable T such that $F(\omega)(T) \geq \beta$. Let $R = \max\{R_1, R_2\}$ we clearly have $\omega \in G_\beta^R$.

Hence for every $\omega \in \Omega$ and $\beta > 0$ there exists a constant $R > 0$ such that $\omega \in G^R_\beta$. Therefore, $\lim_{R\to\infty} \mu(G^R_\beta) = 1$. Fix R such that $\mu(G^R_\beta) \geq 1 - \frac{\epsilon}{2}$.

We can apply the same considerations once more with respect to the constant $\frac{1}{R}$ to find a large enough constant $L > R$ such that $\mu(G^L_{\frac{1}{R}}) \geq 1 - \frac{\epsilon}{2}$. Let $K^\epsilon_\beta = G^R_\beta \cap G^L_{\frac{1}{R}}$. We claim that the set K^ϵ_β has the desired properties with respect to the constant L. First note that $\mu(K^\epsilon_\beta) \geq 1 - \epsilon$ as desired. Let $\omega, \omega' \in K^\epsilon_\beta$ and let $T \subset S$ be a measurable subset such that $F(\omega)(T) \geq \beta$. Since $\omega \in G^R_\beta$ the second condition in the definition of G^R_β above implies that $\mu^*(T) \geq \frac{1}{R}$. Since $\omega' \in G^L_{\frac{1}{R}}$ the first condition implies that $\frac{1}{L} \leq F(\omega')$, as desired. □

Lemma 9. *For every* $0 < \alpha \leq \frac{1}{5}$, $p_t(\mathrm{B}) \geq 1 - 3\alpha$ *infinitely often with* $\mathbf{P}^{\mathrm{B}}_\sigma$ *probability one.*

Proof. Fix $\epsilon > 0$. We shall show that $p_t(\mathrm{B}) \geq 1 - 3\alpha$ infinitely often with $\mathbf{P}^{\mathrm{B}}_\sigma$ probability at least $1 - \epsilon$. By Lemma 7 there exists a $\delta > 0$ and a time t_0 such that $\mu_t \in M^\alpha_\delta$ for every $t > t_0$ with probability at least $1 - \frac{\epsilon}{3}$.

By Lemmas 8 for every η there exists a set $K^\eta_{\frac{\delta}{4}} \subset \Omega$ and a corresponding constant L such that such that the following conditions hold (i) $\mu(K^\eta_{\frac{\delta}{4}}) \geq 1 - \eta$, (ii) for every $\omega, \omega' \in K^\eta_{\frac{\delta}{4}}$ and for every a measurable subset $T \subset S$ if $F(\omega) \geq \frac{\delta}{4}$ then $F(\omega')(T) \geq \frac{1}{L}$.

In addition, by Lemma 5 we can choose $\eta > 0$ such that the following conditions hold,

1. $\forall t \; \mathbf{P}^{\mathrm{B}}_t(\mu_t(K^\eta_{\frac{\delta}{4}}) \geq 1 - \frac{\delta}{4}) \geq 1 - \frac{\epsilon}{3}$,
2. $\mu(K^\eta_{\frac{\delta}{4}}|\mathrm{B}) \geq 1 - \frac{\epsilon}{3}$.

Condition 1 implies that $\mu_t(K^\eta_{\frac{\delta}{4}}) \geq 1 - \frac{\delta}{4}$ infinitely often with $\mathbf{P}^{\mathrm{B}}_\sigma$ with probability at least $1 - \frac{\epsilon}{3}$. It follows that the following conditions holds with probability at least $1 - \epsilon$ under $\mathbf{P}^{\mathrm{B}}_\sigma$:

$$\mu_t \in M^\alpha_\delta \text{ for every } t > t_0, \tag{17}$$

$$\mu_t(K^\eta_{\frac{\delta}{4}}) \geq 1 - \frac{\delta}{4} \text{ infinitely often}, \tag{18}$$

$$\text{The true state of the world lies in } K^\eta_{\frac{\delta}{4}}. \tag{19}$$

We claim that for every $\lambda \in M_\delta^\alpha$ for which $\lambda(K^\eta_{\frac{\delta}{4}}) \geq 1 - \frac{\delta}{4}$ there exists a measurable subset of signals T such that, $\lambda^*(\mathrm{B}|s) \geq 1 - 3\alpha$ for every $s \in T$ and $F(\omega)(T) \geq \frac{1}{L}$ for every $\omega \in K^\eta_{\frac{\delta}{4}}$. To see this note first that since $\lambda \in M_\delta^\alpha$ there exists a measurable set T' such that, $\lambda^*(\mathrm{B} \times T') > (1-\alpha)\lambda^*(T') > \delta$. Let

$$T = \{s \in T' | \lambda^*(\mathrm{B}|s) \geq 1 - 3\alpha\}.$$

Since,

$$\lambda^*(\mathrm{B} \times T') = \int_{T'} \lambda^*(\mathrm{B}|s) \mathrm{d}\lambda^*(s),$$

a simple calculation shows that the set T must have a probability strictly greater than $\frac{2}{3}\lambda^*(T')$. Hence $\lambda^*(T) > \frac{2\delta}{3}$. Moreover, since $\lambda(K^\eta_{\frac{\delta}{4}}) \geq 1 - \frac{\delta}{4}$ there must be $\omega \in K^\eta_{\frac{\delta}{4}}$ such that, $F(\omega)(T) \geq \frac{\delta}{4}$. To see this note first that by definition of T and since $\alpha \leq \frac{1}{5}$ it holds that

$$\lambda^*(\mathrm{B} \times T) \geq (1 - 3\alpha)\lambda^*(T) > \frac{\delta}{2}.$$

Assume by contradiction that $F(\omega)(T) < \frac{\delta}{4}$ for every $\omega \in K^\eta_{\frac{\delta}{4}}$. Then,

$$\begin{aligned}
\frac{\delta}{2} &< \lambda^*(\mathrm{B} \times T) = \int_{\mathrm{B} \cap K^\eta_{\frac{\delta}{4}}} \lambda^*(T|\omega) \mathrm{d}\lambda(\omega) + \int_{\mathrm{B} \cap (K^\eta_{\frac{\delta}{4}})^c} \lambda^*(T|\omega) \mathrm{d}\lambda(\omega) && (20) \\
&< \frac{\delta}{4} + \lambda((K^\eta_{\frac{\delta}{4}})^c) < \frac{\delta}{2}. && (21)
\end{aligned}$$

This leads to a contradiction. Therefore $F(\omega)(T) \geq \frac{1}{L}$ for every $\omega \in K^\eta_{\frac{\delta}{4}}$.

Hence if at time t it holds that (i) $\mu_t \in M_\delta^\alpha$, (ii) $\mu_t(K^\eta_{\frac{\delta}{4}}) \geq 1 - \frac{\delta}{4}$, and (iii) the true state lies in $K^\eta_{\frac{\delta}{4}}$ then $p_t(\mathrm{B}) \geq 1 - 3\alpha$ with probability at least $\frac{1}{L}$.

Let V_t be the event that the these three conditions hold at time t. By equations (17)-(19) V_t holds infinitely often with probability at least $1 - \epsilon$. Let U_t be the event that V_t holds and in addition $p_t(\mathrm{B}) \geq 1 - 3\alpha$. Let $\mathcal{G}$ be the Borel sigma algebra on Ω. For every time $t > 0$ let $\mathcal{G}_t$ be the sigma algebra generated by $\mathcal{G}$ and $\{\mathcal{F}_k\}_{k \leq t}$, where $\mathcal{F}_k$ is the sigma algebra generated by the information of agent k. Note that both V_t and U_t are $\mathcal{G}_t$ measurable. In addition, by the considerations above $\mathbf{P}_\sigma^\mathrm{B}(U_t|\mathcal{G}_{t-1}) \geq \mathbf{P}_\sigma^\mathrm{B}(V_t|\mathcal{G}_{t-1}) \cdot \frac{1}{L}$. Therefore,

$$\sum_{t \geq 1} \mathbf{P}_\sigma^\mathrm{B}(U_t|\mathcal{G}_{t-1}) \geq \sum_{t \geq 1} \frac{1}{L} \mathbf{P}_\sigma^\mathrm{B}(V_t|\mathcal{G}_{t-1}).$$

Hence,

$$\sum_{t\geq 1} \mathbf{P}^{\mathrm{B}}_{\sigma}(V_t|\mathcal{G}_{t-1}) = \infty \text{ implies } \sum_{t\geq 1} \mathbf{P}^{\mathrm{B}}_{\sigma}(U_t|\mathcal{G}_{t-1}) = \infty. \tag{22}$$

Levy's extension of the Borel-Cantelli lemma (see Corollary on page 486 in Shiyayev [9]) implies that,

$$\sum_{t\geq 1} \mathbf{1}_{V_t} = \infty \text{ iff } \sum_{t\geq 1} \mathbf{P}^{\mathrm{B}}_{\sigma}(V_t|\mathcal{G}_{t-1}) = \infty \text{ and }, \sum_{t\geq 1} \mathbf{1}_{U_t} = \infty \text{ iff } \sum_{t\geq 1} \mathbf{P}^{\mathrm{B}}_{\sigma}(U_t|\mathcal{G}_{t-1}) = \infty.$$

Note that U_t holds infinitely often iff $\sum_{t\geq 1} \mathbf{1}_{U_t} = \infty$. Therefore by (22) V_t holds infinitely often with probability at least $1-\epsilon$ implies that U_t holds infinitely often with probability at least $1-\epsilon$. By definition of the event U_t it holds that $p_t(B) \geq 1-3\alpha$ infinitely often with probability at least $1-\epsilon$. □

Let $\mathbf{1}_{\omega}$ be the random measure that assigns a probability one to the true state of the world. The following is a corollary from Lemma 9.

Corollary 1. *For every $\epsilon > 0$, the distance $d_w(\mathbf{1}_{\omega}, p_t) \leq \epsilon$ infinitely often.*

Proof. Fix any $\gamma > 0$. By definition of the Prokhorov metric for every ϵ there exists a δ such that for every B of radios δ or less if $\mu(\mathrm{B}) \geq 1-\delta$ then $d_w(1_{\omega}, \mu) \leq \epsilon$ for every $\omega \in \mathrm{B}$. Let $\{\mathrm{B}_k\}_{k=1}^{n}$ be a finite disjoint collection of balls all with a radios less than δ such that $\mu(\cup_{k=1}^{n}\mathrm{B}_k) \geq 1-\gamma$. Given that $\omega \in \mathrm{B}_k$ for some $1 \leq k \leq n$ one can deduce from the proposition that $p_t(\mathrm{B}_k) \geq 1-\delta$ infinitely often. Hence $d_w(\mathbf{1}_{\omega}, p_t) \leq \epsilon$ infinitely often with probability at least $1-\gamma$. As γ was arbitrary we can deduce that $d_w(\mathbf{1}_{\omega}, p_t) \leq \epsilon$ infinitely often with probability one. □

6.3 Proof of Theorem 2

Fix $\epsilon > 0$. For every $\delta > 0$, let C^{ϵ}_{δ} be the set of continuous functions $u \in C(A\times\Omega)$, such that for every two pairs $(a,\omega), (a',\omega') \in A\times\Omega$ for which $d((a,\omega),(a',\omega')) \leq \delta$ it holds that,

$$u(a,\omega) - u(a',\omega') \leq \epsilon.$$

Since every continuous function u on a compact domain is uniformly continuous we have that for every function $u \in C(A\times\Omega)$, and $\epsilon > 0$ there exists $\delta > 0$ such that, $u \in C^{\epsilon}_{\delta}$.

Let $\mathbf{Q} \in \Delta(C(A\times\Omega))$ be the probability distribution over the payoff functions. By the above considerations we can find small enough $\delta > 0$ and a set, $D \subset C_{\delta}^{\frac{\epsilon^2}{12}}$ such that $\mathbf{Q}(D) \geq 1 - \frac{\epsilon}{4}$, and $|u(a,\omega)| \leq M$ for every $u \in D$ and every $(a,\omega) \in A \times \Omega$.

Claim 3. *There exists a finite partition $\mathcal{P}_D$ of D such that, for every partition element $P \in \mathcal{P}_D$ and every $u, v \in P$ it holds that,*

$$\forall (a,\omega) \in A \times \Omega,\ |u(a,\omega) - v(a,\omega)| \leq \frac{\epsilon^2}{4}.$$

Proof of the Claim. Since $A \times \Omega$ is compact we can find a finite set of points $K \subset A \times \Omega$ such that, $\cup_{(a,\omega)\in K} \mathrm{B}_\delta(a,\omega) = A \times \Omega$. Since all function in D are uniformly bounded by M we can construct a finite partition $\mathcal{P}_D$ of D such that, for every partition element $P \in \mathcal{P}_D$ every $u, v \in P$, and every $(a,\omega) \in K$, it holds that, $|u(a,\omega) - v(a,\omega)| \leq \frac{\epsilon^2}{12}$. To see this note that we can partition the interval $[-M, M]$ into at most $\frac{24M}{\epsilon^2}$ closed intervals with disjoint interior of length bounded by $\frac{\epsilon^2}{12}$. Let $\mathcal{P}_D$ be a finite partition of D such that, for every $P \in \mathcal{P}_D$, *every* $u, v \in P$ and $(a,\omega) \in K$, it holds $u(a,\omega), v(a,\omega)$ lies in the same interval. I.e., $u(a,\omega), v(a,\omega)$ lies in a distance at most $\frac{\epsilon^2}{12}$.

We claim that $\mathcal{P}_D$ has the desired properties. To see this let $u, v \in P$ for some $P \in \mathcal{P}_D$, and let $(a',\omega') \in A \times \Omega$.

By construction there exists $(a,\omega) \in K$ such that $d((a,\omega),(a',\omega')) \leq \delta$, and $|u(a,\omega) - v(a,\omega)| \leq \frac{\epsilon}{3}$. In addition, since $u, v \in C_{\delta}^{\frac{\epsilon}{3}}$ we have that, $|u(a,\omega) - u(a',\omega')| \leq \frac{\epsilon^2}{12}$, and $|u(a,\omega) - u(a',\omega')| \leq \frac{\epsilon^2}{12}$.

Hence by the triangle inequality,

$$|u(a',\omega') - v(a',\omega)'| \leq |u(a,\omega) - u(a',\omega')| + |v(a,\omega) - v(a',\omega')| + |u(a,\omega) - v(a,\omega)| \leq \frac{\epsilon^2}{4}.$$

This concludes the proof of the claim. □

Let $\mathcal{P}$ be the partition over $C(A\times\Omega)$ obtained from $\mathcal{P}_D$ by adding $C(A\times\Omega)\setminus D$ as an additional element. Consider a fixed partition cell $P \in \mathcal{P}_D$ such that $\mathbf{Q}(P) > 0$. Let us fix a representative element $\tilde{u} \in P$. Let $\{\tau_k\}_{k=1}^{\infty}$ be a sequence of stopping times where τ_k is the kth time that a utility $u \in P$ was drawn. For each k, τ_k

is finite with probability 1 and has finite expectation. For every k let $\mathcal{H}_k$ be the sigma algebra generated by τ_k and the history h_{τ_k} up to time τ_k. Let $\mathcal{F}_k$ be the sigma algebra generated by $\mathcal{H}_k$ and the private signal of agent τ_k.

For every k let $\tilde{v}_k$ be the random variable representing the maximal expected payoff of an agent with a utility $\tilde{u}$ conditional on $\mathcal{H}_k$. That is, $\tilde{v}_k = \max_{a\in A} E_\sigma[\tilde{u}(a,\omega)|\mathcal{H}_{\tau_k}]$. Similarly let $\tilde{r}_k = r_{\tau_k} = E_\sigma[u_{\tau_k}(a_{\tau_k},\omega)|\mathcal{F}_k]$ is defined as above and represents the expected utilities conditional on the optimal action given the current utility u_{τ_k} and the sigma algebra $\mathcal{F}_{k},$.

Lemma 10. *The following holds for every k in every perfect Bayesian equilibrium σ,*

$$|\tilde{v}_k - \tilde{r}_k| \leq \frac{\epsilon^2}{4}.$$

Proof. By definition of the partition $\mathcal{P}$ and since $u_k, \tilde{u} \in P$ it holds for every pair $(a,\mu) \in A \times \Delta(\Omega)$ and every k that,

$$|\int_\Omega [u_k(a,\omega) - \tilde{u}(a,\omega)]\mathrm{d}\mu(\omega)| \leq \frac{\epsilon^2}{4}.$$

By the imitation principle it holds that for every k

$$|\tilde{v}_{k+1} - \tilde{r}_k| \leq \frac{\epsilon^2}{4}$$

□

Since $\tilde{v}_k$ is a converges submartingale we have that $\tilde{v}_k$ converges almost surely to a limit $\tilde{v}$. This implies that

$$\limsup_{k\to\infty} |\tilde{v} - r_k| \leq \frac{\epsilon^2}{4} \ a.s. \tag{23}$$

Let O_k be the maximal payoff of agent τ_k conditional on the true state of the world.

Lemma 11. *The following holds with $\mathbf{P}_\sigma$ probability 1,*

$$\liminf_{k\to\infty} |\tilde{r}_k - O_k| = 0.$$

Proof. This follows from our main lemma, as the private belief of agents $\tau_1, \tau_2, \ldots$ gets arbitrarily close to the point belief infinitely often with probability one. Since utility functions are continuous the expected payoff gets arbitrarily close to the optimal payoff infinitely often. □

In the following we provide the proof of Theorem 2.

Proof of Theorem 2. The theorem is established in four steps.

1. The convergence of $\tilde{v}_k$ to $\tilde{v}$ and (23) imply that there exists a time m_1^P such that
$$E_\sigma[r_k - \tilde{v}] \leq \frac{\epsilon^2}{3}$$
for all $k > m_1^P$.

2. Step 1 together with Lemma 11 imply that there exists $m_2^P > m_1^P$ such that
$$E_\sigma[O_k - \tilde{v}] \leq \frac{\epsilon^2}{3}$$
for all $k > m_2^P$.

3. Step 1 and 2 imply that
$$E_\sigma[O_k - \tilde{r}_k] \leq \frac{2\epsilon^2}{3}$$
for all $k > m_2^P$. In particular by definition of $\tilde{r}_k$ this implies that for every $k > m_2^P$,
$$E_\sigma[O_k - \tilde{y}_k] \leq \frac{2\epsilon^2}{3}.$$
Therefore, for every $k > m_2^P$ it must be the case that $\tilde{y}_k + \epsilon \geq O_k$ with probability at least $\frac{2\epsilon}{3}$.

4. Since there are finitely many partition elements P, there exists a finite t_0 such that the probability that there where m_2^P appearances of members of P before time t_0 for any P that obtains with positive probability, is at least $1 - \frac{\epsilon}{12}$. Therefore for any time $t > t_0$ the probability that a type is drawn that belongs to D and this type will be ϵ best replying is at least $1 - \epsilon$. This concludes the proof of the Theorem.

□

References

[1] D. Acemoglu, M.A. Dahleh, I. Lobel, A. Ozdaglar (2010), “Bayesian Learning in Social Networks”, *Review of Economic Studies* 78, 1-34.

[2] I. Arieli, M. Mueller-Frank (2014), ”Non-Sequential Learning and the Wisdom of Crowds”, working paper, IESE Business School.

[3] A.V. Banerjee (1992), “A Simple Model of Herd Behavior”, *Quarterly Journal of Economics* 107, 797-817.

[4] S. Bikhchandani, D. Hirshleifer, I. Welch (1992), “A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades”, *The Journal of Political Economy* 100, 151-170.

[5] A. S. Kechris, (1995) *Classical Descriptive Set Theory*, Springer-Verlag, New York.

[6] I. Lobel, E. Sadler (2013), ”Social Learning and Aggregate Network Uncertainty”, working paper, New York University.

[7] I. Lobel, E. Sadler (2013), ”Preferences, Homophily, and Social Learning”, working paper, New York University.

[8] D. Rosenberg, E. Solan, E. Vieille (2009), ”Information externalities and emergence of consensus,” *Games and Economic Behavior* 66, 979-994.

[9] A. N. Shiryayev, (1984), “Probability,” *Springer.*

[10] L. Smith, P. Sorensen (2000), “Pathological Outcomes of Observational Learning”, *Econometrica* 68, 371-398.