FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

DOUGLAS BOKLIANG ANG CUNHA

SOCIAL LEARNING, DIVERSITY AND POLARIZATION

SÃO PAULO
2020

DOUGLAS BOKLIANG ANG CUNHA

SOCIAL LEARNING, DIVERSITY AND POLARIZATION

Tese apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para a obtenção do título de Doutor em Economia de Empresas. Campo de conhecimento: Teoria Econômica

Supervisor: Daniel Monte

Co-supervisor: Bruno Ferman

SÃO PAULO

2020

Cunha, Douglas Bokliang Ang.

Social learning, diversity and polarization / Douglas Bokliang Ang Cunha. - 2020.

49 f.

Orientador: Daniel Monte.

Co-orientador: Bruno Ferman.

Tese (doutorado CDEE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Economia - Aspectos psicológicos. 2. Aprendizagem social. 3. Probabilidades. 4. Algoritmos. I. Monte, Daniel. II. Ferman, Bruno. III. Tese (doutorado) – Escola de Economia de São Paulo. IV. Fundação Getulio Vargas. V. Título.

CDU 33

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O

Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

DOUGLAS BOKLIANG ANG CUNHA

SOCIAL LEARNING, DIVERSITY AND POLARIZATION

Tese apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para obtenção do título de Doutor em Economia de Empresas.

Campo de Conhecimento:
Teoria Econômica

Data de Aprovação:
05 / 05 / 2020

Banca examinadora:

Prof. Dr. Daniel Monte
FGV-EESP

Prof. Dr. Bruno Ferman
FGV-EESP

Prof. Dr. Braz Camargo
FGV-EESP

Prof. Dr. Eduardo Faingold
INSPER

Prof. Dr. Felipe Shalders
FEA-USP

Acknowledgements

À Liz, sem a qual eu nunca teria a vontade e a força para ir atrás daquilo que realmente importa para mim.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

Resumo

Esta tese é composta por dois ensaios sobre aprendizagem social, um ramo da teoria econômica. Ambos consideram agentes que precisam escolher, em cada período, uma entre duas ações com resultados desconhecidos. Eles também observam o que um outro agente (escolhido de forma aleatória) decidiu fazer. Nesse modelo, os agentes aprendem o quão boa é uma ação através de suas próprias experiências e pelas observações que fazem dos outros.

No primeiro capítulo, nós estudamos a importância da diversidade em um ambiente social. Mostramos que a sociedade asintoticamente aprende o efeito de cada ação quando os agentes são heterogêneos. Isso não acontece em geral quando os agentes são homogêneos, já que eles eventualmente optam pela mesma ação (possivelmente a pior) e, portanto, observar o que outros escolhem é mais fraco para ajudar a distinguir entre estados da natureza no longo prazo.

No segundo capítulo, nosso objetivo é mostrar que aprendizagem social pode contribuir para polarização de crenças. Se a observação sobre os outros é viesada, os agentes podem acreditar que a proporção de pessoas escolhendo cada opção é diferente daquela verdadeira. Atualmente, esse tipo de ambiente é provável de ocorrer por causa de algoritmos de mídias sociais que selecionam o que os usuários observam. Se os agentes não são conscientes disso, nós mostramos em nosso modelo que a sociedade pode ficar divididas em dois grupos de pessoas, cada um fazendo uma escolha diferente e achando que a outra opção é muito pior do que realmente é.

Palavras-chaves: Economia - Aspectos psicológicos, Aprendizagem social, Probabilidades, Algoritmos

Abstract

This thesis consists of two essays on social learning, a branch from the economic theory field. Both consider agents who need to choose, in each period, one between two actions with unknown expected payoffs. They also observe what another (randomly chosen) agent decided to do. In this model, agents learn how good is each action through their own experience and their observations of others.

In the first chapter, we study the importance of diversity in a social environment. We show that society asymptotically learns the true effect of each action with heterogeneous agents. This does not happen in general when agents are homogeneous, since they eventually opt for the same action (possibly the worst one) and, thus, observing what others choose is weaker to help distinguish between states of the world in the long run.

In the second chapter, our goal is to show that social learning may contribute to polarization of beliefs. If the observation of others are biased, agents will misperceive the true proportion of people choosing each option. Nowadays, this kind of environment is likely to happen because of social media's algorithms that select what users observe. We show that, if agents are not fully aware of this, society may be split in two groups of people, each making a different choice and believing that the other option is much worse than it really is.

Key-words: Economics - Psychological aspects, Social Learning, Probabilities, Algorithms.

Contents

1 Does diversity improve social learning?

1.1 Introduction

In an environment where people have to decide what to do without knowing the result of each action, they must consider not only what they believe to be the option with the highest payoff, but also what they learn by taking that action. This trade-off between current payoff and information acquisition has been studied by the experimentation literature[1]. Another way of getting information, in many situations, is by observing others, since others' actions reveal their own experiences. Examples include a student observing that his colleagues already started studying to the next exam, an investor observing in which assets others are investing and a voter observing the election polls. However, this channel may not be very informative when agents are heterogeneous[2]: it is harder to tell whether an action is observed because it is good, it may be because the other agent has different preferences. An investor, for instance, may believe that someone is investing in a risky asset because this person is less risk-averse, and not because that asset has great expected returns.

This paper studies social learning in this kind of situation. We want to know whether agents eventually choose the best choice for themselves when society is heterogeneous. We find that, in equilibrium, the answer is yes. Moreover, we show that agents learn the true payoff distribution of each alternative in the long run. These results highlight the importance of diversity in a social environment.

Our approach is close to that of Aoyagi (1998) and Camargo (2014). Both of them show that, in equilibrium, all players eventually settle on the same alternative, although not necessarily on the best one. This possibility of all agents dropping the best choice is called "Rothschild effect".[3] Aoyagi (1998) studies a two-armed bandit model with finite homogeneous players who observe each other's actions in every period. Camargo (2014) considers a multi-armed bandit model with a continuum of homogeneous players that observe, in each period, the action of another randomly chosen player, which he refers to as *observation in society*. Moreover, he derives a sufficient condition on the distribution of priors beliefs so that the society overcomes the "Rothschild effect", that is, players eventually settle on the best alternative.

[1] See Hörner e Skrzypacz (2017) for a survey.
[2] For instance, Munshi (2004), with data from the Indian Green Revolution, shows that rice growers respond less to neighbors' experience than wheat growers since rice-growing regions are more heterogeneous in growing conditions and rice varieties are more sensitive to unobserved farm characteristics.
[3] Due to the work of Rothschild (1974).

Besides the fact that we consider a two-armed bandit model[4], our main difference from that of Camargo (2014) is that, in our environment, players are heterogeneous in the sense that they may have different payoff distribution for the same action. One could either reason that agents have different preferences for the same results or that they have different results for the same action. Our finding that players eventually settle on the correct arm for themselves may seem counter-intuitive, especially after Aoyagi (1998) and Camargo (2014) show that, in general, homogeneous players do not take the best arm. However, in a society with heterogeneity, players have another source of high quality information: asymptotic distribution of actions. Although players cannot observe the distribution of actions, observations in society allow them to infer about the proportion of players taking each action. States of the world in which asymptotic proportions are different from the proportions in the true state are asymptotically disregarded. In an environment with homogeneous players, many states of the world have the same asymptotic distribution of actions, so this source of information is much less informative. This is the main ingredient that makes diversity of agents overcome the "Rothschild effect".

Although with objectives different from ours, social learning with heterogeneous agents has already been studied in different frameworks. Ellison e Fudenberg (1993) consider, for example, that agents are heterogeneous in the sense that each one has a different (real-valued) parameter. In each period, agents can choose a technology and observe choice and payoff from those with parameters close to theirs. This parameter also affects payoffs, which makes the best technology not to be the same for everyone. Among other results, Ellison e Fudenberg (1993) show that, in general, agents do not eventually take the best alternative for themselves. A key difference from our paper is that the learning process in their environment do not follow Bayes' rule, but an exogenous rule that do not consider all the past experiences.

On the other hand, Bala e Goyal (2001) and Young (2009) consider agents that update their beliefs with Bayes' rule. One difference from our paper is that agents are myopic in their environments. Bala e Goyal (2001) consider a network in which agents can observe the actions and outcomes of the ones connected to them. Heterogeneity comes from the fact that there are two types of agents, with different preferences. In general, agents do not take the best action for themselves in the long run. Young (2009) considers that individuals have different costs to adopt a new technology with unknown payoff. Different from us, he is interested in studying the diffusion of such technology in the short term.

Finally, Smith e Sørensen (2000) consider finite types of rational agents. They follow the framework introduced in Banerjee (1992) and Bikhchandani, Hirshleifer e Welch (1992), considering short-lived individuals who decide sequentially, while we consider long-

4 In section 1.6, we also study the case with a three-armed bandit.

lived agents. In the environment of Smith e Sørensen (2000), players do not necessarily converge to the best alternative for themselves.

We believe our paper contributes to the literature especially in two ways: first, we study social learning in a environment with heterogeneous agents that are fully rational and long-lived; second, our results highlight the importance of diversity in social environments.

1.2 Example

In this section, we adapt the example of Camargo (2014) to illustrate our main results. Suppose there is a continuum I of agents. In the beginning of each period, they have a task and have to choose between making effort ($k = 2$) or not ($k = 1$). They know they succeed in the task with probability p_1 without making any effort. But they do not know the probability p_2 of succeeding in the task after making effort, they just know it must be in the set $\{p_2^1, p_2^2, p_2^3\}$, where $0 < p_2^1 < p_2^2 < p_2^3 < 1$. Agents get payoff 1 after succeeding in the task; after a failure in the task, they get payoff 0. In every period, each agent observes what another randomly chosen anonymous agent decided to do. Suppose that agents' cost to make effort follows a uniform distribution $U(\underline{c}, \bar{c})$ with cdf G, where $\underline{c} < p_2^1 - p_1$ and $\bar{c} > p_2^3 - p_1$, which is common knowledge. This means that some agents want to make effort ($k = 2$) even if the effect of effort is the lowest possible, p_2^1, and some agents do not want to make effort ($k = 1$) even if the effect of effort is the highest possible, p_2^3. Let $c(i)$ be the cost of agent $i \in I$. Players discount the future by a factor $\delta \in [0, 1)$ per period.

The set of states of the world is denoted by $\Theta = \{p_2^1, p_2^2, p_2^3\}$ such that $p_2 = \theta$ if the state is $\theta \in \Theta$. Initial priors about the state of the world are common knowledge and may also be heterogeneous. Let $\pi_t^i(\theta)$ be the probability player i assigns to state θ at period t. We assume that $\eta := \inf\{\pi_1^i(\theta) : i \in I, \theta \in \Theta\} > 0$, that is, players assign probability strictly positive, and not arbitrarily close to 0, for any state of the world at the beginning of the game.

Let Σ be the set of possible strategies. A strategy profile $F : I \to \Sigma$ is a $\mathcal{I}$-measurable function such that $F(i) \in \Sigma$ maps any possible history to a (possibly mixed) action to player $i \in I$. When F is the strategy profile under play, the mass of players who play $k \in \{1, 2\}$ in period t when the state of the world is θ is deterministic and well defined. We denote it by $m_t(k, \theta)$. Although individual actions are stochastic, the mass of players choosing each action is not. Observe that $m_t(k, \theta)$ also gives the probability of observing an agent who plays k in period t when the state of the world is θ. We refer to m_t as *observation likelihood* in period t. In equilibrium, the sequence $m = \{m_t\}$ is known and enables players to update their beliefs after observations in society using Bayes' rule.

We begin following some results of Camargo (2014). First, players who play $k = 2$ infinitely often asymptotically discover the true state θ, by the Strong Law of Large Numbers; second, players eventually behave myopically; third, players eventually choose only one action (they do not keep changing actions forever); fourth, $\{m_t(\theta)\}$ is convergent for any $\theta \in \Theta$. We define $m_\infty(\theta)$, for each $\theta \in \Theta$, such that $m_t(\theta) \to m_\infty(\theta)$.

Next we show that almost all players eventually take the best action for themselves. Stronger than that, we show that they asymptotically learn the true state of the world.

Assume $\theta \in \Theta$ is the true state of the world. Let I_k be the set of players who play $k \in A$ infinitely many times. Let $\mathbb{E}_{\pi_t^i}$ be the expected value operator according to player i's beliefs in period t. Using the first result of Camargo (2014), we get that $\mathbb{E}_{\pi_\infty^i}[p_2] = \theta$, for almost all $i \in I_2$.

Since they eventually behave myopically, for almost all $i \in I_2$, $\mathbb{E}_{\pi_\infty^i}[p_2] - c(i) = \theta - c(i) \geq p_1$. Then, almost all players with cost greater than $\theta - p_1$ must be in I_1:

$$c(i) > \theta - p_1 \Rightarrow i \in I_1. \tag{1.1}$$

This means that players with high cost ($c(i) > \theta - p_1$) must choose the best action for themselves, which is action $k = 1$. We want to show that players with low cost ($c(i) < \theta - p_1$) will choose action $k = 2$. Equivalently, we want to show that almost all players who choose action $k = 1$ are players with high cost.

Since players eventually behave myopically, for all $i \in I_1$, $\mathbb{E}_{\pi_\infty^i}[p_2] - c(i) \leq p_1$. Then

$$i \in I_1 \Rightarrow c(i) \geq \mathbb{E}_{\pi_\infty^i}[p_2] - p_1. \tag{1.2}$$

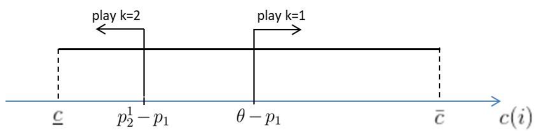
Since p_2^1 is the lowest state possible, we can infer that

$$i \in I_1 \Rightarrow c(i) \geq p_2^1 - p_1. \tag{1.3}$$

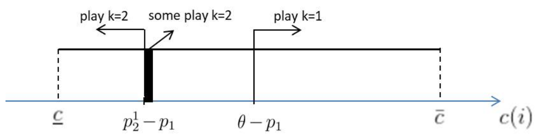
Assume $\theta = p_2^1$ is the true state of the world. Equations 1.1 and 1.3 imply that almost all players eventually take the best action for themselves. Then $m_\infty(2, p_2^1) = G(p_2^1 - p_1)$. Moreover, next we argue that $m_\infty(2, \theta) \neq m_\infty(2, p_2^1)$ if $\theta > p_2^1$.

When $\theta > p_2^1$ is the true state, equations 1.1 and 1.3 do not define the actions of players with cost between $p_2^1 - p_1$ and $\theta - p_1$ (figure 1.1a). Next we show that a strict positive mass of players with costs sufficiently close to $p_2^1 - p_1$ must play $k = 2$ (figure 1.1b). Consider $I(\varepsilon) := \{i \in I : p_2^1 - p_1 \leq c(i) \leq p_2^1 - p_1 + \varepsilon\}$, for $\varepsilon > 0$, the set of players with cost higher than $p_2^1 - p_1$, but smaller than $p_2^1 - p_1 + \varepsilon$. Assume almost all $i \in I(\varepsilon)$ play $k = 1$ infinitely often. Equation 1.2 and the definition of $I(\varepsilon)$ imply that $p_2^1 - p_1 + \varepsilon \geq c(i) \geq \mathbb{E}_{\pi_\infty^i}[p_2] - p_1$. Taking ε close enough to 0, $\mathbb{E}_{\pi_\infty^i}[p_2]$ must be close enough to p_2^1, which implies that $\pi_\infty^i(p_2^1)$ must be close enough to 1 and $\pi_\infty^i(\theta)$ close enough to

0. However, since players update their beliefs with Bayes' Rule, $\pi_t^i(\theta)$ is increasing in average since θ is the true state.[5] In other words, it cannot be the case that almost all players $i \in I(\varepsilon)$ have $\pi_\infty^i(\theta)$ arbitrarily close to 0, since $\pi_1^i(\theta) \geq \eta > 0$ for all $i \in I$. Hence, there is a strict positive mass of players in $I(\varepsilon)$ that must play $k = 2$. Therefore, the total mass of players that eventually chooses $k = 2$ is strictly greater in state θ: $m_\infty(2, \theta) > G(p_2^1 - p_1) = m_\infty(2, p_2^1)$.



(a) From equations (1.1) and (1.3).



(b) Sufficiently close to $p_2^1 - p_1$, a positive mass of players must play $k = 2$.

Figure 1.1 – Players' choice when $\theta > p_2^1$.

With infinite many observations in society, almost every player can tell that the proportion of players choosing $k = 2$ converges to $m_\infty(2, p_2^1)$ and not to $m_\infty(2, \theta)$, $\theta > p_2^1$, when the true state is p_2^1. That's why we say that the state p_2^1 is identified: players know when the true state is p_2^1 and when it is not.

Now, assume $\theta = p_2^2$ is the true state. Almost all players asymptotically assign probability arbitrarily close to 0 to $p_2 = p_2^1$ since it is identified. Therefore $\mathbb{E}_{\pi_\infty^i}[p_2] \geq p_2^2$ and (1.2) would become

$$i \in I_1 \Rightarrow c(i) \geq \mathbb{E}_{\pi_\infty^i}[p_2] - p_1 \geq p_2^2 - p_1. \tag{1.4}$$

Equations 1.1 and 1.4 imply that almost all players eventually take the best action for themselves. Then $m_\infty(p_2^2) = G(p_2^2 - p_1)$. We follow the same argument from before to say that $\theta = p_2^2$ is identified.

Finally, assume the true state is $\theta = p_2^3$. Almost all players eventually assign probability arbitrarily close to 0 to $p_2 = p_2^1$ and $p_2 = p_2^2$ since they are identified states. Then p_2^3 is also identified and almost all players also take the best actions for themselves in this state.

With this example in mind, it is important to highlight what is really driving our results: asymptotic distributions of actions are different across states, which allow players to asymptotically learn the true state of the world through their infinite many observations

[5] $\pi_t^i(\theta)$ is a submartingale conditioned on the true state θ.

in society. In a homogeneous society, Camargo (2014) shows that players converge to the same action. That means there are much less possible asymptotic distributions: only the same number of actions. If our example had homogeneous agents, the asymptotic distribution of actions would be either almost all players choosing $k = 1$ or almost all of them choosing $k = 2$. With three possible states of the world, at least two would have the same asymptotic distribution, and players would not be able to differentiate these states with observations in society. That is the main aspect that makes learning in a heterogeneous society asymptotically better: observations in society are more informative.

There are two important factors that make the asymptotic distributions of actions be different across states: first, there must be a positive mass of players close to the indifference between actions under complete information, for any possible state of the world; second, players update their beliefs with Bayes' rule. Players close to the indifference between actions make different choices depending on the state of the world (figure 1.1b). Comprehending this allows us to understand that the same results would obtain with other specifications that make sense economically, for instance: players could have more than one observation per period, they could also observe other's outcomes, agents could not be anonymous and there could be a (not perfect) correlation between player's action or cost and observed action. None of these aspects change the fact that asymptotic distributions of actions are different across states.

Next session formalizes and generalizes our model.

1.3 Model

A continuum of anonymous players is identified with a probability space $(I, \mathcal{I}, \lambda)$. In every period $t \in \mathbb{N}$, they have to choose one action in the set $A = \{1, 2\}$. The set of states of the world is $\Theta = \Theta_1 \times \Theta_2$, where, for each $k \in A$, $\Theta_k = \{\theta_k^1, \theta_k^2, ..., \theta_k^{n_k}\}$ for some $n_k \in \mathbb{N}$. We define a linear order on Θ_k such that $\theta_k^1 < \theta_k^2 < ... < \theta_k^{n_k}$ and a partial order on Θ such that $\theta \geq \theta'$ if, and only if, $\theta_1 \geq \theta_1'$ and $\theta_2 \geq \theta_2'$. For each agent $i \in I$, let Y^i denote the set of his finite possible payoffs.[6] He gets payoff y with probability $g^i(y|k, \theta_k)$ when he chooses action k and the state of the world is $\theta = (\theta_1, \theta_2)$. Players are heterogeneous since Y^i and g^i may be different for each $i \in I$. Player i's expected payoff is denoted by $r_k^i(\theta_k)$, for each $k \in A$ and $\theta_k \in \Theta_k$. We assume that, for each state $\theta \in \Theta$,

$$\forall \varepsilon > 0, \lambda\{0 < r_k^i(\theta_k) - r_{k'}^i(\theta_{k'}) < \varepsilon\} > 0, \text{ for } k, k' \in A \text{ s.t } k \neq k', \tag{A1}$$

and that

$$\lambda\{r_1^i(\theta_1) = r_2^i(\theta_2)\} = 0. \tag{A2}$$

[6] Although we could consider infinite many different payoffs without changing the results, we choose finite payoffs to ease notation. In the example of section 1.2, $Y^i = \{0, 1, -c(i), 1 - c(i)\}$ for each $i \in I$.

For each state, assumptions A1 and A2 imply that, under complete information, there is a positive mass of players arbitrarily close to the indifference between the two actions, but there is no mass of players exactly in the indifference. Assumption A1 is important to make asymptotic distributions different depending on the state of the world, and assumption A2 is important for the convergence of the observation likelihood (Lemma 1.4).

We also assume that r_k^i is strictly increasing. Stronger than that, we consider, for each $k \in A$,

$$\inf\{r_k^i(\theta_k) - r_k^i(\theta_k') : \theta_k, \theta_k' \in \Theta_k \text{ s.t. } \theta_k > \theta_k', i \in I\} > 0. \tag{A3}$$

Assumption A3 implies that agents are not arbitrarily close to the indifference between states θ and θ' with $\theta_k \neq \theta_k'$ when they choose $k \in A$.[7]

Let $\Pi = \Delta(\Theta)$ be the set of possible beliefs about the state of the world. Let $\phi : I \to \Pi$ be such that $\phi(i) \in \Pi$ is the initial prior of player $i \in I$. Denote $\pi_t^i(\theta)$ the probability player i assigns to $\theta \in \Theta$ in period t and $\pi_t^i(k, \theta_k) := \sum_{\{\theta' : \theta_k' = \theta_k\}} \pi_t^i(\theta')$ the probability he assigns to $\theta_k \in \Theta_k$. We assume that players do not assign a probability arbitrarily close to 0 to any state in the beginning of the game:

$$\eta := \inf\{\pi_1^i(\theta) : i \in I, \theta \in \Theta\} > 0. \tag{A4}$$

In a given period t, a player i chooses an action $k \in A$, observes its outcome $y \in Y^i$ and, finally, a choice $\tilde{k} \in A$ of another randomly chosen anonymous player. The set of histories in period t is $H_t^i = (A \times Y^i \times A)^{t-1}$. The set of infinite histories is $H_\infty^i = (A \times Y^i \times A)^\infty$. A strategy for player i is a sequence $\sigma = \{\sigma_t\}$ such that $\sigma_t : H_t^i \to \Delta(A)$ maps any history in H_t^i to an action (possibly mixed) in period t. Let Σ^i be the set of all possible strategies for player $i \in I$ and define $\Sigma := \cup_{i \in I} \Sigma^i$. A strategy profile $F : I \to \Sigma$ is a $\mathcal{I}$-measurable function[8] that maps each player $i \in I$ to a strategy $F(i) \in \Sigma^i$. The set of all possible strategy profiles is denoted by $\mathcal{F}$.

Given a strategy profile F, the proportion of players choosing each action $k \in A$ at period $t \in \mathbb{N}$ when the state of the world is $\theta \in \Theta$ is well defined. We denote it by $m_t(k, \theta)$ and define $m_t(\theta) := (m_t(1, \theta), m_t(2, \theta))$.[9] Let $\mathcal{M}$ be the set of all possible sequences $m = \{m_t\}$ from $A \times \Theta$ into $[0, 1]$. We denote by $M : \mathcal{F} \to \mathcal{M}$ the map such

---

7 Assuming only that r_k^i is strictly increasing for each $k \in A$ and $i \in I$ is not enough. To see that, consider states θ and θ' such that $\theta_k > \theta_k'$ and $\theta_{\bar{k}} = \theta_{\bar{k}}'$, $\bar{k} \neq k$. Consider that either $r_k^i(\theta_k') - r_{\bar{k}}^i(\theta_{\bar{k}}) < r_k^i(\theta_k) - r_{\bar{k}}^i(\theta_{\bar{k}}) < 0$ or $0 < r_k^i(\theta_k') - r_{\bar{k}}^i(\theta_{\bar{k}}) < r_k^i(\theta_k) - r_{\bar{k}}^i(\theta_{\bar{k}})$ for each $i \in I$. Different asymptotic distributions of actions are important in our model, but this example shows they could be the same in states θ and θ' (under complete information) without breaking strictly monotonicity of r_k^i.

8 F must be a $\mathcal{I}$-measurable function so we can aggregate individual actions.

9 Although individual behavior may be stochastic, aggregate is not. Appendix A.2 of Camargo (2014) shows how to aggregate individual behavior to find $\{m_t\}$.

that $m = M(F)$ gives us the proportion of players choosing each action for each state of the world and period when the strategy profile is F.

Fix a player $i \in I$ and his strategy $\sigma \in \Sigma^i$. His (possibly mixed) action is defined by strategy σ in the first period, θ defines the outcome distribution $g^i(y|k,\theta)$, and $m_1(k,\theta)$ the probability of observing a player who plays k in period $t = 1$. Hence, σ, θ and m define a probability distribution on H_1. For each $h_1 \in H_1$, σ defines the player's action in the second period, and so on. Therefore, σ, θ and m define a probability distribution on H_∞, which we denote by $\mu(\sigma|\theta,m)$. Since i does not know θ, he considers his prior $\pi \in \Pi$ with $\mu(\sigma|\theta,m)$ to find a distribution probability on $\Theta \times H_\infty$, which we denote by $\mu(\sigma|\pi,m)$.

Let y_t^i be player i's stochastic payoff at period t and $R^i = \sum_{t=1}^{\infty} \delta^{t-1} y_t^i$, where $\delta \in [0,1)$ is the discount factor. Given m, we denote the individual learning problem in which player i has prior π_1 by $ILP^i(\pi_1,m)$. An optimal experimentation strategy σ^* for $ILP^i(\pi_1,m)$ is such that

$$\mathbb{E}_{\mu(\sigma^*|\pi_1,m)}[R^i] = \sup_{\sigma\in\Sigma^i} \mathbb{E}_{\mu(\sigma|\pi_1,m)}[R^i]. \tag{1.5}$$

Hence, the observation likelihood m affects optimal strategies and a strategy profile F defines the observation likelihood through $m = M(F)$. This is the idea behind the Nash equilibrium for non-atomic games we adapt to our environment, following Camargo (2014).

Definition 1.1. *An equilibrium is a pair (m^*, F^*) such that $F^*(i)$ is an optimal experimentation strategy for $ILP^i(\phi(i), m^*)$ for λ-almost all $i \in I$ and $m^* = M(F^*)$.*

Although we make some asymptotic characterizations about possible equilibria, their existence is difficult to prove and beyond the scope of this paper.[10]

1.4 Results

We begin with four lemmas of Camargo (2014). Let $I_k \in I$ be the set of player who play $k \in A$ infinitely many times. Let F_k^i be the event such that $i \in I_k$. If i plays k infinitely many times, he will have infinitely many outcome observations and, by the Strong Law of Large Number (and assumption A3), will asymptotically learn the true expected payoff of action k with probability 1. This is what Lemma 1.1 states. The proof is omitted since it is essentially that of Lemma 1 of Aoyagi (1998).

Lemma 1.1. *Assume θ is the true state of the world. Suppose the sequence of observation likelihoods is m. Consider a player $i \in I$ with prior π_1^i and strategy σ^i. If $\mu(\sigma^i|\pi_1^i,m)(F_k^i) > 0$, then $\mu(\sigma^i|\pi_1^i,m)\left(\lim_{t\to\infty} \pi_t^i(k,\theta_k) = 1 | F_k^i\right) = 1$.*

[10] See Appendix B of Camargo (2014) for a proof of existence of equilibria in an environment with homogeneous players.

Consider a player i with belief π. Let $\mathbb{E}_\pi[r_k^i] := \sum_{j=1}^{n_k} \pi(k, \theta_k^j) r_k^i(\theta_k^j)$ be his current expected outcome when he plays k. Denote $BR^i(\pi)$ the set of his myopically optimal actions: $k \in BR^i(\pi) \Leftrightarrow k \in \arg\max_k \mathbb{E}_\pi[r_k^i]$.

Since beliefs π_t^i are martingales, they converge almost surely to a (random) limit π_∞^i. Let A_∞^i be the (random) set of actions player i chooses infinitely many times. Lemma 1.2 states that almost all players will eventually choose actions that are myopically optimal according to their beliefs. Intuitively, players do not want to experiment forever, since they experiment when they want to acquire information for the future, giving up current expected payoff. This result is a consequence of Proposition 2.1 of Rosenberg, Solan e Vieille (2009), and we state without proof.

Lemma 1.2. *Suppose the sequence of observation likelihoods is* m *and* σ^i *is an optimal strategy. Then* $\mu(\sigma^i|\pi_1^i, m)(A_\infty^i \subseteq BR^i(\pi_\infty^i))) = 1$ *for each player* $i \in I$.

Consider a player $i \in I_1 \cap I_2$. Lemma 1.1 implies that he asymptotically learns the true state of the world. Lemma 1.2 requires that $\{1, 2\} \subseteq BR(\pi_\infty^i)$ and, therefore, $r_1^i(\theta_1) = r_2^i(\theta_2)$ with probability 1. By assumption A2, we can state our next result.

Lemma 1.3. *Suppose almost all players follow optimal strategies. Then the mass of players that choose two actions infinitely many times is zero.*

If the mass of players who choose $k = 1$ did not converge, there would be a positive mass of players alternating their choice infinitely many times, which cannot occur, according to Lemma 1.3. This result is our Lemma 1.4, which comes from Lemma 6 of Camargo (2014) and, therefore, we state without proof.

Lemma 1.4. *Let* (m^*, F^*) *be an equilibrium. Then* $\{m_t^*(\theta)\}$ *is convergent for all* $\theta \in \Theta$.

We define $m_\infty(\theta)$, for each $\theta \in \Theta$, such that $m_t(\theta) \to m_\infty(\theta)$.

Assume θ is the true state. Lemma 1.4 states that the fraction of players choosing action k converges to $m_\infty(k, \theta)$. Since players make infinitely many observations in society, the fraction of players they observe choosing k must also converge to $m_\infty(k, \theta)$. Lemma 1.5 below states that players will asymptotically know that the true state cannot be any state θ' such that $m_\infty(k, \theta') \neq m_\infty(k, \theta)$. For a formal proof, see Appendix A.1.

Lemma 1.5. *Let* (m^*, F^*) *be an equilibrium and* $\theta, \theta' \in \Theta$ *such that* $m_\infty^*(\theta') \neq m_\infty^*(\theta)$. *Then* $\mu(F^*(i)|\theta, m^*)(\pi_\infty^i(\theta') = 0) = 1$, *for* λ*-almost all* $i \in I$.

Before next result, it is appropriate to establish some definitions. Suppose that, when the true state is θ, almost all players asymptotically discover that the state is not θ', and vice-versa. In this case, we say that θ is distinguishable from θ'.

Definition 1.2. *Let (m^*, F^*) be an equilibrium. We say θ is distinguishable from θ', and vice-versa, if*

1. $\mu(F^*(i)|\theta, m^*)(\pi^i_\infty(\theta') = 0) = 1$ *and*
2. $\mu(F^*(i)|\theta', m^*)(\pi^i_\infty(\theta) = 0) = 1$,

for λ-almost all $i \in I$.

Next we provide a stronger definition.

Definition 1.3. *θ is identified if θ is distinguishable from θ', for each $\theta' \neq \theta$.*

We denote by ID the set of identified states.

Lemma 1.6 provides a sufficient condition so that two states are distinguishable from each other. Assume θ is the true state. Because of Lemma 1.1, almost all players asymptotically discover that a state θ' such that $\theta_1 \neq \theta'_1$ and $\theta_2 \neq \theta'_2$ cannot be the true state. Because of Lemma 1.5, almost all players asymptotically learn that the state cannot be θ' such that $m_\infty(\theta) \neq m_\infty(\theta')$.

Lemma 1.6. *Assume (m^*, F^*) is an equilibrium. Let $\theta, \theta' \in \Theta$. If at least one of the following conditions is true:*

1. $\theta_k \neq \theta'_k$, *for each $k \in A$;*
2. $m^*_\infty(\theta) \neq m^*_\infty(\theta')$.

Then θ is distinguishable from θ'.

Let $B^i_\infty(k, \tilde{\theta}_k)$ be the event such that player i asymptotically believes the true state θ is such that $\theta_k \geq \tilde{\theta}_k$, that is, $\pi^i_\infty(\theta') = 0$, for each θ' such that $\theta'_k < \tilde{\theta}_k$.

We follow making a definition with regard to beliefs:

Definition 1.4. *Let (m^*, F^*) be an equilibrium. We say players have efficient beliefs in θ if $\mu(F^*(i)|\theta, m^*)(B^i_\infty(k, \theta_k)) = 1$, for each $k \in A$ and λ-almost all $i \in I$.*

Intuitively, players do not underestimate the effect of any action if they have efficient beliefs. We denote by EB the set of states in which players have efficient beliefs. It is straightforward to see that $ID \subseteq EB$.

A player i who plays k infinitely often and do not underestimate $\theta_{k'}$, $k' \neq k$, certainly eventually chooses the best action for himself: he asymptotically learns θ_k (Lemma 1.1) and chooses k even though he can only overestimate the payoff of k'. Suppose almost

all players do not underestimate both θ_1 and θ_2 in the long run. Then, they must eventually play the best action for themselves. Next lemma is about this fact, which inspires the name used in the definition above.

Let $BR^i(\theta)$ denote the set of player i's best actions when the state is θ.

Lemma 1.7. *Assume θ is the true state and players follow optimal strategies. If $\theta \in EB$, then $A^i_\infty \subseteq BR^i(\theta)$, for λ-almost all players $i \in I$.*

Assume θ is the true state and $\underline{\theta} \in ID$, for each $\underline{\theta} < \theta$. Lemma 1.8 below asserts θ is also distinguishable from states $\bar{\theta} > \theta$, that is, $\theta \in ID$.[11]

First we show $\theta \in EB$. Consider players who play $k = 1$ infinitely often. Lemma 1.1 implies that they asymptotically learn θ_1. Since they distinguish θ from any $\underline{\theta} < \theta$, they cannot asymptotically underestimate action $k = 2$. Analogously, players who play $k = 2$ also do not asymptotically underestimate action $k = 1$, thus $\theta \in EB$.

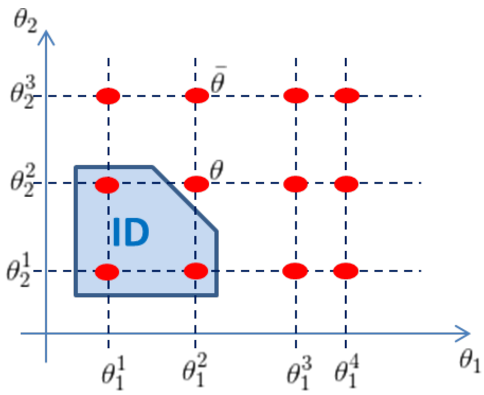


Figure 1.2 – $\theta = (\theta_1^2, \theta_2^2)$ is the true state and $\underline{\theta} \in ID$ for all $\underline{\theta} < \theta$.

Now we argue that θ is distinguishable from any $\bar{\theta} > \theta$. Suppose, for instance, θ is not distinguishable from $\bar{\theta}$ such that $\bar{\theta}_1 = \theta_1$ and $\bar{\theta}_2 > \theta_2$ (see figure 1.2), then $m_\infty(\bar{\theta}) = m_\infty(\theta)$ by Lemma 1.6. Since almost all players eventually play correctly when the state is θ (Lemma 1.7), a strictly positive mass who would be better off playing $k = 2$ when the state is $\bar{\theta}$ have to play $k = 1$ infinitely often. When the true state is $\bar{\theta}$, players asymptotically discover the true state cannot be states $\underline{\theta} < \theta$ since they are identified. We show that a strictly positive mass (players close to indifference between $k = 1$ and $k = 2$ when the state is θ who choose $k = 1$) have to assign probability arbitrarily close to 0 to the true state $\bar{\theta}$ in the long run. However, since beliefs in the true state are submartingales, this cannot happen. This means that $m_\infty(2, \bar{\theta}) > m_\infty(2, \theta)$ and, therefore, θ is also distinguishable from a state $\bar{\theta} > \theta$ (Lemma A.1). Hence, $\theta \in ID$. A formal proof can be found in Appendix A.2.

[11] Note that states $\tilde{\theta}$ such that neither $\tilde{\theta} < \theta$ nor $\tilde{\theta} > \theta$ are distinguishable from θ because both $\tilde{\theta}_1 \neq \theta_1$ and $\tilde{\theta}_2 \neq \theta_2$.

Lemma 1.8. *Consider an equilibrium and let $\theta \in \Theta$. Assume $\forall \underline{\theta} \in \Theta$, $\underline{\theta} < \theta \Rightarrow \underline{\theta} \in ID$. Then $\theta \in ID$.*

Lemma 1.8 implies that $(\theta_1^1, \theta_2^1) \in ID$. With a simple induction argument, Lemma 1.8 is the key to prove Theorem 1.1, our main result.

Theorem 1.1. *In equilibrium, $\theta \in ID$, for each $\theta \in \Theta$.*

Theorem 1.1 shows that players asymptotically learn the true state of the world. Corollary 1.1 asserts that players eventually choose the best action for themselves.

Corollary 1.1. *In equilibrium, $A^i_\infty \subseteq BR^i(\theta)$, for each $\theta \in \Theta$ and λ-almost all players $i \in I$.*

1.5 Discussion

In this section we discuss the important assumptions for efficiency and compare our results with those of Aoyagi (1998) and Camargo (2014).

First, we adapt the definition of efficiency of Camargo (2014):

Definition 1.5. *In an equilibrium, social learning is efficient when the state is θ if $A^i_\infty \subseteq BR^i(\theta)$ for λ-almost all $i \in I$.*

Consider an environment where players have all the same payoff distribution $g(y|k, \theta)$. Then, Aoyagi (1998) and Camargo (2014) show that social learning is not efficient in general. Suppose there are only two states, $\Theta = \{\theta, \theta'\}$, such that $r_1(\theta) > r_2(\theta)$ and $r_2(\theta') > r_1(\theta')$ for all players. Suppose, for each action $k \in \{1, 2\}$, there is a positive mass of players who myopically believe, initially, that k is the best action for them: $\lambda\{i \in I : \mathbb{E}_{\pi^i_1}[r_1] > \mathbb{E}_{\pi^i_1}[r_2]\} > 0$ and $\lambda\{i \in I : \mathbb{E}_{\pi^i_1}[r_2] > \mathbb{E}_{\pi^i_1}[r_1]\} > 0$. Camargo (2014) shows that this condition is sufficient to make social learning efficient.

Our environment with heterogeneous players satisfies an adapted version of this condition, when we replace r_k by r^i_k since expected payoffs are different. However, it is important to highlight that this is not what drives our results. Example 1.1 shows this adapted condition is not sufficient with heterogeneous players.

Example 1.1. *Let $\Theta_1 = \{p_1\}$ and $\Theta_2 = \{p_2^1, p_2^2\}$. Similar to the example in Section 1.2, assume player i has cost $c(i)$ to choose $k = 2$. Let p_k denote the probability of success after choosing k. Players get payoff 1 after a success, and 0 after a failure. Let $L := \{i \in I : c(i) < p_2^1 - p_1\}$ and $H := \{i \in I : c(i) > p_2^2 - p_1\}$ be the sets with, respectively, low-cost and high-cost players. The adapted "sufficient" condition would hold if $\lambda(L)$ and $\lambda(H)$ are strictly positive, but social learning could still be inefficient. Imagine every player in*

$M := I \setminus (L \cup H)$ has the same cost c such that $p_2^1 - p_1 < c < p_2^2 - p_1$. If all players in M believe playing $k = 1$ is better in $t = 1$ and decide not even experiment action $k = 2$, their outcomes will not be informative about the state of the world. Moreover, observations in society are also not informative since the fraction of players who choose a certain action does not depend on θ. Hence, players in M always play $k = 1$ even if $\theta_2 = p_2^2$.

To guarantee efficiency, assumption A1 plays an important role. It makes the distribution of players choosing each action different depending on the state,[12] which enables players to have more information and discover the true state of the world.

We consider this assumption somewhat natural for an environment with a continuum of heterogeneous players. It might even be milder than some assumptions from the literature. Aoyagi (1998), for example, discusses his assumption that actions have different probability of success. Although we may be interested in situations when this happens, it would be desirable that players *learn* that actions have different payoffs. Another example is the sufficient condition for efficiency of Camargo (2014), which requires heterogeneous priors. If there are many states of the world, players will have to have very different priors.

1.6 Extension

In this section we extend our model to an environment where agents can choose among three actions. We explain why the proof we presented for two actions cannot be used and how to adapt it to the new setting. We still could not extend the results for any finite number of actions.

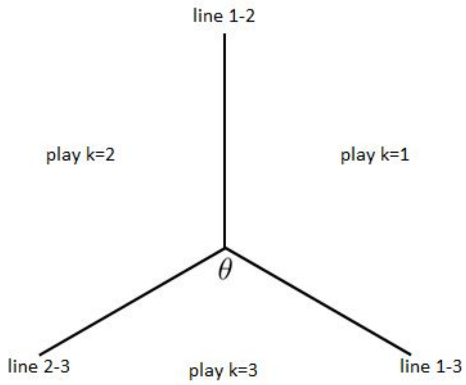
We assume the same setting as before, except that there are three possible actions, $A = \{1, 2, 3\}$, and then the state of the world is an element of $\Theta = \Theta_1 \times \Theta_2 \times \Theta_3$. It is a three-armed bandit model. We show that our main result, Theorem 1.1, still holds.

The proof is more difficult since we cannot use Lemma 1.8. When the true state of the world is θ and $\underline{\theta} \in ID$ for each $\underline{\theta} < \theta$, it is not straightforward to see that $\theta \in EB$, as it was with two actions. For instance, consider $i \in I_1$. Player i asymptotically learns θ_1 (Lemma 1.1), but θ could still not be distinguishable from, for instance, θ' such that $\theta'_1 = \theta_1$, $\theta'_2 > \theta_2$ and $\theta'_3 < \theta_3$. Therefore, player i could asymptotically underestimate action 3. Lemma 1.9 below provides sufficient conditions with regard to beliefs of players such that θ is distinguishable from such θ'.

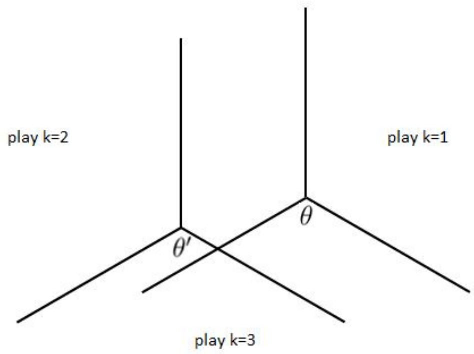
For intuition, consider $\theta = (\theta_1, \theta_2, \theta_3)$. Figure 1.3a represents players as points in $\mathbb{R}^2$. Lines 1-2, 1-3 and 2-3 split the players according to the best action for them. For

[12] To be more precise, it is possible that two states $\theta' \neq \theta''$ are such that the fraction of players choosing a certain action is asymptotically the same. However, this would require $\theta'_1 \neq \theta''_1$ and $\theta'_2 \neq \theta''_2$ and, since players asymptotically learn either θ_1 or θ_2 from the true state $\theta = (\theta_1, \theta_2)$, they would asymptotically disregard at least one of these states.

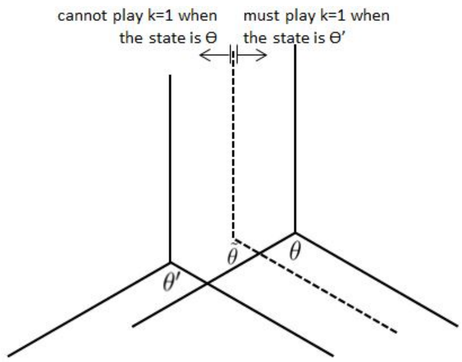
instance, players above line 1-3 and on the right side of line 1-2 would prefer to play $k = 1$ if they knew that the state was θ. The closer to the lines, the more indifferent players are between two actions. Now, let $\theta' = (\bar{\theta}_1, \underline{\theta}_2, \underline{\theta}_3)$ such that $\bar{\theta}_1 > \theta_1$, $\underline{\theta}_2 < \theta_2$ and $\underline{\theta}_3 < \theta_3$. If players could observe the true state, certainly there would be more players choosing $k = 1$ when the state was θ' than when the state was θ (figure 1.3b[13]).



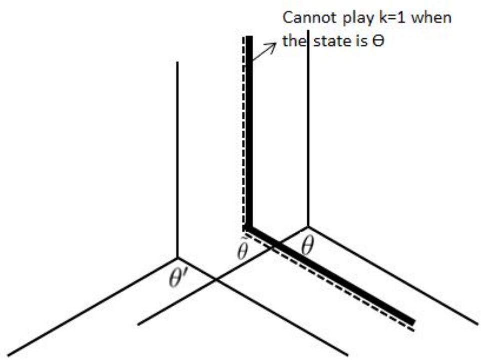
(a) Players' choice under complete information.



(b) Players' choice under complete information in θ and θ'.



(c) Bounds for players who choose $k = 1$ in each state.



(d) Mass of players choosing $k = 1$ must be different.

However, as players cannot observe the true state, we add some assumptions to guarantee more players choose $k = 1$ when the state is θ'. Assume that,

1. when the state is θ, players $i \in I_1$ asymptotically believe $r_2^i \geq r_2^i(\underline{\theta}_2)$ and $r_3^i \geq r_3^i(\underline{\theta}_3)$ and,

2. when the state is θ', players $i \in I_2 \cup I_3$ asymptotically believe $r_1^i \geq r_1^i(\theta_1)$.

Under the first assumption, there is an upper bound for the mass of players who eventually choose $k = 1$ when the state is θ, which is defined by θ_1 and the worst possible

[13] This figure is just for intuition, since we cannot consider players as fixed points in $\mathbb{R}^2$ when we consider more than one state. For instance, player i may be indifferent between $k = 1$ and $k = 2$ while player j is better off playing $k = 2$ in state θ. In state $\bar{\theta} = (\theta_1, \bar{\theta}_2, \theta_3)$ such that $\bar{\theta}_2 > \theta_2$, player i may have stronger preference on $k = 2$ against $k = 1$ than player j. Assumption A3 only implies that both prefer $\bar{\theta}$ instead of θ if they choose $k = 2$.

state for action $k = 2$ and $k = 3$ according to the beliefs of players in I_1. Moreover, this upper bound is a lower bound (defined by the second assumption) for the mass of players who choose $k = 1$ when the state is θ'. To illustrate this bound, consider $\tilde{\theta} = (\theta_1, \underline{\theta}_2, \underline{\theta}_3)$ and see figure 1.3c. We show that this bound is not reachable, otherwise players arbitrarily close to lines 1-2 and 1-3 for the state $\tilde{\theta}$ would have beliefs arbitrarily wrong in the long run (for a formal proof, see Appendix A.3). Next we formalize the result.

Lemma 1.9. *Let (m^*, F^*) be an equilibrium and $\theta, \theta' \in \Theta$ such that $\theta \neq \theta'$. Assume there exists $k \in A$ such that $\theta_k \leq \theta'_k$ and $\theta_{\bar{k}} \geq \theta'_{\bar{k}}$, for each $\bar{k} \in A \setminus \{k\}$. Suppose, for almost all $i \in I$ and each $\bar{k} \in A \setminus \{k\}$,*

1. *$\mu(F^*(i)|\theta, m^*)(B^i_\infty(\bar{k}, \theta'_{\bar{k}})|F^i_k) = 1$ and*
2. *$\mu(F^*(i)|\theta', m^*)(B^i_\infty(k, \theta_k)|F^i_{\bar{k}}) = 1$.*

Then $m^*_\infty(k, \theta_k) < m^*_\infty(k, \theta'_k)$.

While Lemma 1.9 provides sufficient conditions so that θ is distinguishable from θ' such that neither $\theta > \theta'$ nor $\theta < \theta'$, next lemma provides sufficient conditions so that θ is distinguishable from $\bar{\theta} > \theta$. The idea behind Lemma 1.10 is similar to that of Lemma 1.8: the fact that a state θ is distinguishable from states θ' that are not "above" is sufficient to guarantee that θ is also distinguishable from states that are "above", that is, in this case θ would be identified. For a formal proof, see Appendix A.4.

Lemma 1.10. *Let (m^*, F^*) be an equilibrium and let $\theta \in \Theta$. Assume at least one of the conditions below is true for each $\theta' \in \Theta$ such that $\theta' \not> \theta$ and $\theta'_k = \theta_k$, for some $k \in A$:*

1. *$\theta' \in ID$;*
2. *$m^*_\infty(\theta) \neq m^*_\infty(\theta')$.*

Then $\theta \in ID$.

Finally we are ready to prove that **Theorem 1.1** also holds in a environment with three actions. In any equilibrium, almost all players not only eventually choose the best action for themselves, but also asymptotically discover the true state of the world. We prove this result by induction. Given a state θ, we assume all states $\underline{\theta} < \theta$ are identified, thus distinguishable from θ. We show the conditions to use Lemma 1.9 are satisfied, then θ is distinguishable from states θ' such that neither $\theta' \not> \theta$ nor $\theta \not> \theta'$. Finally, we use Lemma 1.10 to show θ is distinguishable from states $\bar{\theta} > \theta$. See Appendix A.5 for a complete proof.

1.7 Conclusion

Considering long-lived fully rational heterogeneous players, this paper fills a gap in the social learning literature. Our main result has important policy implications, as we show that heterogeneity of agents can improve learning in a social environment.

Although we prove the results with a specific model, we argue that they hold for other specifications that makes sense economically. What drive our results are the existence of players around the indifference between actions, in each state of the world, and the fact that players update their beliefs using Bayes' rule. These two factors make the distributions of actions be different across states, allowing players to asymptotically discover the true state through their observations in society. Many other model specifications do not change this fact. For instance, we could add more observations in society per period, include observation of others' outcome, let agents not be anonymous and include a (not perfect) correlation between players' actions or preferences and observed actions.

In particular, correlation between chosen actions and observed actions could help explain polarization especially in a social media environment. Suppose agents may choose between two opinions, L and R, and there is an algorithm that selects whose opinions each agent is observing, in a way that agents are more likely to observe opinions equal to theirs. If they believe in a miss specified model in which their observations are completely independent from their opinions, they may be more confident about the opinions they already had. We leave the development of this idea for further research.

2 Social Learning and Polarization

2.1 Introduction

People have different preferences. For instance, even if they knew how good is each political party, they could still vote differently. In this complete information setting, we would not expect to observe people concentrated in the extremes: highly preferring either a left party or a right party. Evidences of such polarization may suggest that people have different beliefs. We want to know whether this divergence of beliefs may arise because of biased observations of others.

Our model begins with the framework in the first chapter, where there is a continuum of long-lived agents who are heterogeneous in the sense that they have different preferences. In each period, each agent chooses between left and right, and has stochastic payoffs that depend on the quality of his choice. He also observes the choice of another randomly chosen agent, which we call his *observation in society*. However, in this paper we consider agents are more likely to observe others with the same choice they made. One could reason that they are users of a social media that selects whose choice each agent observe. We consider that players do not account for this bias in the observation of others.[1]

We show that society may be asymptotically split in two groups of people, each making a different choice and being sure that the other option is much worse than it really is. This result highlights how social media may confuse learning, inducing polarization. If observations in society were representative of the true proportion of choices or agents fully accounted by the bias in their observations, we show in the first chapter that agents would asymptotically learn the true state of the world. Although they could choose differently because of preferences heterogeneity, they would at least converge in beliefs.

However, since players in this paper do not know their observations in society are biased, they believe they are in an equilibrium of a game where agents eventually choose correctly. Therefore, in the long run, they believe the proportion of agents choosing each action is highly informative about the state of the world. However, if they are more likely to observe others with the same choice, they will underestimate the proportion of agents choosing the other option and, therefore, will believe the other choice is worse than it really is. A difficulty we find to show this result is that, since our players are not fully rational, we cannot use some strong asymptotic results from the literature, such as the

[1] As an empirical evidence, Pogorelskiy e Shum (2019) show that subjects in a laboratory experiment share their signals selectively, but take information from others at face value.

convergence of beliefs.

This paper is closely related to the work in the first chapter. Despite the fact that now we consider a particular kind of players heterogeneity, the main difference is that this paper considers players' observations in society depend on their choice in that period, although players do not know about this. In turn, the model in the first chapter is highly related to those of Aoyagi (1998) and Camargo (2014), mainly the latter. Aoyagi (1998) studies a two-armed bandit model with finite homogeneous players who observe each other's actions in every period. Camargo (2014) considers a multi-armed bandit model with a continuum of homogeneous players that observe, in each period, the action of another randomly chosen player. Both Aoyagi (1998) and Camargo (2014) show that players eventually choose the same option, but not necessarily the best one. The main difference in the model of the first chapter is that players are heterogeneous. In that paper we show that observations of others are more informative in the long run compared with the case where agents are homogeneous. This fact makes players eventually choose the best action for themselves and asymptotically learn the true state of the world.

Apart from the social learning literature, this paper is also related to the polarization literature, in which the concept of echo chambers plays an important role. Baumgaertner (2014) defines echo chamber as a "sociological setting where people's beliefs are 'echoed back' giving the impression that their beliefs are correct". It naturally appears when people gather in groups with similar interests. Especially because of social media, echo chambers appear much more easily nowadays. In the past, it could be hard for an agent with very unusual belief to find other people around him that think similarly; now, he is likely to find an on-line community that reinforces this belief. Furthermore, social media' algorithms also play a role in selecting shared content that this user will enjoy. And even when he is exposed to dissenting viewpoints, he is more likely to choose not to see it.[2]

To develop models to better understand the appearance of echo chambers, economists have been using some concepts from Psychology, such as cognitive dissonance and confirmation bias. The former is the idea that people avoid to update beliefs in a direction that are not agreeable with past actions (FESTINGER; CARLSMITH, 1959) and the latter is the seeking or interpretation of evidences that reinforces what one's already believes (NICKERSON, 1998). For instance, Yariv (2002) builds a model where agents *choose* their beliefs trying to avoid the discovery that a past decision was wrong and Rabin e Schrag (1999) consider agents who have a positive probability to misread signals that are against their priors.

[2] On a study on Facebook, Bakshy, Messing e Adamic (2015) show that users' choices, when compared with algorithmic ranking, play a more important role in the consumption of dissenting viewpoint content.

Using a different approach, Jann e Schottmüller (2018) consider that an agent's payoff depends on others' actions. They find that agents want to share their true information with those who are similar to them, which causes segregation.

We do not intend to explain the formation of echo chambers. Instead, we assume they exist to show how social learning can cause polarization of beliefs. It is not surprising that a model considering echo chambers, like ours, leads to polarization. We believe that our main contribution is to connect the polarization and social learning literatures, providing a solid reason in how (boundedly rational) bayesian agents wrongly update their beliefs, since this would not happen with fully rational agents.

Although there are evidences from Psychology that people suffer from confirmation bias and that agents have preferences that consider beliefs consistency and what others choose to do, factors that may imply in polarization, we study a different channel. We want to know how social learning may also cause polarization if agents are not fully aware of their biased observations of others.[3]

2.2 Model

A continuum of anonymous players is identified with a probability space $(I, \mathcal{I}, \lambda)$. In each period of time $t \in \mathbb{N}$, they must choose one action $k \in \{l, r\}$. Their gross payoffs after choosing action k are stochastic and depend on an unknown parameter $\theta_k \in \Theta$, where $\Theta = \{G, B\}$. The true state of the world is a pair $(\theta_l, \theta_r) \in \Theta^2$, which is drawn once and for all in the beginning of the game from an unknown distribution. Actions are either good (G) or bad (B). For each $k \in \{l, r\}$, gross current payoff is either 1 (success) or 0 (failure). For each period, an agent who chooses action k gets 1 with probability 0.75 if k is good ($\theta_k = G$) and probability 0.25 if k is bad ($\theta_k = B$).

Players are heterogeneous in the sense that they have different costs to choose each action. Let $c(i, k)$ be player i's cost to choose action k. In period t, if he chooses action k and gets gross payoff equals to y_t, his net payoff is $\tilde{y_t}^i = y_t - c(i, k)$. Let $b(i) = c(i, l) - c(i, r)$ be player i's bias towards action r. We assume b is uniformly distributed between -1 and 1, which is common knowledge. This implies that a positive mass of players want to choose each action no matter the state of the world.

Let $\Pi = \Delta(\Theta^2)$ be the set of possible beliefs about the state of the world. Priors are common knowledge and may also be heterogeneous. Let $\phi : I \to \Pi$ be such that $\phi(i) \in \Pi$ is the prior of player $i \in I$. Denote $\pi_t^i(\theta)$ the probability player i assigns to $\theta \in \Theta^2$ in period t and $\pi_t^i(k, \theta_k) := \sum_{\{\theta' : \theta'_k = \theta_k\}} \pi_t^i(\theta')$ the probability he assigns to $\theta_k \in \Theta$. We assume that players do not assign a probability arbitrarily close to 0 to any state in

[3] One could also reason that, in our model, observations of others are not biased but agents have positive probability of misunderstanding what they observe because of confirmation bias.

the beginning of the game.

$$\inf\{\pi_1^i(\theta) : i \in I, \theta \in \Theta^2\} > 0. \tag{A1}$$

In each period t, after choosing an action $k \in \{l, r\}$ and seeing his gross payoff $y \in \{0, 1\}$, each player observes the action $\tilde{k} \in \{l, r\}$ of another randomly chosen player. The set of possible histories until period t is $H_t = (\{l, r\} \times \{0, 1\} \times \{l, r\})^{t-1}$. The set of infinite histories is $H_\infty = (\{l, r\} \times \{0, 1\} \times \{l, r\})^\infty$. A strategy is a sequence $\sigma = \{\sigma_t\}$ such that $\sigma_t : H_t \to \Delta(\{l, r\})$ maps any history in H_t to an action (possibly mixed) in period t. Let Σ be the set of all possible strategies. A strategy profile $F : I \to \Sigma$ is a $\mathcal{I}$-measurable function[4] that maps each player $i \in I$ to a strategy $F(i) \in \Sigma$. The set of all possible strategy profiles is denoted by $\mathcal{F}$.

Although individual actions may be stochastic, aggregated behavior for each period and state is completely determined by players' strategies. We denote by $\tilde{m}_t(k, \theta)$ the proportion of agents choosing k in period t when the state of the world is θ.[5] If a player chooses action k in that period, to simplify our analysis we consider that the probability of observing another agent choosing k is that of a fair draw of agent in the society incremented by $\eta > 0$, provided it is not greater than 1. We define $\tilde{l}_t^k(k, \theta) := \min\{\tilde{m}_t(k, \theta) + \eta, 1\}$ and $\tilde{l}_t^k(\bar{k}, \theta) := 1 - \tilde{l}_t^k(k, \theta)$, $\bar{k} \neq k$, and we call $\tilde{l}^k = \{\tilde{l}_t^k\}$ the *observation likelihood* for players who choose k. Although $\eta > 0$, players naively believe $\eta = 0$ and do not learn about this parameter.

Given a strategy profile, we can also determine the proportion of players choosing action k at period t when the state of the world is θ *if* η was zero. We denote it by $m_t(k, \theta)$. Since players believe their observations of others are representative of the true proportion of actions, $m = \{m_t\}$ is the sequence that matters for belief updating after an observation in society. We call m the *incorrect observation likelihood.* Denote by $\mathcal{M}$ the set of all possible sequences $m = \{m_t\}$ from $\{l, r\} \times \Theta^2$ into $[0, 1]$. We denote $M : \mathcal{F} \to \mathcal{M}$ the map such that $m = M(F)$ gives us the proportion of players choosing each action, if η was zero, for each state of the world and period when the strategy profile is F.

Consider a player $i \in I$. As outcome y_t depend on the state of the world, he is able to update his prior after choosing an action. If he knows the incorrect observation likelihood m, he will be able to (wrongly) update his beliefs after an observation in society. For any history $h_t \in H_t$, player i updates his prior π_1 to $\pi_t^i(h_t | \pi_1, m)$ through Bayes' Rule. Then, given the incorrect observation likelihood m, we denote the individual problem in which player i has prior π_1 by $IP^i(\pi_1, m)$. We consider agents are myopic.[6] Then player i

[4] F must be a $\mathcal{I}$-measurable function so we can aggregate individual actions.

[5] See Appendix A.2 of Camargo (2014) for more details about how to aggregate individual behavior.

[6] For a game where observations in society are not biased, Proposition 2.1 of Rosenberg, Solan e Vieille (2009) shows that even non-myopic players eventually choose myopically. However, we decide to consider myopic players since we cannot directly use this result to our environment with not fully rational agents.

chooses an optimal strategy σ^* for $IP^i(\pi_1, m)$ such that $\sigma^*_t(h_t)$ maximizes $\mathbb{E}_{\pi^i_t(h_t|\pi_1,m)}[\tilde{y}_t]$, for each $h_t \in H_t$ and $t \in \mathbb{N}$.

Therefore, m affects the optimal strategy each $i \in I$ will choose through $IP^i(\phi(i), m)$; in turn, strategy profile F defines $m = M(F)$. This idea is behind the definition of Nash equilibrium for non-atomic games we adapt to our environment.[7]

Definition 2.1. *A biased equilibrium is a pair (m^*, F^*) such that $F^*(i)$ is an optimal strategy for $IP^i(\phi(i), m^*)$, for λ-almost all $i \in I$, and $m^* = M(F^*)$.*

Since observations of others are biased in our model, agents believe they are in the equilibrium of a game without correlation between action taken and action observed, but they are not. This equilibrium notion uses the idea that almost all agents *think* they have best responses to others' strategies. However, since their model of the world is wrong, aggregated actions do not coincide with the one given by m^* from the equilibrium. In this environment, social learning may be misleading. And this characterizes the echo chamber in our model.

2.3 Results

We start observing what would happen under complete information about the true state of the world. Let p_k be the true probability of success after choosing action $k \in \{l, r\}$. Then, for each player i,

$$\begin{cases} b(i) > p_l - p_r \Rightarrow i \text{ chooses action } r \\ b(i) < p_l - p_r \Rightarrow i \text{ chooses action } l \end{cases}.$$

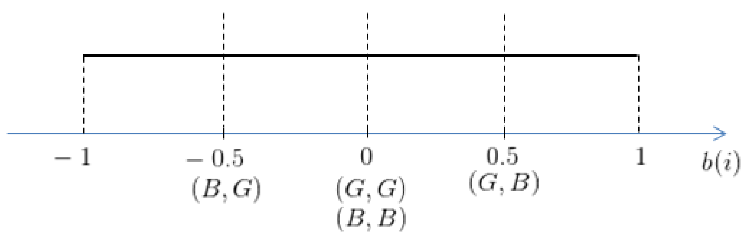


Figure 2.1 – With complete information, we can determine thresholds, for each state of the world, that split players between those who choose action l (players in the left) and those who choose action r (players in the right).

Let $\bar{m}(l, \theta)$ and $\bar{m}(r, \theta)$ be the proportion of agents choosing, respectively, action l and r under complete information when the true state of the world is $\theta \in \Theta$. Let

[7] Equilibria existence is difficult to prove and beyond the scope of this paper. See Appendix B of Camargo (2014) for a proof of existence of equilibria in an environment with homogeneous players whose observations of others are not biased.

$\bar{m}(\theta) = (\bar{m}(l,\theta), \bar{m}(r,\theta))$. Then (see figure 2.1),

$$\begin{aligned} &\bar{m}(B,B) = \bar{m}(G,G) = (0.5, 0.5), \\ &\bar{m}(B,G) = (0.25, 0.75) \quad \text{and} \quad \bar{m}(G,B) = (0.75, 0.25). \end{aligned} \tag{2.1}$$

Now we follow with incomplete information about the state of the world and observations in society. If $\eta = 0$, our main theorem in the first chapter shows that players eventually take the best action for themselves. Stronger than that, players learn the true state of the world through their actions and observations of others. Hence, the distribution of actions converges to that of a game with complete information. Let $m_\infty := \lim_{t\to\infty} m_t$. Lemma 1.4 states this conclusion.

Lemma 2.1. *Let (m^*, F^*) be a biased equilibrium. Then $m^*_\infty = \bar{m}$.*

Therefore, in our setting with $\eta > 0$, players *naively* believe that the true asymptotic distribution of actions is in the set $\{\bar{m}(\theta) : \theta \in \Theta\}$.

For illustration, from now on we consider the true state is $\theta = (G, G)$. Let I_k be the set of players who choose $k \in \{l, r\}$ infinitely often. In an environment with biased observations, it is not clear whether a player $i \in I_k$ asymptotically learns that $\theta_k = G$ with probability 1. By the Strong Law of Large Numbers, his infinitely many outcomes from action k would suffice for him to discover $\theta_k = G$. However, observations in society may be misleading. Lemma 2.2 states that, at least, players are asymptotically sure that the state of the world is not (B, B). By Lemma 2.1, observations in society are not significantly informative to distinguish between states (G, G) and (B, B), since $m_\infty(G, G) = m_\infty(B, B)$. Thus, the information from outcome observations leads them to asymptotically disregard (B, B) when compared to (G, G). Find a formal proof in Appendix B.1.

Lemma 2.2. *Assume $\theta = (G, G)$ and consider a biased equilibrium. Then $\lim_{t\to\infty} \pi^i_t(B, B) = 0$ for almost all players $i \in I$.*

Consider a player i who chooses $k = l$ in period t, his outcome is informative towards $\theta_l = G$ (states (G, G) and (G, B)), but his observation in society may be misleading towards $\theta_l = B$ (states (B, B) and (B, G)). Lemma 2.2 implies that (B, B) is asymptotically disregarded. Lemma 2.3 below states that the log likelihood ratio between player i's belief in states (B, G) and his beliefs in state (G, θ_r), for each $\theta_r \in \{G, B\}$, is more likely to decrease. This implies that the expected total effect in period t is towards learning $\theta_l = G$.

Let $\alpha^i_t(\theta', \theta) := \log\left(\frac{\pi^i_{t+1}(\theta')}{\pi^i_{t+1}(\theta)}\right) - \log\left(\frac{\pi^i_t(\theta')}{\pi^i_t(\theta)}\right)$ be the increase in the log likelihood ratio between player i's beliefs in states θ' and θ. For $\varepsilon > 0$, let $\xi_1(\varepsilon)$ and $\xi_2(\varepsilon)$ be two random

variables such that

$$\xi_1(\varepsilon) = \begin{cases} -2\log 3 + \varepsilon, \text{ with prob. } \bar{p} \\ \varepsilon, \text{ with prob. } 1 - \bar{p} - \underline{p} \\ 2\log 3 + \varepsilon, \text{ with prob. } \underline{p} \end{cases}$$

and

$$\xi_2(\varepsilon) = \begin{cases} -2\log 3 + \log 2 + \varepsilon, \text{ prob. } \bar{p} \\ -\log 2 + \varepsilon, \text{ prob. } 0.75 - \bar{p} \\ \log 2 + \varepsilon, \text{ prob. } 0.25 - \underline{p} \\ 2\log 3 - \log 2 + \varepsilon, \text{ prob. } \underline{p} \end{cases},$$

where $0 \leq \underline{p} := 0.25\max\{0, 0.75 - \eta\} < \bar{p} := 0.75\min\{1, 0.25 + \eta\}$.

Now we can state the next lemma (proof in Appendix B.2).

Lemma 2.3. *Assume $\theta = (G, G)$ and consider a biased equilibrium. Let $\theta_k = \theta'_{\bar{k}} = G$ and $\theta_{\bar{k}} = \theta'_k = B$, for $k, \bar{k} \in \{l, r\}$ such that $\bar{k} \neq k$. Then*

$$\forall \varepsilon > 0, \exists T \in \mathbb{N} \text{ s.t. } \xi_1(\varepsilon) \succsim_{FD} \alpha_t^i(\theta', \theta) \text{ and } \xi_2(\varepsilon) \succsim_{FD} \alpha_t^i(\theta', (G, G))$$

for any player $i \in I$ who chooses k in period $t \geq T$.

Since $\xi^n(\varepsilon)$ has negative expected value for ε small enough, Lemma 2.3 and the Strong Law of Large Numbers imply that players who eventually choose only one action k asymptotically learn that $\theta_k = G$ with probability 1. We state this result below.

Lemma 2.4. *Assume $\theta = (G, G)$ and consider a biased equilibrium. If player $i \in I$ eventually chooses only one $k \in \{l, r\}$, then $\lim_{t\to\infty} \pi_t^i(k, G) = 1$ almost surely.*

But Lemma 2.3 does not imply that a player who chooses, for instance, $k = l$ infinitely many times asymptotically learns $\theta_l = G$. He may also choose $\bar{k} = r$ infinitely many times and, every time he does, may have greater probability to learn that $\theta_l = B$.

However, Lemma 2.5 below states that almost all players do not play both actions infinitely many times. For instance, consider a player i with $b(i) < 0$ who chooses $k = l$ infinitely often. Lemma 2.2 implies he asymptotically disregard (B, B). Using Lemma 2.3, every time he plays l in some period t, there is a strict positive probability that the likelihood ratios $\log\left(\frac{\pi_t^i(B,G)}{\pi_t^i(G,\theta_l)}\right)$ do not ever return to the previous level (they become dominated by a random walk with higher probability of decreasing). Then, we show that there is a strict probability that his belief in the state (B, G) compared to belief in (G, B) does not ever return to the level from period t. Since $b(i) < 0$, he would choose r only if his belief in state (B, G) increased compared to his belief in (G, B). Since every time he

plays l there is a positive probability, bounded away from zero, that he will never play r again, this means that he eventually stops playing r, with probability 1, if he plays l infinitely often. See the proof in Appendix B.3.

Lemma 2.5. *Assume $\theta = (G, G)$ and consider a biased equilibrium. Almost all players eventually choose only one action.*

Because of Lemma 2.5, the proportion of agents choosing each action cannot change forever. Hence, $\{\tilde{m}_t(G, G)\}$ must converge. This is our Lemma 2.6, which comes from Lemma 6 of Camargo (2014) and we state without proof.

Lemma 2.6. *Consider a biased equilibrium. Then $\{\tilde{m}_t(G, G)\}$ is convergent.*

We denote $\tilde{m}_\infty := \lim_{t\to\infty} \tilde{m}_t$ and $\tilde{l}_\infty := \lim_{t\to\infty} \tilde{l}_t$.

Consider players who eventually choose l, they asymptotically learn this action is good and hence the true state of the world is either (G, G) or (G, B). Observations in society may help (or not) these players to disregard (G, B). There exists a threshold of the proportion of *observed* players choosing l below which observations in society for those who eventually choose l, in the long run, are signals towards (G, G) and above which they are signals towards (G, B). We find this threshold is $\frac{\log 2}{\log 3} \approx 0.63$. Thus, players who eventually choose an action k asymptotically believe in the wrong state of the world if they eventually *observe* a proportion greater than 0.63 of agents choosing the same action (see figure 2.2a). Lemma 2.7 states this result (proof in Appendix B.4).

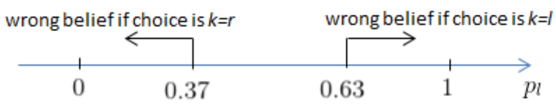
Lemma 2.7. *Assume $\theta = (G, G)$. Consider a biased equilibrium and a player $i \in I_k$ for some $k \in \{l, r\}$. Let $\bar{k} \in \{l, r\}$ such that $\bar{k} \neq k$. If player i eventually observes a proportion of agents choosing k*

- *greater than $\frac{\log 2}{\log 3}$, then $\lim_{t\to\infty} \pi_t^i(\bar{k}, B) = 1$ almost surely;*
- *lower than $\frac{\log 2}{\log 3}$, then $\lim_{t\to\infty} \pi_t^i(\bar{k}, G) = 1$ almost surely.*

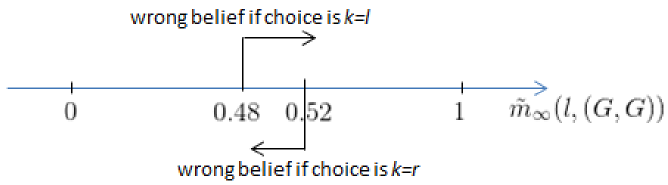
Players who eventually choose action k asymptotically observe a proportion $\tilde{l}_\infty(k, (G, G))$ of players choosing k. If the connection between action taken and action observed is high enough, it will be impossible that almost all players asymptotically learn that the true state of the world is (G, G).

Suppose, for instance, $\eta = 0.15$. If $\tilde{m}_\infty(l, (G, G)) > 0.48$, then players who eventually choose $k = l$ will believe the true state of the world is (G, B). If $\tilde{m}_\infty(l, (G, G)) < 0.52$, then players who eventually choose $k = r$ will believe the true state of the world is (B, G). This means that there is no proportion $\tilde{m}_\infty(l, (G, G))$ such that almost all players learn the true state of the world, while there is an interval (between 0.48 and 0.52) such that

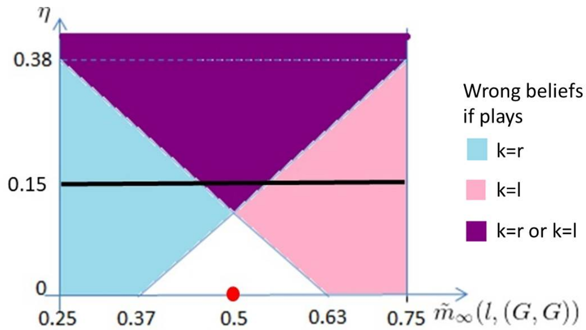
almost all players asymptotically believe in wrong states with probability 1 (see figure 2.2b). Figure 2.2c shows what happens for different values of η.



(a) Observation in society may be misleading depending on the proportion of agents is *observed* choosing action $\tilde{k} = l$.



(b) High correlation between action taken and action observed may prevent that almost all players learn the true state of the world.



(c) Depending on η and on the true proportion of agents asymptotically choosing $k = l$, we can say whether agents who eventually choose l or r have wrong beliefs in the long term, underestimating the other choice. The red dot shows what happens with $\eta = 0$, when half of the agents eventually choose each action. The horizontal line shows the situation when $\eta = 0.15$.

Theorem 2.1 below states these results.

Theorem 2.1. *Assume $\theta = (G,G)$. Consider a biased equilibrium. Let $k, \bar{k} \in \{l, r\}$ such that $\bar{k} \neq k$. Then*

- *if $\tilde{m}_\infty(k,(G,G)) > \frac{\log 2}{\log 3} - \eta$, $\pi^i_\infty(\bar{k}, B) = 1$ for almost all player $i \in I_k$;*
- *if $\eta > \frac{\log 2}{\log 3} - 0.5$, $\pi^i_\infty(\bar{k}, B) = 1$ for almost all $i \in I_k$ or $\pi^j_\infty(k, B) = 1$ for almost all $j \in I_{\bar{k}}$.*

Considering that, with $\eta = 0$, almost all players asymptotically learn the true state of the world, the results in Theorem 2.1 highlights how biased observations of others may disturb learning.

Besides the fact that $\tilde{m}_\infty(l,(G,G))$ must be in the interval $[0.25, 0.75]$, we still do not know which values for $\tilde{m}_\infty(l,(G,G))$ are possible in a biased equilibrium. If we

imagine a symmetric equilibrium,[8] where for each $i \in I$ there is a $j \in I$ with bias to the right $b(j) = -b(i)$, prior $\phi(j) = \phi(i)$ and strategy equals to i's strategy changing actions r by l and vice-versa, half of the agents would eventually choose each action. In this situation, almost all players believe in the wrong state, always underestimating how good the action not taken is.

2.4 Discussion

2.4.1 Political Polarization

We have especial interest in trying to better understand polarization in politics. In this case, the interpretation of some elements of our model may not be so clear.

One could reason that, depending on the agent's preferences and beliefs, he could choose to either follow media that is biased towards the left-wing or media that is biased towards the right-wing. For illustration, assume he chooses left. Then, he gets some news about how good the left party is. It would be expected that the information from the media he chose selects more good information about the left-wing (payoff 1) than the correct proportion between good and bad news. However, the objective of this paper is not about understanding polarization emerged by unknown media bias. Hence, we could imagine that people know about these biases so that they correctly learn the quality of the party they choose.

Moreover, it would be also expected that, even choosing a left biased media, the agent would receive some news about the right-wing. We could imagine that the agent just disregards such news simply because the left media may have incentives to only give bad news about the right party and, thus, those news are not informative (babbling) about the state of the world.

Therefore, choosing left makes the agent learn about the left-wing but not about the right. He needs to make observations in society to learn about the action not chosen. However, although he is aware about the media bias, he is not about how his observation of others are impacted by the decision he made. One could reason, for example, that he does not fully understand how algorithms from social media work, so that he believes that what is presented to him on the Internet is closer to the true proportion of choices than it really is.

2.4.2 Model Generalization

Although we try to make our model the simplest possible to better illustrate our findings, it is important to understand that qualitative results obtain for more general

[8] We do not know whether such equilibrium exists.

settings. In particular, if there were more states of the world, it could be easier for agents to underestimate the effect of the action not taken. For instance, if there were nine possible different probabilities that each θ_k could assume, $p_j = \frac{j}{10}$ for each $j \in \{1, 2, ..., 9\}$, each p_j would be much closer to p_{j+1} than 0.25 is to 0.75 from the model we presented. This means that η that assures not almost all agents will learn the state of the world would be much less than $\frac{\log 2}{\log 3} - 0.5 \approx 0.13$ (around 0.03). Hence, a small correlation between action taken and action observed, not fully accounted by the agents, suffices to make them underestimate the action not taken.

Our results depend on the possible probabilities 0.25 and 0.75 we considered for success (payoff 1) and the uniform distribution of b between -1 and 1, especially Lemma 2.3. However, it is important to highlight that, even with different probabilities and distribution of players, we obtain the same qualitative results conditioned on some η high enough. If correlation between action chosen and action observed is strong, observations in society will help any player to discover his action is good and to (wrongly) believe the other is bad.

2.4.3 Irrationality

In our model, agents are not fully rational since they are not aware that $\eta > 0$. They act as if they were in an unbiased equilibrium. A natural concern is whether it would bring some inconsistency to the agents' behavior. The dynamic in our model helps in this question. Observe that, no matter that observations in society have different probabilities from those the players believe, they never see a history that would be impossible if η was *zero*. They never see themselves as being off-equilibrium. It is true that most of them are on path with relatively low probability of happening according to their probability measures. But, individually, even if η was *zero* they would believe that a positive mass of players would be in paths with relatively low probability. What happens is that most of them think this is their situation.

For instance, suppose that an agent eventually chooses $k = l$ and observes a proportion of agents choosing $\tilde{k} = l$ converging to 0.60. This proportion is not one of the possible proportions in the equilibrium he believes in, so he will always think that his observations in society are not representative of the true distribution of actions, but that sooner or later the proportion observed will converge either to 0.50 or 0.75 (or even 0.25).

However, it is reasonable to believe that, after observing a history with so low probability of happening, an agent should question whether η is zero. To adjust this, we would have to include this parameter in the agents' probability measure, which would make our model much more complicated. Moreover, making players fully rational, learning about η, would certainly avoid the possibility of them being asymptotically sure about the wrong state of the world.

2.5 Conclusion

This paper shows how observations in society may lead to polarization of beliefs if (boundedly rational) bayesian agents do not fully account for correlation between action they take and action they observe. We show that, depending on how high this correlation is, it may be impossible that almost all agents learn the true state of the world. In this situation, some or all agents believe that the decision they did not take is worse than it really is.

Although we can define what happens with learning depending on the asymptotic distribution of actions, we still have to understand what possible values it can have in a biased equilibrium. Also as a next step, we should generalize the probabilities in our model.

Another question we could think about is what would happen if it was presented to agents some statistics with the true proportion of choices of the whole society (an election poll, for instance). They could learn that η is not zero or that the distribution of agents' costs is different from what they thought.

Bibliography

AOYAGI, M. Mutual observability and the convergence of actions in a multi-person two-armed bandit model. *Journal of Economic Theory*, Elsevier, v. 82, n. 2, p. 405–424, 1998.

BAKSHY, E.; MESSING, S.; ADAMIC, L. A. Exposure to ideologically diverse news and opinion on facebook. *Science*, American Association for the Advancement of Science, v. 348, n. 6239, p. 1130–1132, 2015.

BALA, V.; GOYAL, S. Conformism and diversity under social learning. *Economic theory*, Springer, v. 17, n. 1, p. 101–120, 2001.

BANERJEE, A. V. A simple model of herd behavior. *The quarterly journal of economics*, MIT Press, v. 107, n. 3, p. 797–817, 1992.

BAUMGAERTNER, B. Yes, no, maybe so: a veritistic approach to echo chambers using a trichotomous belief model. *Synthese*, Springer, v. 191, n. 11, p. 2549–2569, 2014.

BIKHCHANDANI, S.; HIRSHLEIFER, D.; WELCH, I. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, The University of Chicago Press, v. 100, n. 5, p. 992–1026, 1992.

CAMARGO, B. Learning in society. *Games and Economic Behavior*, Elsevier, v. 87, p. 381–396, 2014.

ELLISON, G.; FUDENBERG, D. Rules of thumb for social learning. *Journal of political Economy*, The University of Chicago Press, v. 101, n. 4, p. 612–643, 1993.

FESTINGER, L.; CARLSMITH, J. M. Cognitive consequences of forced compliance. *The journal of abnormal and social psychology*, American Psychological Association, v. 58, n. 2, p. 203, 1959.

HÖRNER, J.; SKRZYPACZ, A. Learning, experimentation, and information design. *Advances in Economics and Econometrics*, v. 1, p. 63–98, 2017.

JANN, O.; SCHOTTMÜLLER, C. *Why Echo Chambers are Useful.* [S.l.], 2018.

MUNSHI, K. Social learning in a heterogeneous population: technology diffusion in the indian green revolution. *Journal of development Economics*, Elsevier, v. 73, n. 1, p. 185–213, 2004.

NICKERSON, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, SAGE Publications Sage CA: Los Angeles, CA, v. 2, n. 2, p. 175–220, 1998.

POGORELSKIY, K.; SHUM, M. News we like to share: How news sharing on social networks influences voting outcomes. *Available at SSRN 2972231*, 2019.

RABIN, M.; SCHRAG, J. L. First impressions matter: A model of confirmatory bias. *The quarterly journal of economics*, MIT Press, v. 114, n. 1, p. 37–82, 1999.

ROSENBERG, D.; SOLAN, E.; VIEILLE, N. Informational externalities and emergence of consensus. *Games and Economic Behavior*, Elsevier, v. 66, n. 2, p. 979–994, 2009.

ROTHSCHILD, M. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, Elsevier, v. 9, n. 2, p. 185–202, 1974.

SMITH, L.; SØRENSEN, P. Pathological outcomes of observational learning. *Econometrica*, Wiley Online Library, v. 68, n. 2, p. 371–398, 2000.

YARIV, L. I'll see it when i believe it? a simple model of cognitive consistency. Cowles Foundation Discussion Paper, 2002.

YOUNG, H. P. Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American economic review*, v. 99, n. 5, p. 1899–1924, 2009.

Appendix

APPENDIX A – First Chapter Proofs

A.1 Lemma 1.5

Proof. Consider $\theta' \neq \theta$. Fix $i \in I_k$ a player with strategy σ and beliefs $\{\pi_t\}$ that plays k infinitely many times. If $\theta'_k \neq \theta_k$, Lemma 1.1 completes the proof. Suppose, then, that $\theta'_k = \theta_k$. Let $\{\tilde{k}_t\}$ be the sequence of observations in society. By Lemma 1.3, there exists T such that i plays k for all $t \geq T$. By Bayes' Rule, for $t \geq T$,

$$\frac{\pi_{t+1}(\theta')}{\pi_{t+1}(\theta)} = \frac{\pi_t(\theta')}{\pi_t(\theta)} \cdot \frac{m_t(\tilde{k}_t, \theta')}{m_t(\tilde{k}_t, \theta)}. \tag{A.1}$$

The outcome y_t does not change the likelihood ratio $\frac{\pi(\theta')}{\pi(\theta)}$ since $\theta'_k = \theta_k$.

In equation A.1, defining $\gamma_t := \log\left(\frac{\pi_t(\theta')}{\pi_t(\theta)}\right)$ and $\zeta_t := \log\left(\frac{m_t(\tilde{k}_t,\theta')}{m_t(\tilde{k}_t,\theta)}\right)$,

$$\gamma_{t+1} = \gamma_t + \zeta_t. \tag{A.2}$$

Using strict concavity of the log function,

$$\begin{aligned}
\mathbb{E}[\zeta_t|\theta] &= m_t(\tilde{k}, \theta) \log\left(\frac{m_t(\tilde{k}_t, \theta')}{m_t(\tilde{k}_t, \theta)}\right) + (1 - m_t(\tilde{k}, \theta)) \log\left(\frac{1 - m_t(\tilde{k}_t, \theta')}{1 - m_t(\tilde{k}_t, \theta)}\right) \\
&< \log\left(m_t(\tilde{k}, \theta)\frac{m_t(\tilde{k}_t, \theta')}{m_t(\tilde{k}_t, \theta)} + (1 - m_t(\tilde{k}, \theta))\frac{1 - m_t(\tilde{k}_t, \theta')}{1 - m_t(\tilde{k}_t, \theta)}\right) = 0.
\end{aligned} \tag{A.3}$$

Consider $Z_n := \frac{1}{n}\sum_{t=T}^{T+n-1} \zeta_t$. As $n \to \infty$, the variance of $\{Z_n\}$ converges to zero since ζ_t's variance is bounded. Chebyshev's Theorem guarantees that $\{Z_n\}$ converges in probability to its negative expected value. Hence, we get that $\sum_{t=T}^{\infty} \zeta_t \xrightarrow{p} -\infty$. Then $\gamma_t \xrightarrow{p} -\infty$ and, thus, $\{\pi_t(\theta')\}$ converges in probability to 0.

Since $\{\pi_t(\theta')\}$ converges almost surely to a random variable π_∞, we get that $\pi_\infty = 0$ with probability 1.

□

A.2 Lemma 1.8

Proof. Fix $\theta \in \Theta$. Assume,

$$\forall \underline{\theta} \in \Theta,\ \underline{\theta} < \theta \Rightarrow \underline{\theta} \in ID. \tag{A.4}$$

Since players can distinguish θ from any $\underline{\theta} < \theta$, Lemma 1.1 implies that $\theta \in EB$. Lemma 1.7 implies that $A^i_\infty = BR^i(\theta)$ for λ-almost all players.

Now we prove that $\theta \in ID$. Assume, by contradiction, there exists $\bar{\theta} \neq \theta$ such that θ is not distinguishable from $\bar{\theta}$. Lemma 1.6 implies $m_\infty(\bar{\theta}) = m_\infty(\theta)$ and either $\bar{\theta}_1 = \theta_1$ or $\bar{\theta}_2 = \theta_2$. Assume, without loss of generality, $\bar{\theta}_1 = \theta_1$. Equation A.4 implies that $\bar{\theta}_2 > \theta_2$.

When the state is $\bar{\theta}$, using Lemmas 1.1 and 1.2 and the fact that states $(\theta_1, \underline{\theta}_2)$ with $\underline{\theta}_2 < \theta_2$ are either inexistent or identified,

$$i \in I_1 \text{ in } \bar{\theta} \Rightarrow r^i_1(\theta_1) = r^i_1(\bar{\theta}) \geq \mathbb{E}_{\pi^i_\infty}[r^i_2] \geq r^i_2(\theta_2), \text{ for almost all } i \in I. \tag{A.5}$$

When the state is θ, $A^i_\infty = BR^i(\theta)$ for almost all players. Then equation A.5 can be rewritten as

$$i \in I_1 \text{ in } \bar{\theta} \Rightarrow i \in I_1 \text{ in } \theta, \text{ for almost all } i \in I. \tag{A.6}$$

Equation A.6 implies that $m_\infty(1, \bar{\theta}) \leq \lambda\{i \in I : BR^i(\theta) = \{1\}\} = m_\infty(1, \theta)$. Since we assumed $m_\infty(\bar{\theta}) = m_\infty(\theta)$, the converse of equation A.6 must be true.

$$i \in I_1 \text{ in } \theta \Rightarrow i \in I_1 \text{ in } \bar{\theta}, \text{ for almost all } i \in I. \tag{A.7}$$

Consider $I(\varepsilon) := \{i \in I : 0 \leq r^i_1(\theta_1) - r^i_2(\theta_2) < \varepsilon\} \subset \{i \in I : BR^i(\theta) = \{1\}\}$. Note that $\lambda\{I(\varepsilon)\} > 0$ according to assumption A1. When the true state is θ, almost all $i \in I(\varepsilon)$ are such that $i \in I_1$. Equation A.7 implies also that almost all $i \in I(\varepsilon)$ are such that $i \in I_1$ when the true state is $\bar{\theta}$.

Consider $\bar{\theta}$ is the true state. For $i \in I(\varepsilon) \cap I_1$,

$$0 \leq r^i_1(\theta_1) - \mathbb{E}_{\pi^i_\infty}[r^i_2] \leq r^i_1(\theta_1) - [\pi^i_\infty(2, \bar{\theta}_2) r^i_2(\bar{\theta}_2) + (1 - \pi^i_\infty(2, \bar{\theta}_2)) r^i_2(\theta_2)]. \tag{A.8}$$

Where the last inequality considers that $r^i_2(\theta_2)$ is the worst possible expected payoff for action $k = 2$ according to equation A.4. Rewriting equation A.8,

$$0 \leq r^i_1(\theta_1) - r^i_2(\theta_2) - \pi^i_\infty(2, \bar{\theta}_2)(r^i_2(\bar{\theta}_2) - r^i_2(\theta_2)) \leq r^i_1(\theta_1) - r^i_2(\theta_2) < \varepsilon. \tag{A.9}$$

Taking ε arbitrarily small, equation A.9 implies that $\pi^i_\infty(2, \bar{\theta}_2)$ must be arbitrarily small (using assumption A3), for almost all $i \in I(\varepsilon)$. We get a contradiction: beliefs on the true state are submartingales, thus cannot be arbitrarily wrong for almost all $i \in I(\varepsilon)$.

□

A.3 Lemma 1.9

Proof. Assume, without loss of generality, that $k = 1$. Define $P := \{i \in I : r^i_1(\theta_1) \geq r^i_{\bar{k}}(\theta'_{\bar{k}}), \bar{k} \in \{2, 3\}\}$. When the state is θ, assumption 1 implies that $i \in I_1 \Rightarrow i \in P$, for

almost all $i \in I$. When the state is θ', assumption 2 implies that $i \in P \Rightarrow i \in I_1$, for almost all $i \in I$. Then $m_\infty(1,\theta) \leq \lambda\{P\} \leq m_\infty(1,\theta')$. Now we show that at least one inequality must be strict. Assume, by contradiction, $m_\infty(1,\theta) = \lambda\{P\} = m_\infty(1,\theta')$, that is, I_1 when the state is θ, P and I_1 when the state is θ' are the same except by a set of players with measure zero.

Case I: $\theta_1 < \theta_1'$. For $\varepsilon > 0$, define $I(\varepsilon) := \{i \in I : 0 < r_2^i(\theta_2') - r_1^i(\theta_1) < \varepsilon \text{ and } r_1^i(\theta_1) \geq r_3^i(\theta_3')\}$[1]. Assume θ' is the true state. Since $I(\varepsilon) \cap P = \emptyset$, almost all $i \in I(\varepsilon)$ must be such that $i \notin I_1$. For almost all $i \in I(\varepsilon)$, assumption 2 and condition $r_1^i(\theta_1) \geq r_3^i(\theta_3')$ imply that $i \in I_2$. Similarly to the arguments we used in the proof of Lemma 1.8, for ε arbitrarily small, $\pi_\infty^i(1,\theta_1')$ must also be arbitrarily small for almost all $i \in I(\varepsilon)$ when the true state is θ' to guarantee $i \in I_2$. Contradiction since beliefs on the true state are submartingales, thus cannot be arbitrarily wrong for almost all $i \in I(\varepsilon)$.

Case II: $\theta_2 > \theta_2'$. For $\varepsilon > 0$, define $I(\varepsilon) := \{i \in I : 0 \leq r_1^i(\theta_1) - r_2^i(\theta_2') < \varepsilon \text{ and } r_1^i(\theta_1) \geq r_3^i(\theta_3')\}$. Assume the true state is θ. Since $I(\varepsilon) \subset P$, almost all $i \in I(\varepsilon)$ must be such that $i \in I_1$. For ε arbitrarily small, $\pi_\infty^i(2,\theta_2)$ must also be arbitrarily small for almost all $i \in I(\varepsilon)$ when the state is θ to guarantee $i \in I_1$. Contradiction since beliefs on the true state are submartingales.

Case III: $\theta_3 > \theta_3'$. Analogous to Case II.

Then $m_\infty(1,\theta) < m_\infty(1,\theta')$ for all cases.

□

A.4 Lemma 1.10

Proof. It is straightforward to see that θ is distinguishable from any θ' such that $\theta' \not> \theta$. Assume, by contradiction, $\bar\theta > \theta$ is not distinguishable from θ. Then $m_\infty(\bar\theta) = m_\infty(\theta)$ and there exists $k \in A$ such that $\bar\theta_k = \theta_k$ (Lemma 1.6). Assume, without loss of generality, that $\bar\theta_1 = \theta_1$, $\bar\theta_2 > \theta_2$ and $\bar\theta_3 \geq \theta_3$.

Any $\theta' \in \Theta$ such that $\theta' \not> \theta$ and $\theta_1' = \theta_1$, if existent, is such that $\theta' \in ID$ or $m_\infty(\theta') \neq m_\infty(\theta) = m_\infty(\bar\theta)$. Therefore, $\bar\theta$ is also distinguishable from such θ' (Lemma 1.5). Then, for each $\bar k \in \{2,3\}$ and almost all $i \in I$,

$$\mu(F^*(i)|\bar\theta, m^*)(B_\infty^i(\bar k, \theta_{\bar k})|F_1^i) = 1. \tag{A.10}$$

If θ is distinguishable from any θ' such that $\theta' \not> \theta$, then $\theta \in EB$. Hence, for each $\bar k \in \{2,3\}$ and almost all $i \in I$,

$$\mu(F^*(i)|\theta, m^*)(B_\infty^i(1, \theta_1)|F_{\bar k}^i) = 1. \tag{A.11}$$

---

[1] We could invert $\bar k = 2$ and $\bar k = 3$ in the definition of $I(\varepsilon)$ and follow analogous arguments.

With equations A.10 and A.11, Lemma 1.9 implies that $m_\infty(\theta) \neq m_\infty(\bar{\theta})$. Contradiction.

$\square$

A.5 Theorem 1.1 for Three Actions

Proof. We prove by induction. Clearly $(\theta_1^1, \theta_2^1, \theta_3^1) \in ID$ by Lemma 1.10. Fix $\theta \in \Theta$. Assume

$$\begin{aligned} \underline{\theta} \in ID, \text{ for each } \underline{\theta} \text{ such that } \underline{\theta}_1 < \theta_1 \text{ or } (\underline{\theta}_1 = \theta_1 \text{ and } \underline{\theta}_2 < \theta_2) \text{ or } \\ (\underline{\theta}_1 = \theta_1 \text{ and } \underline{\theta}_2 = \theta_2 \text{ and } \underline{\theta}_3 < \theta_3). \end{aligned} \tag{A.12}$$

Then, for almost all $i \in I$,

$$\mu(F^*(i)|\theta, m^*)(B_\infty^i(2, \theta_2)|F_1^i) = 1 \text{ and} \tag{A.13}$$

$$\mu(F^*(i)|\theta, m^*)(B_\infty^i(1, \theta_1)|F_2^i \cup F_3^i) = 1. \tag{A.14}$$

We want to prove that the conditions in Lemma 1.10 are satisfied. We already know that $\theta' \in ID$, for each $\theta' \in \Theta$ such that

- $\theta_1' < \theta_1$ or
- $\theta_2' < \theta_2$ and $\theta_1' = \theta_1$.

Claim A.1 below proves that $m_\infty^*(\theta) \neq m_\infty^*(\theta')$, for each $\theta' \in \Theta$ such that $\theta \not> \theta'$, $\theta_3' < \theta_3$ and $(\theta_1' = \theta_1$ or $\theta_2' = \theta_2)$.

Claim A.2 below shows that $m_\infty^*(\theta) \neq m_\infty^*(\theta')$, for each $\theta' \in \Theta$ such that $\theta \not> \theta'$, $\theta_2' < \theta_2$ and $\theta_3' = \theta_3$.

Finally, Lemma 1.10 implies that $\theta \in ID$.

$\square$

Claim A.1. *Let $\theta' \in \Theta$ such that $\theta \not> \theta'$, $\theta_3' < \theta_3$ and $(\theta_1' = \theta_1$ or $\theta_2' = \theta_2)$. Then $m_\infty^*(\theta) \neq m_\infty^*(\theta')$.*

Proof. We prove by induction. Fix $m \in \{1, 2, ..., n_3\}$ such that $\theta_3^m < \theta_3$. Let $\theta' \in \Theta$ such that $\theta_3' = \theta_3^m$. Assume that

$$m_\infty(\theta) \neq m_\infty(\theta''), \text{ for each } \theta'' \in \Theta \text{ such that } \theta \not> \theta'' \text{ and } \theta_3'' < \theta_3^m. \tag{A.15}$$

Lemma 1.5 implies that, for almost all $i \in I$,

$$\mu(F^*(i)|\theta, m^*)(B_\infty^i(3, \theta_3^m)|F_1^i \cup F_2^i) = 1. \tag{A.16}$$

That is, almost all players $i \in I_2 \cup I_3$ believe r_3^i is at least $r_3^i(\theta_3^m)$ when the state is θ. Next we show that they must also believe that r_3^i is at least $r_3^i(\theta_3^{m+1})$.

Assume, by contradiction, this is not true. Then there exists $\tilde{\theta} \in \Theta$ such that $\tilde{\theta}_3 = \theta_3^m$ and θ is not distinguishable from $\tilde{\theta}$. Lemma 1.5 implies that $m_\infty^*(\tilde{\theta}) = m_\infty^*(\theta)$. Lemma 1.1 implies that either $\tilde{\theta}_1 = \theta_1$ or $\tilde{\theta}_2 = \theta_2$.

<u>Case I</u>: $\tilde{\theta}_1 = \theta_1$. Then $\theta \not> \tilde{\theta}$ implies that $\tilde{\theta}_2 > \theta_2$. We want to show that $m_\infty^*(\tilde{\theta}) \neq m_\infty^*(\theta)$ using Lemma 1.9. We claim that, for almost all $i \in I$,

$$\mu(F^*(i)|\tilde{\theta}, m^*)(B_\infty^i(2, \theta_2)|F_3^i) = 1. \tag{A.17}$$

That is, almost all players $i \in I_3$ believe r_2^i is at least $r_2^i(\theta_2)$ when the state is $\tilde{\theta}$. Assume, by contradiction, this is not true. Then there exists $\theta'' \in \Theta$ such that $\theta_3'' = \tilde{\theta}_3$, $\theta_2'' < \theta_2$ and θ'' is not distinguishable from $\tilde{\theta}$. Lemma 1.5 implies that $m_\infty^*(\theta'') = m_\infty^*(\tilde{\theta}) = m_\infty^*(\theta)$. Equation A.12 implies that $\theta_1'' > \theta_1$ and, for all $i \in I$,

$$\mu(F^*(i)|\theta'', m^*)(B_\infty^i(1, \theta_1)|F_2^i \cup F_3^i) = 1. \tag{A.18}$$

Equations A.13, A.16 and A.18 imply that $m_\infty^*(\theta'') \neq m_\infty^*(\theta)$ by Lemma 1.9. Contradiction. Hence equation A.17 is true for almost all $i \in I$.

Equation A.12 implies that, for almost all $i \in I$,

$$\mu(F^*(i)|\tilde{\theta}, m^*)(B_\infty^i(1, \theta_1)|F_3^i) = 1. \tag{A.19}$$

Equations A.16, A.17 and A.19 imply $m_\infty^*(\tilde{\theta}) \neq m_\infty^*(\theta)$ by Lemma 1.9. Contradiction.

<u>Case II</u>: $\tilde{\theta}_2 = \theta_2$. Then $\theta \not> \tilde{\theta}$ implies that $\tilde{\theta}_1 > \theta_1$. Equation A.12 implies that, for almost all $i \in I$,

$$\mu(F^*(i)|\tilde{\theta}, m^*)(B_\infty^i(1, \theta_1)|F_2^i \cup F_3^i) = 1. \tag{A.20}$$

Equations A.13, A.16 and A.20 imply that $m_\infty^*(\tilde{\theta}) \neq m_\infty^*(\theta)$ by Lemma 1.9. Contradiction.

The induction is complete.

□

Claim A.2. *Let $\theta' \in \Theta$ such that $\theta_1' > \theta_1$, $\theta_2' < \theta_2$ and $\theta_3' = \theta_3$. Then $m_\infty^*(\theta) \neq m_\infty^*(\theta')$.*

Proof. Assume, by contradiction, there exists $\tilde{\theta} = (\bar{\theta}_1, \underline{\theta}_2, \theta_3)$ with $\bar{\theta}_1 > \theta_1$ and $\underline{\theta}_2 < \theta_2$ such that $m_\infty^*(\theta) = m_\infty^*(\tilde{\theta})$. Equation A.12 implies that, for all $i \in I$,

$$\mu(F^*(i)|\tilde{\theta}, m^*)(B_\infty^i(1, \theta_1)|F_2^i \cup F_3^i) = 1. \tag{A.21}$$

Claim A.1 and Lemma 1.5 imply that, for all $i \in I$,

$$\mu(F^*(i)|\theta, m^*)(B^i_\infty(3, \theta_3)|F^i_1) = 1. \tag{A.22}$$

Equations A.13, A.21 and A.22 imply that $m^*_\infty(\tilde{\theta}) \neq m^*_\infty(\theta)$ by Lemma 1.9. Contradiction. $\square$

APPENDIX B – Second Chapter Proofs

B.1 Lemma 2.2

Proof. Fix a player $i \in I$. Let $\{k_t\}$ be his sequence of actions, $\{y_t\}$ be the sequence of his gross payoffs, $\{\tilde{k}_t\}$ be the sequence of his observations in society.

By Bayes' Rule, for any $\theta, \theta' \in \Theta^2$ and $t \in \mathbb{N}$,

$$\frac{\pi^i_{t+1}(\theta')}{\pi^i_{t+1}(\theta)} = \frac{\pi^i_t(\theta')}{\pi^i_t(\theta)} \cdot \frac{\mathrm{P}(y_t|\theta')}{\mathrm{P}(y_t|\theta)} \cdot \frac{m_t(\tilde{k}_t, \theta')}{m_t(\tilde{k}_t, \theta)}.$$

Taking the logarithm function in both sides, we get the law of motion for the likelihood ratio between states θ' and θ according to player i's beliefs,

$$\gamma_{t+1}(\theta', \theta) = \gamma_t(\theta', \theta) + \lambda_t(\theta', \theta) + \zeta^{\tilde{k}}_t(\theta', \theta), \tag{B.1}$$

where

$$\gamma_t(\theta', \theta) := \log\left(\frac{\pi^i_t(\theta')}{\pi^i_t(\theta)}\right), \tag{B.2}$$

$$\lambda_t(\theta', \theta) := \log\left(\frac{\mathrm{P}(y_t|\theta')}{\mathrm{P}(y_t|\theta)}\right) \text{ and} \tag{B.3}$$

$$\zeta^{\tilde{k}_t}_t(\theta', \theta) := \log\left(\frac{m_t(\tilde{k}_t, \theta')}{m_t(\tilde{k}_t, \theta)}\right). \tag{B.4}$$

Then,

$$\gamma_t(\theta', \theta) = \gamma_1(\theta', \theta) + \sum_{s=1}^{t-1} \alpha_s(\theta', \theta), \tag{B.5}$$

where, for each $s \in \mathbb{N}$,

$$\alpha_s(\theta', \theta) = \lambda_s(\theta', \theta) + \zeta^{\tilde{k}}_s(\theta', \theta).$$

Since $\{m_t\}$ converges to $\bar{m}$ (Lemma 2.1), for each $\tilde{k} \in \{l, r\}$,

$$\lim_{t\to\infty} \zeta^{\tilde{k}}_t(\theta', \theta) = \log\left(\frac{\bar{m}(\tilde{k}, \theta')}{\bar{m}(\tilde{k}, \theta)}\right) =: \zeta^{\tilde{k}}(\theta', \theta). \tag{B.6}$$

Let $\theta' = (B, B)$ and $\theta = (G, G)$. Equations 2.1 and B.6 imply that $\zeta^{\tilde{k}}((BB), (GG)) = 0$,[1] for each $\tilde{k} \in \{l, r\}$.

[1] To simplify notation, we sometimes use $(\theta_l \theta_r)$ instead of (θ_l, θ_r).

Equation B.3 becomes

$$\lambda_t((BB),(GG)) = \begin{cases} \log 3, \text{ if } y_t = 0 \\ -\log 3, \text{ if } y_t = 1 \end{cases}.$$

When the true state of the world is (G,G), we get that $\mathbb{E}[\lambda_t((BB),(GG))] = -0.5\log 3$. Hence, for t large enough, $\mathbb{E}[\alpha_t((BB),(GG))] = \mathbb{E}[\lambda_t((BB),(GG))+\zeta_t^{\tilde{k}}((BB),(GG))]$ is negative and bounded away from *zero*. Then, by the Strong Law of Large Numbers, equation B.5 implies that $\lim_{t\to\infty} \gamma_t((BB),(GG)) = -\infty$ almost surely, which implies that $\lim_{t\to\infty} \pi_t^i(B,B) = 0$ with probability 1.

□

B.2 Lemma 2.3

Proof. We use the same definitions as in the proof in Appendix B.1. Without loss of generality, consider a player $i \in I$ who chooses $k_t = l$ in some period $t \in \mathbb{N}$. Considering $\theta = (G,B)$ and $\theta' = (B,G)$, equation B.3 yields

$$\lambda_t((BG),(GB)) = \begin{cases} \log 3, \text{ if } y_t = 0 \\ -\log 3, \text{ if } y_t = 1 \end{cases}. \tag{B.7}$$

When the true state of the world is (G,G),

$$\mathrm{P}(y_t = 0) = 0.25 \text{ and } \mathrm{P}(y_t = 1) = 0.75. \tag{B.8}$$

Define $a_t := \log\left(\frac{m_t(r,(B,G))}{m_t(r,(G,B))}\right)$, for each $t \in \mathbb{N}$. Equation B.4 implies that

$$\zeta_t^{\tilde{k}}((BG),(GB)) = \begin{cases} -a_t, \text{ if } \tilde{k}_t = l \\ a_t, \text{ if } \tilde{k}_t = r \end{cases}. \tag{B.9}$$

When the true state of the world is (G,G),

$$\mathrm{P}(\tilde{k}_t = l) = \tilde{l}_t(l,(G,G)) \text{ and } \mathrm{P}(\tilde{k}_t = r) = \tilde{l}_t(r,(G,G)). \tag{B.10}$$

From equations B.7-B.9, the distribution of $\alpha_t((BG),(GB)) = \lambda_t((BG),(GB)) + \zeta_t^{\tilde{k}}((BG),(GB))$ is given by

$$\mathrm{P}(\alpha_t((BG),(GB)) = x) = \begin{cases} 0.75\tilde{l}_t(l,(G,G)), \text{ if } x = -\log 3 - a_t \\ 0.75(1-\tilde{l}_t(l,(G,G))), \text{ if } x = -\log 3 + a_t \\ 0.25\tilde{l}_t(l,(G,G)), \text{ if } x = \log 3 - a_t \\ 0.25(1-\tilde{l}_t(l,(G,G))), \text{ if } x = \log 3 + a_t \end{cases}. \tag{B.11}$$

$\{m_t\}$ converges to $\bar{m}$ (Lemma 2.1), then $\lim_{t\to\infty} a_t = a$, where $a := \log\left(\frac{\bar{m}(r,(B,G))}{\bar{m}(r,(G,B))}\right) = \log 3$. Fix $\varepsilon > 0$ and find $T \in \mathbb{N}$ such that $|a_t - \log 3| < \varepsilon$, for all $t \geq T$.

Moreover, the proportion of agents choosing an action l in period t, $\tilde{m}_t(l,(G,G))$, is also between 0.25 and 0.75. Then, using

$$\begin{aligned}
&0.75\tilde{l}_t(l,(G,G)) \geq 0.75\min\{1, 0.25+\eta\} =: \bar{p},\\
&0.25(1-\tilde{l}_t(l,(G,G))) \leq 0.25\max\{0, 0.75-\eta\} := \underline{p} \quad \text{and}\\
&\log 3 - \varepsilon < a_t \leq \log 3 + \varepsilon,
\end{aligned}$$

for $t \geq T$, equation B.11 implies $\xi_1(\varepsilon) \succsim_{FD} \alpha_t^i((BG),(GB))$. An analogous argument shows that $\xi_2(\varepsilon) \succsim_{FD} \alpha_t^i((BG),(GG))$.

□

B.3 Lemma 2.5

Proof. We use the same definitions as in the proof in Appendix B.1. Consider a player $i \in I$ such that $b(i) < 0$. He chooses his actions according to his beliefs, then

$$k_t = r \Rightarrow \pi_t^i(B,G) \geq \pi_t^i(G,B) - 0.5b(i) \quad \text{and} \tag{B.12}$$

$$k_t = l \Rightarrow \pi_t^i(B,G) \leq \pi_t^i(G,B) - 0.5b(i). \tag{B.13}$$

While player i chooses only $k = l$, Lemma 2.3 implies that, for t large enough, $\{\gamma_t((BG),(GG))\}$ and $\{\gamma_t((BG),(GB))\}$ are processes dominated by random walks that have strict positive probability of not ever returning to the previous level above, since $\bar{p} > \underline{p}$. Hence, every time player i chooses l, there is a strict positive probability such that both $\gamma_t((BG),(GG))$ and $\gamma_t((BG),(GB))$ do not ever return to the previous level above. Let's assume this probability is greater than some $\rho \in (0,1)$.

Assume player i chooses $k_t = l$ in period t and $\gamma_s((BG),(GG)) < \gamma_t((BG),(GG))$ and $\gamma_s((BG),(GB)) < \gamma_t((BG),(GB))$, for each $s \geq t$. Condition B.13 does not imply whether $\pi_t^i(B,G) \leq \pi_t^i(G,B)$ or $\pi_t^i(B,G) > \pi_t^i(G,B)$. Next we show that, in both cases, condition B.12 is not satisfied for $s \geq t$.

<u>Case 1</u>: Assume $\pi_t^i(B,G) \leq \pi_t^i(G,B)$, then $\gamma_s((BG),(GB)) < \gamma_t((BG),(GB)) \leq 0$ and condition B.12 is never satisfied. Player i does not choose $k_s = r$ for any $s \geq t$.

<u>Case 2</u>: Assume $\pi_t^i(B,G) > \pi_t^i(G,B)$. For t large enough, Lemma 2.2 states that $\pi_t^i(B,B)$ is arbitrary close to 0. For each $s \geq t$, $\gamma_s((BG),(GG)) < \gamma_t((BG),(GG))$ and $\gamma_s((BG),(GB)) < \gamma_t((BG),(GB))$ imply that $\frac{\pi_s^i(B,G)}{\pi_s^i(G,G)} < \frac{\pi_t^i(B,G)}{\pi_t^i(G,G)}$ and $\frac{\pi_s^i(B,G)}{\pi_s^i(G,B)} < \frac{\pi_t^i(B,G)}{\pi_t^i(G,B)}$. Thus, $\pi_s^i(B,G) < \pi_t^i(B,G)$ for each $s \geq t$. Belief in state (B,G) decreased, but we still need to be sure that belief in state (G,B) did not decrease more, even though the likelihood ratio also decreased. Since $\pi_t^i(B,G) > \pi_t^i(G,B)$, $\frac{\pi_s^i(B,G)}{\pi_s^i(G,B)} < \frac{\pi_t^i(B,G)}{\pi_t^i(G,B)}$ implies that

$\pi_t^s(B,G)-\pi_t^i(B,G) < \pi_t^s(G,B)-\pi_t^i(G,B)$.[2] With condition B.13, we get that $\pi_s^i(B,G) < \pi_s^i(G,B) - 0.5b(i)$ for each $s \geq t$. Condition B.12 is never satisfied and player i does not choose $k_s = r$ for any $s \geq t$.

We conclude that there is a strict positive probability that equation B.12 is never satisfied whenever player i chooses l.

For each $t \in \mathbb{N}$, let E_t be the event where he chooses $k_{t-1} = r$ and $k_t = l$. Then $\sum_{t=1}^{\infty} \mathrm{P}(E_t) \leq \sum_{n=1}^{\infty}(1-\rho)^{n-1} < \infty$, which implies that $\mathrm{P}(E_t\ i.o.) = 0$ by the First Borel-Cantelli Lemma. Thus we conclude that player i must eventually choose only one action with probability 1.

The proof for a player $j \in I$ such that $b(j) > 0$ is analogous.

$\square$

B.4 Lemma 2.7

Proof. We use the same definitions as in the proof in Appendix B.1. Fix a player i that eventually plays $k = l$ (proof for $k = r$ is analogous) and, therefore, know the true state of the world is either (G,B) or (G,G) (Lemma 2.4). Since i eventually plays l, there exists T such that $k_t = l$ for all $t \geq T$, with probability 1. After T, player i updates these states likelihood ratio only because of observations in society.

Using $\theta = (G,G)$ and $\theta' = (G,B)$, equation B.5 becomes, for each $t \geq T$,

$$\gamma_t((GB),(GG)) = \gamma_T((GB),(GG)) + \sum_{\tau=T}^{t} \zeta_t(\tilde{k}_t((GB),(GG))) \tag{B.14}$$

and, from equation B.6,

$$\zeta^{\tilde{k}}((GB),(GG)) = \begin{cases} \log\left(\frac{3}{2}\right), & \text{if } \tilde{k} = l \\ -\log 2, & \text{if } \tilde{k} = r \end{cases}.$$

Fix $\varepsilon > 0$ and find $T' \in \mathbb{N}$ such that $|\zeta_t(\tilde{k})((GB),(GG)) - \zeta(\tilde{k})((GB),(GG))| < \varepsilon$, for all $t \geq T'$ and $\tilde{k} \in \{l,r\}$.

Consider the sequence $\{p_t\}_{t \geq T}$ such that $p_t = \frac{\sum_{s=T}^{t} \mathbb{1}_{\{l\}}(\tilde{k}_t)}{t-T+1}$, which gives us the proportion of *observed* players by player i choosing action l. Assume $\{p_t\}$ is eventually below some $\bar{p} \in [0,1]$. Then, there exists $T'' > T$ such that $p_t \leq \bar{p}$, for each $t \geq T''$.

[2] Intuitively, $\pi^i(G,B)$ percentual decrease is limited by $\pi^i(B,G)$ percentual decrease. Since $\pi_t^i(B,G) > \pi_t^i(G,B)$, $\pi^i(B,G)$ absolute decrease is higher.

Then, for $\bar{T} := \max\{T', T''\}$, using the Strong Law of Large Numbers and equation B.14,

$$\begin{aligned}
\lim_{t\to\infty} \gamma_t((GB),(GG)) &= \gamma_{\bar{T}}((GB),(GG)) + \sum_{\tau=\bar{T}}^{\infty} \zeta_\tau^{\tilde{k}_t}((GB),(GG)) \\
&\leq \gamma_{\bar{T}} + \sum_{t=\bar{T}}^{\infty} \left\{ \mathbb{1}_{\{l\}}(\tilde{k}_t) \left[\log\left(\frac{3}{2}\right) + \varepsilon \right] + \mathbb{1}_{\{r\}}(\tilde{k}_t) \left[-\log 2 + \varepsilon \right] \right\} \\
&\leq \gamma_{\bar{T}} + \sum_{t=\bar{T}}^{\infty} \left\{ \bar{p} \left[\log\left(\frac{2}{3}\right) + \varepsilon \right] + (1-\bar{p}) \left[-\log 2 + \varepsilon \right] \right\}
\end{aligned}$$

with probability 1.

Taking $\varepsilon \to 0$, we get almost surely that

$$\lim_{t\to\infty} \gamma_t \leq \gamma_{\bar{T}} + \sum_{t=\bar{T}}^{\infty} \left[\bar{p} \log\left(\frac{3}{2}\right) - (1-\bar{p}) \log 2 \right]. \tag{B.15}$$

Observe that $\bar{p} \log\left(\frac{3}{2}\right) - (1-\bar{p}) \log 2 < 0 \Leftrightarrow \bar{p} < \frac{\log 2}{\log 3}$. Equation B.15 implies that, if $\bar{p} < \frac{\log 2}{\log 3}$, $\lim\limits_{t\to\infty} \gamma_t = -\infty$ and player i asymptotically believes that the true state of the world is (G,G) with probability 1.

If the proportion of *observed* players choosing $\tilde{k} = l$ is greater than $\underline{p}$, an analogous argument shows that $\lim\limits_{t\to\infty} \gamma_t = \infty$ if $\underline{p} > \frac{\log 2}{\log 3}$ and player i asymptotically believes that the true state of the world is (G,B) with probability 1.

$\square$