FACULTY OF LAW, ECONOMICS AND FINANCE

CREA Discussion Paper 2015-18

Economics

Center for Research in Economics and Management
University of Luxembourg

Endogeneous Quantal Response Equilibrium for Normal Form Games

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers

Oswaldo Gressani, CREA, Université du Luxembourg
CORE, Université catholique de Louvain

November, 2015

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
162A, avenue de la Faïencerie
L-1511 Luxembourg

UNIVERSITÉ DU LUXEMBOURG

Endogeneous Quantal Response Equilibrium for Normal Form Games[1]

Gressani Oswaldo[2]

November 2015

Abstract

We develop an equilibrium concept coined Endogeneous Quantal Response Equilibrium (EQRE) based on heterogeneous players and endogeneous learning in a logistic quantal choice model. Each player has an asymptotically consistent estimate of his rival's rationality index and is able to choose his own rationality level according to a cost-benefit tradeoff. This approach allows to enrich bounded rationality models by incorporating heterogeneous skills and by bridging the gap between stylized facts on the rationality index dynamics and a learning dimension.

Keywords: Quantal Response Equilibrium, bounded rationality, learning.

[1]The present project is supported by the National Research Fund, Luxembourg. The author would like to thank Prof. Pierre Picard and Prof. Vincent Vannetelbosch for their precious comments.

[2]CREA, Université du Luxembourg; CORE, Université catholique de Louvain.
E-mail: oswaldo.gressani@uni.lu, oswaldo.gressani@uclouvain.be

1 Introduction

1.1 From bounded to perfect rationality and back

Decision making is an important aspect of our daily lives. In some cases the decisional process is rooted in the unconscious or genetically preprogrammed, entailing an automatized selection among a set of available alternatives; in most circumstances however, a sophisticated reasoning and a certain level of computational capacities are required. Although the reasoning process is innate in the human mind since the dawn of mankind, the genesis of rationalism can be traced back to the work of ancient Greek philosophers such as Aristotle's Prior Analytics on deductive reasoning or the dialogues of Plato and in a less distant past René Descartes and his Discourse on the Method. In economic theory, the term "rationality" has not the same meaning across time, such that if you question Adam Smith and John Hicks about rational behavior, you will almost surely get different interpretations; the selfish wealth seeker for the classical era and the utility maximizer for the neoclassical one.

Indeed, the idea of *homo economicus* spreads back to the work of John Stuart Mill's essay *"On the Definition of Political Economy; and on the Method of Investigation Proper to It"* [1] , in which the (rational) economic man is defined as a being in the quest of wealth accumulation and leisure, endowed with the capability of "judging the comparative efficacy of means for obtaining that end". Thus, as highlighted by Persky [2], it would be a mistake to confound Mill's "rational man" abstraction with the neoclassical conception of rationality whose core lies in the optimizing behavior. Likewise Adam Smith's, *The Wealth of Nations* [3] mentions individuals using their human capital in labor to acquire *reasonable* compensations without mentioning any utility maximization scheme.

Towards the end of the 19th century arises a paradigm shift from economic agents targeting reasonable compensations from their labor as inspired by Adam Smith and his contemporaries to a utility maximizing rational agent, the latter being the building block of methodological individualism. The neoclassical vision of a perfectly rational agent becomes predominant during the first half of the 20th century if not to say the unique rule in economic decision making. Paraphrasing Herrnstein [4], the rationality assumption in economic theory can be seen equivalent to the Newtonian theory of matter in the physical sciences, i.e. the law that behavior would follow if not disturbed by idiosyncratic forces or cognitive limitation, the behavioral analogues of friction or measurement error. During this period, the distance between what Simon [5] labels "substantive" and "procedural" rationality becomes abyssal. A behavior is said to be substantively rational when it aims at achieving well defined goals under a set of given constraints; hence rationality is in this definition tightly linked to the objectives to be attained and remains silent about the processes of decision to achieve them. On the other side, "behavior is procedurally rational when it is the outcome of appropriate deliberation", encompassing the reasoning and cognitive processes of humans.

In the fifties, the papers of Arrow [6] and Debreu [7] brought a successful axiomatic reformulation of the Walrasian general equilibrium theory based on the perfect rationality assumption. At the same time Savage [8], proposes an elegant Bayesian approach in which behavioral rationality is identified with subjective expected utility maximization. Another breakthrough in the theory of decision making is the pioneering work of John von Neumann and Oskar Morgenstern [9] in the *Theory of Games and Economic Behavior* that presents a new analytical apparatus to cope with the interaction of rational agents in strategic situations. A couple of years later, John Forbes Nash [10] developed the solution concept of equilibrium points for non-cooperative games which is well known as Nash equilibrium and is extensively used in game-theoretic analysis. Ironically, when John Nash walked into the office of the Chair of the Princeton Mathematics Department and presented his theory to von Neumann, the latter scornfully remarked: "that's trivial, you know, that's just a fixed point theorem"[3] without being fully aware of the power of the theory and its future applications to economics, political sciences, psychology and biology.

While the substantive rationality facet reached its climax in economics; in other social sciences and especially in psychology, the arrival of improved experimental methods and new tools for statistical inference opened the gates to procedural rationality, i.e. the study of cognitive processes. The critical punch of Allais [11] towards the von Neumann-Morgenstern expected utility theory, namely that preferences of individuals are insconsistent with the axioms of rational choice, offered without doubt an enthusiastic promotion to the study of psychological factors to enrich the theory of decision making. In the late 1950s, while economic theory was molded in a mathematized shape of perfect rationality, cognitive psychology started to explore the mechanisms of thoughts by means of experiments and new computer programs. According to Simon [12], it is a different conceptualization of rationality where the study of computational strategies in the reasoning process is not less important than the understanding of how emotions and other feelings interact to drive behavior.

The ambition of Simon [13] was to find the center of gravity between the mythical fully rational man and an over-simplification of the latter, in other words the aim was to marry substantive and procedural rationality to give birth to a theory of bounded rationality. Important dimensions of bounded

[3] Nasar, S. (1998) A Beautiful Mind (Simon and Schuster, New York, p.94)

rationality are the limited computational capacities of humans as well as the restricted perception of their environment. To formalize this dimension, Simon [13, 14] proposes a decision scheme based on "aspiration levels" rather than classic maximization illustrated as follows. Let A be a point set of (objectively) available alternatives. Subjectively, the human mind may be limited to perceive a restricted set $\mathring{A} \subset A$. The set of all possible outcomes is denoted by S and we denote its power set by $\mathcal{P}(S) = \{S_a | S_a \subseteq S\}$, such that under incomplete information we have a mapping $\mathring{f} : A \to \{S_a \in \mathcal{P}(S) | S_a \neq \emptyset\}$, meaning that a given behavioral alternative $a \in \mathring{A}$ will lead to an outcome in the subset S_a.

Thus, limited information is here synonym of uncertainty among states of the world in S_a. In order to simplify the decisional process, we can assume a binary nature of the payoff mapping, namely an indicator function $\mathbb{I}_S(s)$ taking the value 1 if the outcome $s \in S$ is satisfactory and 0 otherwise. For instance in chess, +1 would be assigned to a move which contributes to the development of a winning strategy, while the 0 payoff can be associated to a non-contributing move towards checkmate, or an erroneous move giving advantage to the opponent. Finally, the bounded rational decision process consists in identifying a subset $S' \subset S$ of satisfactory outcomes and then search for a behavior alternative in $\mathring{A}$ that maps to outcomes being all elements of S'. This procedure leads to a satisfactory outcome with certainty (if it exists) and highlights Simon's important contrast between satisficing and maximizing strategies.

The introduction of psychological limits in economic decision making is less associated to a behavioral revolution in the field than to a refined comeback to the older classicist conception of limited rationality. In addition, the arrival of laboratory experiments seriously shattered the flawed paradigm of "rational man" and corroborated the bounded rationality theory pioneered by Simon. Of particular importance is the experimental work of Tversky and Kahneman [15, 16] that demonstrates how heuristic principles used by individuals are responsible for the biases in judging probabilities. Another challenge to the perfect rationality assumption is the so called framing effect as highlighted by Tversky and Kahneman [17] where two logically equivalent alternatives are not necessarily treated equally depending on the context of choice, hence violating a core principle of rational choice called invariance. The development of behavioral economics along with the accumulated empirical evidence for deviations from the perfect rationality paradigm makes the sharp divide between psychology and economics belong to the past and gives a chance to a synergistic collaboration between the substantive and procedural rationality dimensions, i.e. a connection between the rational and the psychological. The next section discusses this connection in game-theoretic settings.

1.2 Bounded rationality in Game Theory

Game theory is a set of mathematical tools for analyzing and predicting outcomes of multiperson strategic interactions. In contrast to standard decision theory where a single decision maker is involved, a game-theoretic setting incorporates a feature of strategic interdependence, such that each player's welfare depends not only on his own initiative, but also on those taken by other participants. Nash equilibrium (NE) is undoubtedly the dominant solution concept in non-cooperative games and is based on the following rationality premises: (1) *strategic thinking*, i.e., players form beliefs about the behavior of their rivals, (2) *optimization*, that is, given their beliefs, players choose the actions that maximize their expected payoff and (3) *mutual consistency* which requires that the belief of each player is consistent with the actual behavior of other players.

The NE solution concept is especially attractive for its generality, in the sense that in a game where each player has a finite pure strategy set and is allowed to play mixed strategies, the existence of a NE is guaranteed. However, equilibrium analysis is often criticized for its unrealistic assumptions on players' high degree of rationality. Driven by the accumulated empirical evidence on the poor predictive power of NE in many settings (see Camerer [18] for a complete experimental work on multiperson interactions), the research community has proposed various theories of bounded rationality that have shown to make more accurate predictions than NE in various experimental games already, at the cost of introducing an additional free parameter to which is attributed psychological and cognitive interpretations.

Instead of abandoning equilibrium analysis, the idea is to enrich it by incorporating limits on rationality as in the tradition of Simon. The central point is to capture "real life" behaviors by considering boundedly rational agents that according to Williamson [19] are limited in their capacity in "formulating and solving complex problems and in processing (receiving, storing, retrieving, transmitting) information". A first game-theoretic attempt to catalyze human behavior via limited rationality and cognitive biases can be found in a paper of Rosenthal [20] who presents an alternative behavioral theory for non-cooperative games, namely that a player's probability of adopting a strategy is a monotone-nondecreasing function of his expected payoff, such that players are not exclusively restricted to the choice of best responses. In the same direction, Beja [21] proposes the concept of "imperfect performance equilibrium" with the idea that players fail to implement their target strategy because of inaccurate computations or random errors. Implications of error-prone behavior on equilibrium is also studied by Chen [22], who develops a rational novice model that allows for non-optimal actions.

The paper of McKelvey and Palfrey [23] gives a statistical facelift to bounded rationality models by weakening the optimization assumption usually made in equilibrium analysis and applying a new solution concept coined "Quantal Response Equilibrium" (QRE) to normal form games. More specifically, they use a parametric logit response function to model individual choice behavior and compare QRE with other equilibrium concepts in two-person experimental games. This approach has been complemented by Rogers, Palfrey and Camerer [24] by introducing "skill" heterogeneity among players, such that the individual payoff responsiveness parameter (or rationality index) varies across players. This leads to a generalization of the QRE concept called Heterogeneous Quantal Response Equilibrium (HQRE) in which the distribution of skill-types is common knowledge across players. Moreover, this paper makes a useful connection between QRE and Cognitive Hierarchy (CH) theory pioneered by Camerer, Ho and Chong [25]. CH theory keeps the assumption of best responding to beliefs, but relaxes the assumption of mutual consistency in the sense that players mistakenly anticipate the behavior of their rivals in the game due to a limited spectrum of strategic thinking. Choi, Gale and Kariv [26] apply QRE to the study of social learning in networks. More precisely, the authors adapt the decision rule of QRE models and undertake an empirical study to determine the ability of agents to learn and rationally process the information available in the network in which they belong.

It is important to emphasize that both QRE and CH class of models are one-parameter empirical alternatives to NE that have proved to be successful in predicting the observed systematic deviations from NE in a variety of games. However, a limited feature of these models is that the sophistication level of players is exogenous and identical across all players in the game. Even though Rogers, Palfrey and Camerer [24] allow for players with heterogeneous types, nothing is said about the formation of this heterogeneity. In light of experiments done by Lieberman [27], O'Neill [28], Rapoport and Boebel [29] and Ochs [30] that reveal a non-monotonic increase in the rationality index when games are repeated, the aim of this paper is to develop a theory to endogenize the learning of players.

2 Modelling bounded rationality with statistical reaction functions

2.1 Motivation

A classical approach to predict human behavior in game theory is to utilize the Nash equilibrium solution concept which relies on strict rationality assumptions. Even if this approach is intuitively appealing and applicable to general settings, there is a plethora of evidence that it cannot rationalize behavior in a variety of experimental games. To circumvent this predictive mismatch McKelvey and Palfrey [23] proposed a statistical theory of equilibrium based on quantal choice. The boundedly rational dimension of this QRE model lies in the relaxation of the optimizing assumption, that is, players do not necessarily succeed in choosing their payoff maximizing alternative, but are more likely to select better actions (i.e. those yielding higher expected payoffs) than worse ones. In short, the framework keeps the mutual consistency assumption of rational expectations, but accommodates for noisiness in optimizing behavior.

According to the work of Luce [31] on individual choice behavior, a common specification of the quantal response structure is the logistic function parametrized by a payoff sensitivity parameter $\lambda \in \mathbb{R}_{++}$ to which is attributed psychological insights. When $\lambda \to 0$ players distribute the probability mass uniformly over their pure strategy set with a total lack of sensitivity towards expected gains. At the other extreme, when $\lambda \to \infty$ players are perfectly rational with a high degree of sensitivity towards expected payoffs, such that they will almost surely choose their best alternative. Thus, QRE is a generalization of NE that nests the latter when the payoff sensitivity is infinitely large.
In addition, the logistic quantal response function possesses the attractive properties of interiority, monotonicity and continuity. Interiority gurarantees that choice probabilities are in the open interval $]0,1[$, allowing for scenarios where literally anything can happen, while monotonicity in expected payoffs implies that a pure strategy with higher expected payoff is chosen with higher probability than an action to which is associated a lower expected gain. The assumption of continuity requires that small disturbances in a player's strategy will only cause slight changes in every other player's strategy, hence avoiding brutal jumps in choice probabilities. Under these conditions, the step functions of best-responding players are replaced by smooth quantal response functions reflecting "better-responding" behavior.

Since the groudbreaking work on QRE [23], there has been several implementations in game-theoretic settings, namely in all-pay auctions (Anderson, Goeree and Holt [32]), alternating-offer bargaining games (Goeree and Holt [33]), posted-offer markets (Acevedo [34]), coordination games (Anderson, Goeree and Holt [35]) and also in the study of individual behavior on the formation of irrational bubbles (Moinas and Pouget [36]). In Section 2.2, we present the formal structure of the quantal choice model and follow the philosophy of McFadden [37] to derive the logistic quantal response function proposed by McKelvey and Palfrey [23]. For the sake of illustration, this section also provides an application of logit QRE (LQRE) to a simple two-player zero-sum game.

2.2 The quantal response model

Let $\Gamma_N = \left[N, \{\Delta A_i\}_{i=1}^n, \{u_i(\cdot)\}_{i=1}^n\right]$ be a strategic form game with $N = \{1, 2, ..., n\}$ a finite set of players , $A_i = \{a_{i1}, a_{i2}, ..., a_{iJ_i}\}$ is the set of strategies of player $i \in N$ consisting of J_i pure strategies and $u_i : A \to \mathbb{R}$ represents the payoff function of player i with A representing the Cartesian product $A = A_1 \times ... \times A_n$, such that a pure strategy profile is denoted by $a = (a_1, ..., a_n) \in A$.

Let $\sigma_i \in \Delta A_i$ be a mixed strategy for player $i \in N$, which assigns to each pure strategy $a_{ik} \in A_i$ a probability $0 \leq \sigma_{ik} \leq 1$ that it will be played and denote by $\Delta A_i = \left\{\sigma_i \in \mathbb{R}^{J_i} | \sigma_{ik} \geq 0 \; for \; k = 1, 2, ..., J_i \; and \; \sum_{k=1}^{J_i} \sigma_{ik} = 1\right\}$ the $(J_i - 1)$ dimensional simplex of player $i \in N$, also called mixed extension of A_i. In addition, we denote the product set of mixed extensions as follows $\Delta A = \Delta A_1 \times ... \times \Delta A_n$, with mixed strategy profiles $\sigma = (\sigma_1, ..., \sigma_n) \in \Delta A$. Accordingly, given a mixed strategy profile $\sigma \in \Delta A$, the expected payoff of player $i \in N$ has a domain extended to ΔA and is given by:

$$u_i(\sigma) = \sum_{a \in A} \Big(\prod_{k=1}^{n} \sigma_k(a_k)\Big) u_i(a)$$

where $\sigma_k(a_k)$ is player k's mixed strategy component associated to his pure strategy a_k in the outcome $a \in A$. Likewise, the expected gain of player $i \in N$ when playing pure strategy $a_{ij} \in A_i$ is:

$$u_i(a_{ij}, \sigma_{-i}) = \sum_{a_{-i} \in A_{-i}} \Big(\prod_{\substack{k=1 \\ k \neq i}}^{n} \sigma_k(a_k)\Big) u_i(a_{ij}, a_{-i})$$

where $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$ is the $(n-1)$ vector of pure strategies for players other than i and σ_{-i} is defined alike.

2.3 The logistic quantal response function

In order to introduce the structure underlying individual choice behavior, let $E_i = \mathbb{R}^{J_i}$ be the space of expected payoffs related to strategies available to player $i \in N$ and let the product set $\Xi = E_1 \times ... \times E_n$ represent the entire payoff space for all players.

This allows us to define the map $\bar{u} : \Delta A \to \Xi$ as $\bar{u}(\sigma) = (\bar{u}_1(\sigma), ..., \bar{u}_n(\sigma))$ with a particular entry being $\bar{u}_{ij}(\sigma) = u_i(a_{ij}, \sigma_{-i})$.

Next, consider that each player's expected gain for each pure strategy is disturbed by a random term as follows:

$$\hat{u}_{ij}(\sigma) = \bar{u}_{ij}(\sigma) + \mathcal{E}_{ij}$$

The rationale behind this disturbance is that players act as statisticians and provide an estimation of the expected payoff for each action, such that the random term $\mathcal{E}_{ij}$ is interpreted as a subjective estimation error. An alternative interpretation following Beja [21] is that players have a target strategy they would like to adopt but fail to do so because of computation errors.

Furthermore, let us denote player i's error vector by $\mathcal{E}_i = (\mathcal{E}_{i1}, ..., \mathcal{E}_{iJ_i})$ with joint probability density function (pdf) $f_i(\varepsilon_i)$ and existing marginal distributions $f_{ij}(\varepsilon_{ij})$ for each $\mathcal{E}_{ij}$ with $E(\mathcal{E}_i) = 0$. When these properties are satisfied for all $i \in N$, we call $f = \{f_1, ..., f_n\}$ a set of *admissible* error distributions. Given this specification, the rule that governs individual choice behavior is that each player will assign a probability mass to a pure strategy $a_{ij} \in A_i$ according to the plausibility of the event $\hat{u}_{ij} \geq \hat{u}_{ik}\ \forall k = 1, ..., J_i,\ k \neq j$. Mathematically, for each $i \in N$ and any $\bar{u} \in \Xi$, the ij-response set $\mathcal{R}_{ij} \subseteq \mathbb{R}^{J_i}$ is defined by

$$\mathcal{R}_{ij}(\bar{u}_i) = \big\{(\varepsilon_{i1}, ..., \varepsilon_{iJ_i}) \in \mathbb{R}^{J_i} | \bar{u}_{ij} + \varepsilon_{ij} \geq \bar{u}_{ik} + \varepsilon_{ik}\ \forall k = 1, ..., J_i,\ k \neq j\big\}$$

It follows that for any given mixed strategy profile $\sigma \in \Delta A$ and a joint PDF f_i, integrating the area above each region and below the density curve, one can compute the probability that each player will assign to a given action. More formally, the probability that player $i \in N$ will choose pure strategy $a_{ij} \in A_i$ is given by:

$$\sigma_{ij}(\bar{u}_i) = \int_{\mathcal{R}_{ij}(\bar{u}_i)} f_i(\varepsilon_i)\ d\varepsilon_i$$

To graphically illustrate the relation between the response set $\mathcal{R}_{ij}$ and the choice probabilities, let us assume for simplicity that player i has two pure strategies $A_i = \big\{a_{i1}, a_{i2}\big\}$ with a vector of expected payoffs $\bar{u}_i = (\bar{u}_{i1}, \bar{u}_{i2}) \in \mathbb{R}^2$ where $\bar{u}_{i1} > \bar{u}_{i2}$.
In addition, let us consider a vector of shocks $\mathcal{E}_i = (\mathcal{E}_{i1}, \mathcal{E}_{i2}) \sim \mathcal{N}_2\big(\mu, \sum\big)$ drawn from a bivariate standard normal distribution with mean $\mu = (0, 0)'$, covariance matrix $\sum = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and joint density $f_i(\varepsilon_i) = (2\pi)^{-1} \exp\big\{-\frac{1}{2}\varepsilon_i^t \varepsilon_i\big\}$. It follows that the response set for the first pure strategy of player i is given by

$$\begin{aligned}\mathcal{R}_{i1} &= \big\{(\varepsilon_{i1}, \varepsilon_{i2}) \in \mathbb{R}^2 | \varepsilon_{i2} \leq (\bar{u}_{i1} - \bar{u}_{i2}) + \varepsilon_{i1}\big\} \\ &= \big\{(\varepsilon_{i1}, \varepsilon_{i2}) \in \mathbb{R}^2 | \varepsilon_{i2} \leq \Delta\bar{u}_i + \varepsilon_{i1}\big\}\end{aligned}$$

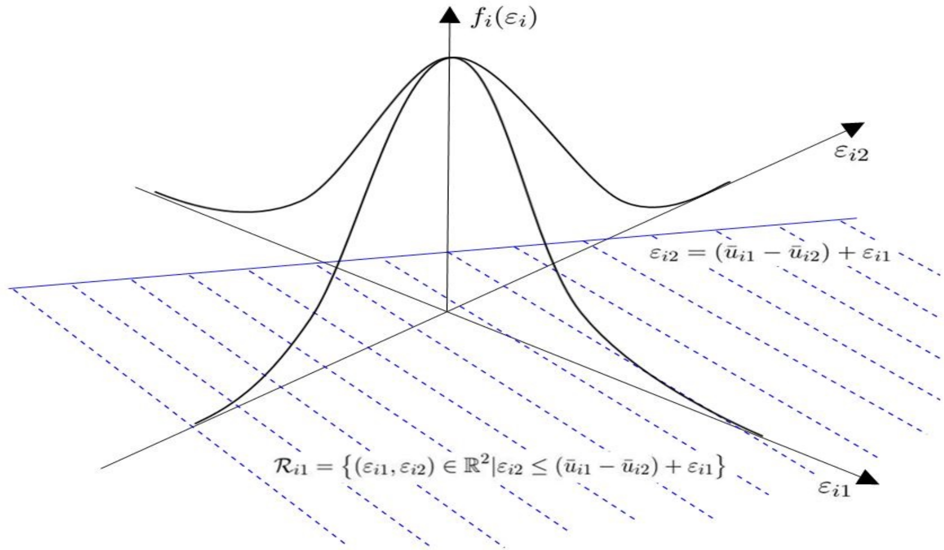

Figure 1: Illustration of a response set in $\mathbb{R}^2$

In Figure 1, the $\mathcal{R}_{i1}$ response set of player i is represented by the dashed surface, such that integrating the joint density over the latter gives us the choice probability associated to pure strategy $a_{i1} \in A_i$, formally we have

$$\sigma_{i1}(\bar{u}_{i1}, \bar{u}_{i2}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{\Delta \bar{u}_i + \varepsilon_{i1}} f_i(\varepsilon_{i1}, \varepsilon_{i2}) \; d\varepsilon_{i2} \; d\varepsilon_{i1}$$

The property of monotonicity in expected payoffs, i.e. that actions with larger expected gains are attributed more probability mass can also be noticed in Figure 1. Indeed, in this particular example when $\bar{u}_{i1} > \bar{u}_{i2}$ the closed half-plane $\mathcal{R}_{i1}$ supports a larger portion of $f_i(\varepsilon_i)$ than $\mathcal{R}_{i2}$, implying that $\sigma_{i1} > \sigma_{i2}$.

It is also important to emphasize that the property of monotonicity is not satisfied by any arbitrary disturbance distribution. In order to avoid anti-rational behavior and have stochastic choice models that are suitable for empirical analysis, an axiomatic approach to QRE has been proposed by Goeree, Holt and Palfrey [38] in which common sense restrictions that are both intuitively and economically convincing are imposed on choice probabilities to obtain a *regular* quantal response function. In fact, without restricting the set of possible error distributions and by considering completely general probabilistic choice models with an additive random utility structure, Haile et al. [39] have shown that any pattern of choice probabilities can be fitted perfectly.

Thus, without restrictions on statistical response functions, quantal choice models would be empirically meaningless in the sense that there always exist an error structure that perfectly matches the observed choice behavior of players.

Definition 1. *(Goeree et al. 2005)* $\sigma_i : \mathbb{R}^{J_i} \to \Delta A_i$ *is a regular quantal response function if it satisfies the following four axioms.*

1. *Interiority:* $\sigma_{ij}(\bar{u}_i) > 0 \quad \forall j = 1, ..., J_i$ *and* $\forall \bar{u}_i \in \mathbb{R}^{J_i}$

2. *Continuity:* $\sigma_{ij}(\bar{u}_i)$ *is continuous and differentiable* $\forall \bar{u}_i \in \mathbb{R}^{J_i}$

3. *Responsiveness:* $\frac{\partial \sigma_{ij}(\bar{u}_i)}{\partial \bar{u}_{ij}} > 0 \quad \forall j = 1, ..., J_i$ *and* $\forall \bar{u}_i \in \mathbb{R}^{J_i}$

4. *Monotonicity:* $\bar{u}_{ij} > \bar{u}_{ik} \Rightarrow \sigma_{ij}(\bar{u}_i) > \sigma_{ik}(\bar{u}_i) \quad \forall j, k = 1, ..., J_i$ *and* $j \neq k$

An interesting parametric class of statistical response functions that emerges frequently in the literature and satisfies all the four properties of Definition 1 is the logistic quantal response function which takes the following form

$$\sigma_{ij}(\bar{u}_i) = \frac{\exp(\lambda \bar{u}_{ij})}{\sum_{k=1}^{J_i} \exp(\lambda \bar{u}_{ik})} \tag{1.1}$$

where $\lambda \in \mathbb{R}_{++}$ is a parameter representing the "degree of rationality" (or payoff responsiveness) of players, so that when $\lambda \to 0$ the probability distribution becomes uniform $\sigma_{ij} \to \frac{1}{J_i} \ \forall j = 1, ..., J_i$ representing a low degree of rationality, while on the other side $\lambda \to \infty$ means that players are infinitely sensible towards their expected payoffs and will play their best response almost surely. Following McFadden [37] and the early work of Marschak [40] and Luce and Suppes [41] one can show that (1.1) is the resulting choice behavior when the underlying error structure has an extreme value type-I distribution also called Gumbel distribution with cumulative distribution function (cdf)

$$F_i(\varepsilon_{ij}) = \exp(-\exp(-\lambda \varepsilon_{ij} - \gamma)) \tag{1.2}$$

where $\lambda > 0$ is the scale parameter and γ is the Euler-Mascheroni constant [4]. In Lemma 1, we explore the relationship between the Gumbel error structure and the logistic quantal response function in (1.1). The proof is given in Appendix.

Lemma 1. *Assume that the error structure of player* $i \in N$ *consists in i.i.d. random shocks* $\mathcal{E}_{i1}, ..., \mathcal{E}_{iJ_i}$ *following the distribution given in* (1.2)*. Then, the quantal response function representing individual choice behavior takes the logistic form as specified in equation* (1.1)*.*

[4] We will consider Gumbel distributions where the location parameter is normalized to 0

Taking the Gumbel distribution as an error structure also allows to capture the property of unbiasedness, i.e. the estimate by player $i \in N$ about the expected gain of action $a_{ij} \in A_i$ will on average be equal to the expected payoff associated to the latter strategy, which is computed from the equilibrium probability distribution of other players. In addition, the variance of a player's estimated gain of adopting a particular pure strategy is $o(\lambda^{-1})$, reflecting that when learning (or experience) takes place through an increase of the rationality index λ, a player is able to make more precise estimates of his expected payoffs. The latter two statistical properties are shown in Appendix of this paper. Using logistic reaction functions, a QRE is defined as follows.

Definition 2. *Let $\Gamma_N = \left[N, \{\Delta A_i\}_{i=1}^n, \{u_i(\cdot)\}_{i=1}^n\right]$ be a game in normal form where individual choice behavior follows a logistic quantal response function for each player. A logit QRE (LQRE) is a mixed strategy profile $\sigma^* \in \Delta A$, such that for all $i \in N$ and $1 \leq j \leq J_i$, we have*

$$\sigma_{ij}^* = \frac{\exp(\lambda \bar{u}_{ij}(\sigma_{-i}^*))}{\sum_{k=1}^{J_i} \exp(\lambda \bar{u}_{ik}(\sigma_{-i}^*))}$$

3 Endogeneous Quantal Response Equilibrium

3.1 Motivation for endogenizing learning effects

In order to assess the goodness of fit of the logistic version of the QRE model, [23] have explored several experiments on normal form games based on [27], [28], [29] and [30]. The analyzed experiments consist in repeated two-person games with unique Nash equilibria, where the total number of repetitions is broken down in various "experience levels" containing a fixed number of successive plays of the game. In each experience level, the rationality index λ is estimated by maximum likelihood and the corresponding QRE is computed. A salient feature regarding the evolution of the (estimated) rationality level is its tendency to increase across game repetitions. This phenomenon can be explained by a learning effect in which players gain insights through repeated observations of the actual payoffs received from choosing different actions, such that their estimation of expected payoffs becomes more accurate across time. It is also important to emphasize that this increase is not systematic and in some experiments λ remains relatively low even in later periods, indicating non-convergence (or very slow convergence) towards the Nash equilibrium of the game.

The study of the evolution of the rationality index over repetitions of a game is tightly related to learning curve theory which is prominent in the literature of psychology and psychometrics. A learning curve also called experience curve can be defined as a graphical representation of how skills, knowledge and experience evolve over time. According to [42], a typical curve of learning displays a strong rate of change in early periods followed by a deceleration to a horizontal slope as practice advances. This approach is used by [43] in an experimental study of the centipede game, where an exponential learning curve of the form $\varepsilon(t) = \varepsilon \exp(-\delta(t-1))$ is used to model the decline of noisy play with experience. In this parametric curve, ε is the initial error level, while δ is the learning rate affecting the convergence speed of errors to zero. [44] use a similar parametric form for the rationality index $\lambda(t) = \underline{\lambda} \exp(\delta(t-1))$ to study a boundedly rational route choice model, where δ is again the rate of learning. Modeling the rationality index dynamics by a learning curve may be a first naive approach to endogenize learning. The exponential patterns proposed by [43] and [44] are nevertheless too simplistic to capture the non-monotonic evolution of λ over game repetitions. Hence, a first attempt to naively describe the learning dynamics in the spirit of learning curve theory would be to propose a refined law of motion for λ that allows for fluctuations/cycles in the error structure governing player behavior. Another attempt to endogenize learning effects would be to allow for heterogeneous rationality where players are able to choose their rationality level according to a cost-benefit tradeoff. More precisely, players are allowed to have different rationality indices and are able to affect their own rationality parameter via the choice of an effort level to which is associated a fixed marginal cost. This two approaches are developed in the next sections.

3.2 Learning curve theory and the rationality index dynamics

A learning curve also called experience curve can be defined as a graphical representation of how skills, knowledge and experience evolve over time.[5]Following the observations of [42], the general shape of a learning curve displays a strong rate of change in early periods and a decline of the improvement rate as the number of periods increases. More recent experiments and results from psychometrics tend to corroborate the flattening of the curve as practice advances. Of course, the forms of the learning curves are wide ranging since they depend on the system under study such that various shapes can be found in the literature, namely, Sigmoid or exponential functions, logistic, Gompertz and von Bertalanffy functions.

[5]For a thorough discussion of the learning curve consult [45].

A refined law of motion for the learning rate is given by the following non-autonomous first-order differential equation:

$$\dot{\lambda}(t) = \delta(\bar{\lambda} - \lambda(t)) + Ae^{-t\tau_d}\left[\sin(\omega t)(\delta - \tau_d) + \cos(\omega t)\omega\right],$$

where $[\underline{\lambda}, \bar{\lambda}]$ is a closed interval representing a player's lower and upper bound of his learning capacity with $0 < \underline{\lambda} < \bar{\lambda} < \infty$. The learning rate of a player is represented by $\delta \in \mathbb{R}_{++}$ and measures the speed of convergence towards his upper bound $\bar{\lambda}$ as periods unfold. $A \in \mathbb{R}_{++}$ measures the amplitude of the cyclical component i.e. of fluctuations in λ and $\tau_d \in \mathbb{R}_{++}$ is called damping coefficient and acts as a reduction in a player's fluctuating behavior. Indeed, as $\tau_d \to \infty$ then the sine and cosine functions disappear implying a simple first order differential equation whose solution is a smooth concave curve. Finally, $\omega \in \mathbb{R}_{++}$ stands for the frequency of the cyclical component and so represents the frequency with which the rationality index undergoes fluctuations.
To find the rationality index dynamics, we need to solve the differential equation $\dot{\lambda}(t) = f(t, \lambda(t))$ given an initial condition $\lambda(0) = \underline{\lambda}$, where the function $f : \mathcal{U} \to \mathbb{R}$ is defined on a compact set $\mathcal{U} = [0, T]$ x $[\underline{\lambda}, \bar{\lambda}]$ in the euclidean space $\mathbb{R}^2$. This means that we have a Cauchy problem which is formally written as:

$$\begin{cases} \dot{\lambda}(t) = \delta(\bar{\lambda} - \lambda(t)) + Ae^{-t\tau_d}\left[\sin(\omega t)(\delta - \tau_d) + \cos(\omega t)\omega\right] \\ \lambda(0) = \underline{\lambda} \end{cases}$$

Proposition 1. *There exists a unique solution to the Cauchy problem involving the law of motion of the rationality index $\lambda(t)$.*

Proof. By rewriting the differential equation as $\dot{\lambda}(t) + \Gamma_1(t)\lambda(t) = \Gamma_2(t)$ with the functions $\Gamma_1(t) = \delta$ and $\Gamma_2(t) = \delta\bar{\lambda} + Ae^{-t\tau_d}\left[\sin(\omega t)(\delta - \tau_d) + \cos(\omega t)\omega\right]$, we see that $\forall \epsilon > 0$ both $\Gamma_1(t)$ and $\Gamma_2(t)$ are continuous functions on the open interval $0 - \epsilon < t_0 < \infty$ and so by the Linear Fundamental Theorem of Existence and Uniqueness, there exists a unique solution to the Cauchy problem on the interval $t \in [0, \infty)$. $\square$

To obtain the analytical solution to the Cauchy problem, we first solve the trivial homogeneous equation to get a general solution:

$$\dot{\lambda}(t) = -\delta\lambda(t)$$

$$\Leftrightarrow \frac{\dot{\lambda}(t)}{\lambda(t)} = -\delta \Leftrightarrow \int \frac{\dot{\lambda}(t)}{\lambda(t)} dt = -\int \delta dt$$

$$\Leftrightarrow ln\lambda(t) = \delta t + c$$

$$\Leftrightarrow \lambda_g(t) = Be^{-\delta t}, B \in \mathbb{R}$$

So, the general solution to the homogenous equation is $\lambda_g(t) = Be^{-\delta t}$ and the second step consists in finding a particular solution with the time-varying parameter $B(t)$ such that we start our second step with $\lambda_p(t) = B(t)e^{-\delta t}$ and taking the time derivative we obtain:

$$\dot{\lambda}(t) = \dot{B}e^{-\delta t} - \delta Be^{-\delta t}$$

and equating the latter with $\dot{\lambda}(t) = \delta(\bar{\lambda} - \lambda(t)) + Ae^{-t\tau_d}\left[\sin(\omega t)(\delta - \tau_d) + \cos(\omega t)\omega\right]$, we get:

$$\dot{B}(t) = \delta\bar{\lambda}e^{\delta t} + Ae^{(\delta-\tau_d)t}\sin(\omega t)(\delta - \tau_d) + Ae^{(\delta-\tau_d)t}\cos(\omega t)\omega$$

$$\Leftrightarrow \int \dot{B}(t)dt = \underbrace{\delta\bar{\lambda}\int e^{\delta t}dt}_{(I)} + A\,(\delta-\tau_d)\underbrace{\int e^{(\delta-\tau_d)t}\sin(\omega t)dt}_{(II)} + A\,\omega\underbrace{\int e^{(\delta-\tau_d)t}\cos(\omega t)dt}_{(III)} \qquad (\star\star)$$

For the sake of clarity, we will solve (I), (II) and (III) separately:

The solution of (I) is trivial and immediate and gives us $\bar{\lambda}e^{\delta t}$.

The solution of (II) however is not immediate and requires twice the application of integration by parts:

Let $u(t) = e^{(\delta-\tau_d)t}$ so that $du(t) = (\delta - \tau_d)e^{(\delta-\tau_d)t}dt$ and let us also take $dv(t) = \sin(\omega t)dt$ such that $v(t) = \frac{-\cos(\omega t)}{\omega}$. Integrating by parts gives us:

$$\int e^{(\delta-\tau_d)t}\sin(\omega t)dt = \frac{-e^{(\delta-\tau_d)t}cos(\omega t)}{\omega} + \int \frac{(\delta - \tau_d)e^{(\delta-\tau_d)t}\cos(\omega t)}{\omega}dt$$

$$\Leftrightarrow \int e^{(\delta-\tau_d)t}\sin(\omega t)dt = \frac{-e^{(\delta-\tau_d)t}\cos(\omega t)}{\omega} + \frac{(\delta - \tau_d)}{\omega}\int e^{(\delta-\tau_d)t}\cos(\omega t)dt$$

Defining $u(t) = e^{(\delta-\tau_d)t}$, $du(t) = (\delta - \tau_d)e^{(\delta-\tau_d)t}dt$ and $dv(t) = \cos(\omega t)dt$, $v(t) = \frac{\sin(\omega t)}{\omega}$ and integrating by parts a second time:

$$\int e^{(\delta-\tau_d)t}\sin(\omega t)dt = \frac{-e^{(\delta-\tau_d)t}\cos(\omega t)}{\omega} + \frac{(\delta-\tau_d)}{\omega}\left\{\frac{e^{(\delta-\tau_d)t}\sin(\omega t)}{\omega} - \frac{(\delta-\tau_d)}{\omega}\int e^{(\delta-\tau_d)t}\sin(\omega t)dt\right\}$$

$$\Leftrightarrow \int e^{(\delta-\tau_d)t}\sin(\omega t)dt = \frac{-e^{(\delta-\tau_d)t}\cos(\omega t)}{\omega} + \frac{(\delta-\tau_d)}{\omega^2}e^{(\delta-\tau_d)t}\sin(\omega t) - \frac{(\delta-\tau_d)^2}{\omega^2}\int e^{(\delta-\tau_d)t}\sin(\omega t)dt$$

$$\Leftrightarrow \left(\frac{(\delta-\tau_d)^2+\omega^2}{\omega^2}\right)\int e^{(\delta-\tau_d)t}\sin(\omega t)dt = \frac{(\delta-\tau_d)}{\omega^2}e^{(\delta-\tau_d)t}\sin(\omega t) - \frac{e^{(\delta-\tau_d)t}\cos(\omega t)}{\omega}$$

and hence the final result for (II) is:

$$\int e^{(\delta-\tau_d)t}\sin(\omega t)dt = \left(\frac{(\delta-\tau_d)}{(\delta-\tau_d)^2+\omega^2}\right)e^{(\delta-\tau_d)t}\sin(\omega t) - \left(\frac{\omega}{(\delta-\tau_d)^2+\omega^2}\right)e^{(\delta-\tau_d)t}\cos(\omega t)$$

To derive the solution of (III), we define $u(t) = e^{(\delta-\tau_d)t}$, $du(t) = (\delta-\tau_d)e^{(\delta-\tau_d)t}dt$ and
$dv(t) = \cos(\omega t)dt$, $v(t) = \frac{\sin(\omega t)}{\omega}$ and we integrate by parts:

$$\int e^{(\delta-\tau_d)t}\cos(\omega t)dt = \frac{e^{(\delta-\tau_d)t}\sin(\omega t)}{\omega} - \frac{(\delta-\tau_d)}{\omega}\int e^{(\delta-\tau_d)t}\sin(\omega t)dt$$

We again define $u(t) = e^{(\delta-\tau_d)t}$, $du(t) = (\delta-\tau_d)e^{(\delta-\tau_d)t}dt$ and $dv(t) = \sin(\omega t)dt$, $v(t) = \frac{-\cos(\omega t)}{\omega}$ and integrate by parts a second time:

$$\int e^{(\delta-\tau_d)t}\cos(\omega t)dt = \frac{e^{(\delta-\tau_d)t}\sin(\omega t)}{\omega} - \frac{(\delta-\tau_d)}{\omega}\left\{\frac{-e^{(\delta-\tau_d)t}\cos(\omega t)}{\omega} + \frac{(\delta-\tau_d)}{\omega}\int e^{(\delta-\tau_d)t}\cos(\omega t)dt\right\}$$

from there, we apply the same manipulations as in (II) and we obtain the final result for (III):

$$\int e^{(\delta-\tau_d)t}\cos(\omega t)dt = \left(\frac{\omega}{(\delta-\tau_d)^2+\omega^2}\right)e^{(\delta-\tau_d)t}\sin(\omega t) + \left(\frac{(\delta-\tau_d)}{(\delta-\tau_d)^2+\omega^2}\right)e^{(\delta-\tau_d)t}\cos(\omega t)$$

Gathering together (I), (II) and (III), we obtain the solution of equation $(\star\,\star)$:

$$B(t) = \bar{\lambda}e^{\delta t} + Ae^{(\delta-\tau_d)t}\sin(\omega t)$$

so that the particular solution to the non-homogenous differential equation is:

$$\lambda_p(t) = B(t)e^{-\delta t} = \bar{\lambda} + Ae^{-\tau_d t}\sin(\omega t)$$

Hence, we can now combine the general solution to the homogeneous equation and the particular solution to the non-homogeneous equation to obtain $\lambda(t) = \lambda_g(t) + \lambda_p(t)$ which gives us:

$$\lambda(t) = Be^{-\delta t} + \bar{\lambda} + Ae^{-\tau_d t}\sin(\omega t)$$

and the value of B can be discovered by the initial condition $\lambda(0) = \underline{\lambda}$ implying that $B = (\underline{\lambda} - \bar{\lambda})$ such that our final solution for the rationality index is the following:

$$\lambda(t) = \underline{\lambda}e^{-\delta t} + \bar{\lambda}(1-e^{-\delta t}) + A\ \sin(\omega t)e^{-\tau_d t}$$

The above dynamics allows to incorporate fluctuation behaviors and cycles after an appropriate calibration of the parameters.

3.3 The QRE model with endogeneous learning

Let $\Gamma_N^{\Lambda} = \left[N, \{\Delta, A_i\}_{i=1}^n, \{u_i(\cdot)\}_{i=1}^n, \Lambda\right]$ be a normal form game where Λ is the population rationality level and belongs to some parametric family denoted by $\mathcal{P} = \left\{P_\theta : \theta \in \Theta \subset \mathbb{R}^d\right\}$, where P_θ is a probability distribution on a non-negative domain parametrized by the vector θ and represents a population of players with heterogeneous skills. In period $t = 1$ each player $i \in N$ is endowed with an initial rationality level $\lambda_i^{(t=1)} = \lambda_i^{(1)}$ drawn from P_θ. The logit QRE in the first period is denoted by $\text{LQRE}^{(1)}$ and is computed using the initial rationality endowment of the players. As the QRE model is a statistical model, we assume that players act as statisticians and estimate the rationality level of their rivals. After the first play of the game each player observes the logit equilibrium probability distributions (his own and that of his opponents). These observations play the same role as a data sample in analogy with a statistical model and will be used by each player to construct an estimator of his rival's rationality index. The estimator of player j of the rationality index of player $i \neq j$ in period $t \geq 1$ is defined as:

$$\hat{\lambda}_i^{(t)} = \lambda_i^{(t)} + \eta_{i|[-\lambda_i^{(t)},+\infty)} \tag{1.3}$$

Equation (1.3) means that player j's estimate of player i's rationality level is equal to the true rationality level of player i in period t plus a noise with a truncated normal distribution $\eta_i \sim \mathcal{N}(0, \sigma_j^2)$ on the domain $[-\lambda_i^{(t)}, +\infty)$, where $\sigma_j^2 = 1/\sqrt{\lambda_j^{(1)}}$. Note that the variance of the noise equals the inverse square root of the initial rationality index of player j, such that the larger the rationality endowment of player j, the less volatile is his estimator around the true value $\lambda_i^{(t)}$. Another desirable characteristic of the estimator in (1.3) is that it should have the property of consistency, i.e. both bias and variance of a player's estimator should decrease with his initial rationality endowment.

Definition 3. *Player j's estimator of player i's rationality index in period $t \in \mathbb{N}$ is asymptotically consistent if $MSE\big[\hat{\lambda}_i^{(t)}\big] = Bias^2\big[\hat{\lambda}_i^{(t)}\big] + Var\big[\hat{\lambda}_i^{(t)}\big] \to 0$ as $\lambda_j^{(1)} \to +\infty$.*

Proposition 2. *The belief structure in (1.3) is asymptotically consistent.*

Proof. Let us denote the lower and upper bounds of the domain of the distribution of the noise by $a = -\lambda_i^{(t)}$ and $b = +\infty$. In addition, let us define $\alpha_1 = \frac{a-\mu}{\sigma_j}$ and $\alpha_2 = \frac{b-\mu}{\sigma_j} = +\infty$ where $\mu = 0$ is the mean of the noise while the standard deviation is $\sigma_j = \sqrt{\sigma_j^2} = \sqrt{1/\sqrt{\lambda_j^{(1)}}}$. In addition, let ϕ and Φ denote the pdf and cdf of a standard normal distribution $Z \sim \mathcal{N}(0, 1)$.

We have $Bias\big[\hat{\lambda}_i^{(t)}\big] = E\big[\lambda_i^{(t)}\big] - \lambda_i^{(t)} = E\big[\eta_i|_{[-\lambda_i^{(t)},+\infty)}\big]$ and for a truncated normal distribution the latter expectation equals:

$$\begin{aligned} E\big[\eta_i|_{[-\lambda_i^{(t)},+\infty)}\big] &= \mu + \sigma_j \frac{\phi(\alpha_1)}{1 - \Phi(\alpha_1)} \\ &= \sqrt{1/\sqrt{\lambda_j^{(1)}}}\frac{\phi(\alpha_1)}{1 - \Phi(\alpha_1)} \end{aligned}$$

Letting $\lambda_j^{(1)} \to +\infty$ implies the following convergence $\phi(\alpha_1) \to 0$, $\Phi(\alpha_1) \to 0$ such that $\lim_{\lambda_j^{(1)} \to +\infty} Bias\big[\hat{\lambda}_i^{(t)}\big] = 0$.

For the variance, we have $Var\big[\hat{\lambda}_i^{(t)}\big] = Var\big[\eta_i|_{[-\lambda_i^{(t)},+\infty)}\big]$ and for a truncated normal distribution on $[-\lambda_i^{(t)}, +\infty)$ this variance becomes:

$$Var\big[\eta_i|_{[-\lambda_i^{(t)},+\infty)}\big] = \sigma_j^2 \left\{ 1 - \frac{\alpha_2\phi(\alpha_2) - \alpha_1\phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} - \left(\frac{\phi(\alpha_2) - \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \right)^2 \right\}$$

and since $\alpha_2 = +\infty$ it follows that:

$$Var\big[\eta_i|_{[-\lambda_i^{(t)},+\infty)}\big] = 1/\sqrt{\lambda_j^{(1)}} \left\{ 1 + \frac{\alpha_1\phi(\alpha_1)}{1 - \Phi(\alpha_1)} - \left(\frac{\phi(\alpha_1)}{1 - \Phi(\alpha_1)} \right)^2 \right\}$$

Accordingly, we have $\lim_{\lambda_j^{(1)} \to +\infty} Var\big[\hat{\lambda}_i^{(t)}\big] = 0$ and so

$$\lim_{\lambda_j^{(1)} \to +\infty} MSE\big[\hat{\lambda}_i^{(t)}\big] = \lim_{\lambda_j^{(1)} \to +\infty} \left(Bias^2\big[\hat{\lambda}_i^{(t)}\big] + Var\big[\hat{\lambda}_i^{(t)}\big] \right) = 0 \quad \square$$

The statement in Proposition 1 is equivalent to the following convergence in probability in a given period t, $\hat{\lambda}_i^{(t)} \xrightarrow{P} \lambda_i^{(t)}$ when $\lambda_j^{(1)} \to +\infty$.

We assume that a player is able to affect his rationality index by choosing an effort level $e_i \geq 0$ to which is associated a fixed marginal cost $\gamma > 0$. In addition, we assume that there exists a "mimick" behavior, i.e. a player's rationality level is attracted towards the estimated average rationality level of his opponents. The expected payoff of player j in period $t+1$ is defined as:

$$\mathbb{E}_{U_j}^{t+1} = \Psi_j\Big(\tilde{\lambda}_j^{(t+1)}(e_j^{(t+1)}, \bar{\lambda}_{-j}^{(t)})\Big) - \gamma e_j^{(t+1)} \tag{1.4}$$

where Ψ_j is player j's reward function of the rationality index $\tilde{\lambda}_j^{(t+1)}$ chosen in period $t+1$ and $\bar{\lambda}_{-j}^{(t)} = \frac{1}{(N-1)} \sum_{\substack{k=1 \\ k \neq j}}^{N} \hat{\lambda}_k^{(t)}$ is the estimated average rationality level of j's opponents. In other words, the expected gain of player j in period $t+1$ is assumed to depend on two key components, namely player j's effort level $e_j^{(t+1)}$ and the estimated average rationality level of his opponents $\bar{\lambda}_{-j}^{(t)}$. Even though at first glance of equation (1.4), there seems to be no relationship between $\mathbb{E}_{U_j}^{t+1}$ and the payoffs of the game, note that each $\hat{\lambda}_k$ is actually an implicit or explicit function of the payoffs of the game. In fact, given the observed probability distribution on the pure strategies of the players at time t and given the payoffs of the game, one is able to determine via the logistic quantal response function the true value of a player's rationality index at time t. In that sense we have a dependence structure between $\mathbb{E}_{U_j}^{t+1}$ and the payoffs of the game. For the sake of clarity, consider the following generic payoff matrix:

$$\begin{array}{cc} & \begin{array}{cc} \sigma_{21} & 1-\sigma_{21} \end{array} \\ \begin{array}{c} \sigma_{11} \\ 1-\sigma_{11} \end{array} & \begin{pmatrix} a \;\; e & b \;\; f \\ c \;\; g & d \;\; h \end{pmatrix} \end{array}$$

Let us define the following quantities $z_1 = c+b-d-a$, $z_2 = d-b$, $z_3 = f+g-h-e$ and $z_4 = h-g$. One can easily show that in this simple setting, there exists an explicit relationship between the rationality index estimators (at time t) and the payoffs as follows:

$$\hat{\lambda}_2^{(t)} = \log\left(\frac{1-\sigma_{21}^{(t)}}{\sigma_{21}^{(t)}}\right)\left\{\frac{1}{\sigma_{11}^{(t)} z_3 + z_4}\right\} + \eta_2|_{[-\lambda_2^{(t)},+\infty)}$$

$$\hat{\lambda}_1^{(t)} = \log\left(\frac{1-\sigma_{11}^{(t)}}{\sigma_{11}^{(t)}}\right)\left\{\frac{1}{\sigma_{21}^{(t)} z_1 + z_2}\right\} + \eta_1|_{[-\lambda_1^{(t)},+\infty)}$$

When the number of players and/or number of pure strategies exceeds 2, the above relation is usually more complex.

The reward function in (1.4) is a map from a player's rationality level in $t+1$ to a real number. Put more simply, it is a function acting as a utility derived from rationality and satisfies the following conditions:

$$\frac{\partial \Psi_j}{\partial \tilde{\lambda}_j^{(t+1)}} > 0 \quad \frac{\partial^2 \Psi_j}{\left(\partial \tilde{\lambda}_j^{(t+1)}\right)^2} < 0$$

In addition, we assume that the rationality index increases with effort and that the marginal increase in rationality from an additional unit of effort falls with the effort level, resulting in a decreasing marginal benefit of effort. The last condition is a positive relationship between the estimated average rationality level and the rationality level of player j.

$$\frac{\partial \tilde{\lambda}_j^{(t+1)}}{\partial e_j^{(t+1)}} > 0 \quad \frac{\partial^2 \tilde{\lambda}_j^{(t+1)}}{\left(\partial e_j^{(t+1)}\right)^2} < 0 \quad \frac{\partial \tilde{\lambda}_j^{(t+1)}}{\partial \bar{\lambda}_{-j}^{(t)}} > 0$$

In period $t \geq 2$ a player j chooses his effort level e_j to maximize his expected payoff, such that the objective can be written as

$$\max_{e_j^{(t+1)}} \mathbb{E}_{U_j}^{t+1} = \Psi_j\left(\tilde{\lambda}_j^{(t+1)}(e_j^{(t+1)}, \bar{\lambda}_{-j}^{(t)})\right) - \gamma e_j^{(t+1)} \tag{1.5}$$

The first order condition imposes that the marginal benefits from a rationality level must be equal to marginal cost:

$$\frac{\partial \Psi_j}{\partial \tilde{\lambda}_j^{(t+1)}} \frac{\partial \tilde{\lambda}_j^{(t+1)}}{\partial e_j^{(t+1)}} = \gamma \tag{1.6}$$

Equation (1.6) allows us to isolate the optimal effort level $e_j^{*(t+1)}$, such that the optimal rationality level is $\tilde{\lambda}_j^{*(t+1)} = \tilde{\lambda}_j^{(t+1)}(e_j^{*(t+1)}, \bar{\lambda}_{-j}^{(t)})$. Finally, we assume that a player's rationality level in period $t+1$ is sensitive towards his previous period level at rate $\phi_j \in [0,1]$, such that the rationality index chosen by player j is written as

$$\lambda_j^{*(t+1)} = \tilde{\lambda}_j^{*(t+1)} + \phi_j \sqrt{\lambda_j^{*(t)}} \tag{1.7}$$

The above model allows us to define an enriched version of logit QRE where the choice of the rationality index is endogenously determined by an effort level.

Definition 4. *Let $\Gamma_N^\Lambda = \left[N, \{\Delta, A_i\}_{i=1}^n, \{u_i(\cdot)\}_{i=1}^n, \Lambda\right]$ be a game in normal form, where individual choice behavior follows a logistic quantal response function and where the population rationality level belongs to a parametric family with a distribution on a non-negative domain. An Endogeneous Quantal Response Equilibrium (EQRE) in period $t \in \mathbb{N}$ consists in a mixed strategy profile $\sigma^{*(t)} \in \Delta A$, such that for all $i \in N$ and $1 \leq j \leq J_i$, we have*

$$\sigma_{ij}^{*(t)} = \frac{\exp\left(\lambda_i^{*(t)} \bar{u}_{ij}(\sigma_{-i}^{*(t)})\right)}{\sum_{k=1}^{J_i} \exp\left(\lambda_i^{*(t)} \bar{u}_{ik}(\sigma_{-i}^{*(t)})\right)}$$

where $\lambda_i^{*(t)} \sim P_\theta$ for $t = 1$, $\lambda_i^{*(t)} = \tilde{\lambda}_i^{*(t)}(e_i^{*(t)}, \bar{\lambda}_{-i}^{(t-1)}) + \phi_i \sqrt{\lambda_i^{*(t-1)}}$ for $t \geq 2$ and $e_i^{*(t)}$ determined by (1.6).

3.4 Simulation study on some 2×2 games

In this section, we explore the quantal response model with endogeneous learning in simulation studies involving simple 2×2 games. There are two players $N = \{1, 2\}$ and $T = 100$ periods. We assume a population of players with a rationality level that is gamma distributed, $\Lambda \sim \mathcal{G}(2, 0.4)$ and a unit marginal cost of effort $\gamma = 1$. In addition, we assume the following reward function $\Psi_i\big(\tilde{\lambda}_i^{(t+1)}\big) = \sqrt{\tilde{\lambda}_i^{(t+1)}}$ where $\tilde{\lambda}_i^{(t+1)} = \big(e_i^{(t+1)}\big)^{1/2}\big(\hat{\lambda}_j^{(t)}\big)^{1/2}$ $i \neq j$ and a sensitivity parameter $\phi_i = 1$ for $i = 1, 2$. Under this specification, it can easily be shown that the optimal effort level resulting from (1.6) is given by:

$$e_i^{*(t+1)} = \frac{\big(\hat{\lambda}_j^{(t)}\big)^{(1/3)}}{(4\gamma)^{(4/3)}} \ i \neq j \tag{1.8}$$

It follows that the rationality index of player i will have the following dynamic:

$$\lambda_i^{*(t+1)} = \left(\frac{\hat{\lambda}_j^{(t)}}{4\gamma}\right)^{2/3} + \phi_i \sqrt{\lambda_i^{*(t)}} \quad i \neq j \quad for \quad t = 1, ..., T-1. \tag{1.9}$$

The first game (Game 1) that we consider is the following zero-sum game:

$$\begin{array}{c} \\ \sigma_{11} \\ 1-\sigma_{11} \end{array} \begin{array}{c} \begin{array}{cc} \sigma_{21} & \quad 1-\sigma_{21} \end{array} \\ \begin{pmatrix} -1 \;\; 1 & 0 \;\; 0 \\ 0.5 \;\; -0.5 & -1 \;\; 1 \end{pmatrix} \end{array}$$

As can be observed in Figure 2, the model is able to reproduce some experimental characteristics of learning. Figure 2 (a) shows the evolution of the rationality index of each player across periods $t = 1, ..., 100$. We see that the rationality index of each player strongly increases in early periods reflecting a large propensity to learn in the first repetitions of the game as suggested by [42]. Note also that there is a stabilization pattern, i.e. after the increase the rationality index of both players has a stochastic fluctuation between two bounds (here between 1.5 and 2.5). Figure 2 (b) represents the time series of the EQRE, i.e. the series of equilibrium probabilities $\sigma_{11}^{*(t)}$ and $\sigma_{21}^{*(t)}$. The equilibrium probabilities with endogeneous learning do not converge to the Nash equilibrium of the game $\mathcal{N}_m = \big((3/5, 2/5), (2/5, 3/5)\big)$ even in later periods. This lack of convergence can be explained by the presence of a marginal cost of effort γ in equation (1.9).

From the latter equation, it can be seen that as the marginal cost of effort vanishes, $\gamma \to 0$, the rationality index of a player will grow larger with the implication that the EQRE will converge to the Nash equilibrium of the game.

This situation is illustrated in Figure 3, where we make the same simulation exercise with a low marginal cost of effort $\gamma = 0.05$. Under these circumstances, the rationality index increases relatively fast as shown in Figure 3 (a). Furthermore, in Figure 3 (b), we see that the EQRE has converged to the Nash equilibrium after 4 periods.

We also apply the endogeneous quantal response model on the coordination game of the battle of the sexes, where the payoff matrix is given by

$$\begin{array}{cc} & \begin{array}{cc} C \ \sigma_{21} & D \ 1-\sigma_{21} \end{array} \\ \begin{array}{c} C \ \sigma_{11} \\ D \ 1-\sigma_{11} \end{array} & \left(\begin{array}{cc} 2 \ \ 1 & 0 \ \ 0 \\ 0 \ \ 0 & 1 \ \ 2 \end{array} \right) \end{array}$$

where C and D stand for coordination and non-coordination respectively. This game has two pure strategy Nash equilibria, namely $\mathcal{N}_p = \{(C, C), (D, D)\}$ and a mixed strategy Nash equilibrium $\mathcal{N}_m = \big((2/3, 1/3), (1/3, 2/3)\big)$. Figure 4 shows the results of the simulation study with a unit marginal cost.

We see in Figure 4 (b) that the EQRE series fluctuates around the pure strategy Nash equilibrium (D, D), meaning that there is a large probability of coordination on this outcome. Interestingly, when the cost of effort becomes large ($\gamma = 10$), we see a tendency of the equilibrium probabilities of playing outcome C to increase for both players, such that the EQRE series has fluctuations closer to the mixed strategy Nash equilibrium of the game. In other words, increasing the marginal cost of effort has decreased the probability of coordination on outcome (D, D).

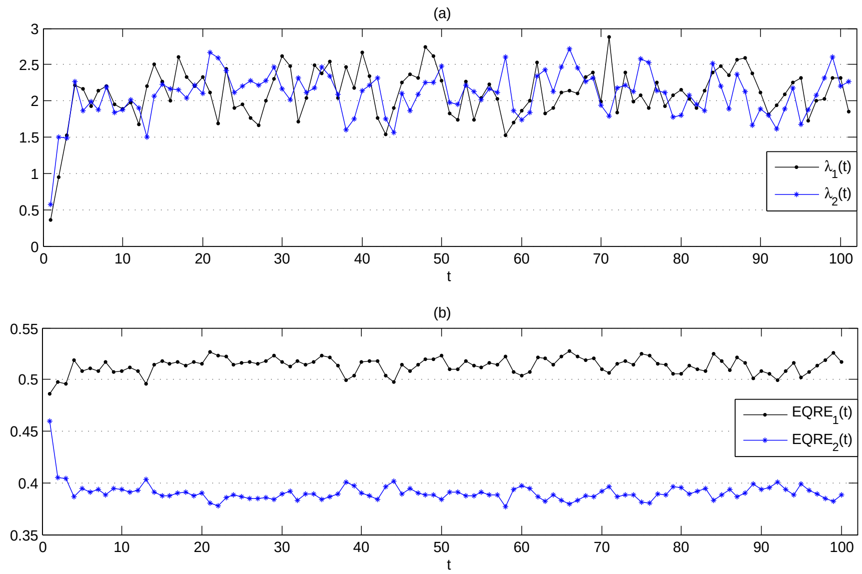

Figure 2: (a) Time series of rationality indices. (b) Evolution of EQRE in Game 1.

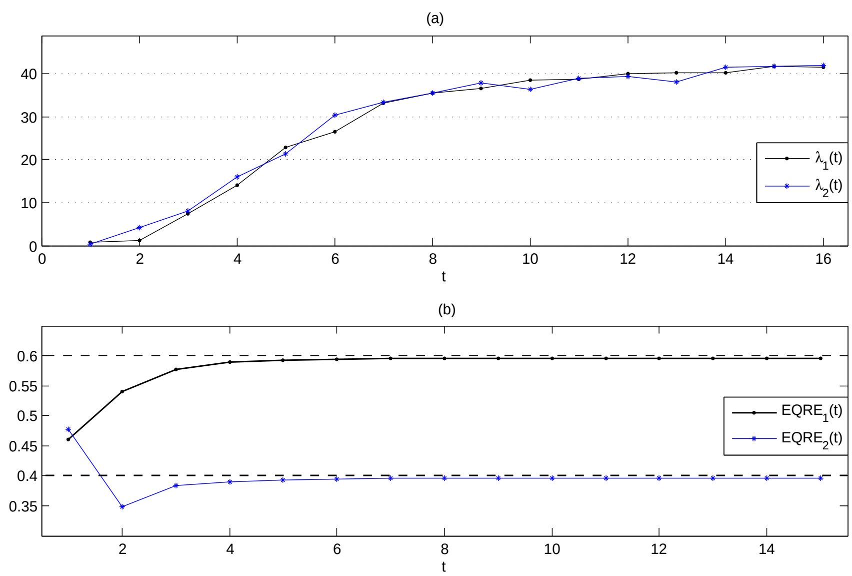

Figure 3: (a) Time series of rationality indices. (b) Evolution of EQRE for Game 1 with small marginal cost of effort.

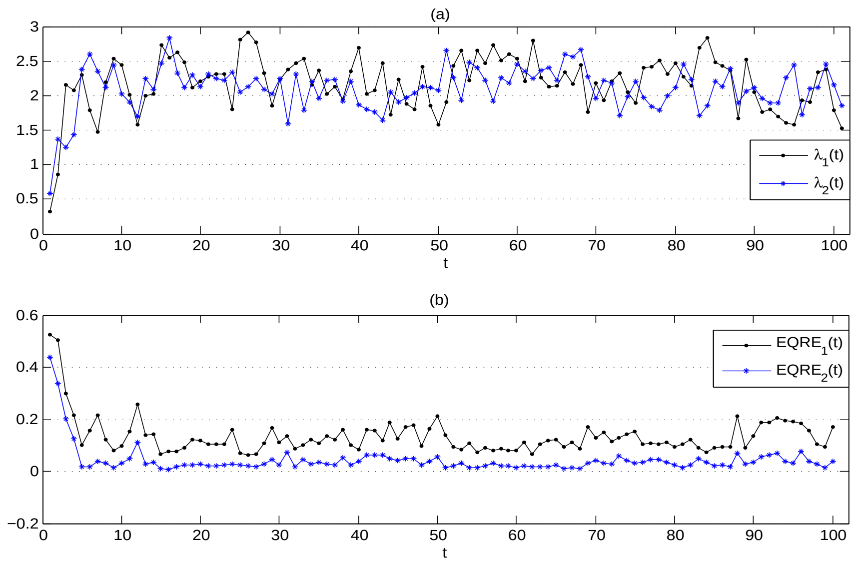

Figure 4: (a) Time series of rationality indices. (b) Evolution of EQRE for the battle of the sexes game.

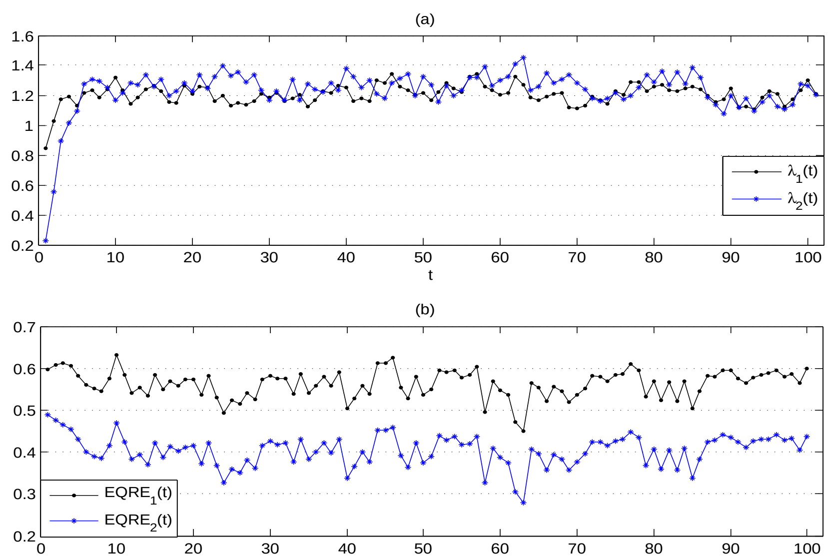

Figure 5: (a) Time series of rationality indices. (b) Evolution of EQRE for the battle of the sexes game with large cost of effort.

Conclusion

This paper aims at enriching bounded rationality models by proposing two different ways for modeling the rationality index dynamics in quantal response settings. After a brief literature review, we present the formal structure of quantal response models and propose two extensions for endogenizing learning effects. The first "naive" approach is based on learning curve theory and consists in giving a parameterized version of the evolution of the rationality index over time. The second approach is more sophisticated and allows for heterogeneous players and endogeneous learning. More formally, each player has an asymptotically consistent estimate of his rival's rationality index and is able to choose his own rationality level according to a cost-benefit tradeoff.
We conclude the paper with a simulation study involving simple 2×2 games. The results are motivating in the sense that the endogeneous quantal response model is able to reproduce some experimental characteristics of learning. An interesting future research direction would be to consider a time-dependent sensitivity framework with respect to previous period rationality levels and to implement laboratory experiments to test the EQRE theory.

Appendix

Proof of Lemma 1

From the ij-response set $\mathcal{R}^{J_i}$ the probability that player $i \in N$ will choose action j is given by:

$$\begin{aligned}
\sigma_{ij}(\bar{u}_i) &= P\big[\hat{u}_{ij} \geq \hat{u}_{ik} \ \ \forall k = 1, ..., J_i \ k \neq j\big] \\
&= P\big[\bar{u}_{ij} + \mathcal{E}_{ij} \geq \bar{u}_{ik} + \mathcal{E}_{ik} \ \ \forall k = 1, ..., J_i \ k \neq j\big] \\
&= P[\mathcal{E}_{ik} \leq \bar{u}_{ij} - \bar{u}_{ik} + \mathcal{E}_{ij} \ \ \forall k = 1, ..., J_i \ k \neq j\big] \\
&= P\big[\mathcal{E}_{i1} \leq \bar{u}_{ij} - \bar{u}_{i1} + \mathcal{E}_{ij}, ..., \mathcal{E}_{i(j-1)} \leq \bar{u}_{ij} - \bar{u}_{i(j-1)} + \mathcal{E}_{ij}, \\
&\quad \mathcal{E}_{i(j+1)} \leq \bar{u}_{ij} - \bar{u}_{i(j+1)} + \mathcal{E}_{ij}, ..., \mathcal{E}_{iJ_i} \leq \bar{u}_{ij} - \bar{u}_{iJ_i} + \mathcal{E}_{ij}\big] \\
&= \int_{-\infty}^{+\infty} \left[\prod_{\substack{k=1 \\ k \neq j}}^{J_i} P\big[\mathcal{E}_{ik} \leq \bar{u}_{ij} - \bar{u}_{ik} + \varepsilon_{ij}\big] \right] f_i(\varepsilon_{ij}) \ d\varepsilon_{ij}
\end{aligned}$$

with pdf $f_i(\varepsilon_{ij}) = \dfrac{dF_i(\varepsilon_{ij})}{d\varepsilon_{ij}} = \lambda \ \exp(-\exp(-\lambda\varepsilon_{ij} - \gamma)) \ \exp(-\lambda\varepsilon_{ij} - \gamma),$ *so we have*

$$\sigma_{ij}(\bar{u}_i) = \int_{-\infty}^{+\infty} \left[\prod_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\exp(-\lambda(\bar{u}_{ij}-\bar{u}_{ik}+\varepsilon_{ij})-\gamma)) \right] \lambda \ \exp(-\exp(-\lambda\varepsilon_{ij}-\gamma)) \ \exp(-\lambda\varepsilon_{ij}-\gamma) \, d\varepsilon_{ij}$$

Let $z = \exp(-\lambda\varepsilon_{ij} - \gamma)$ *, such that the above equation becomes*

$$\sigma_{ij}(\bar{u}_i) = \int_{0}^{+\infty} \left[\prod_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\exp(-\lambda(\bar{u}_{ij} - \bar{u}_{ik}))z) \right] \exp(-z) \ dz$$

$$= \int_{0}^{+\infty} \exp\left(-z \left\{ \sum_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\lambda(\bar{u}_{ij} - \bar{u}_{ik})) \right\} \right) \exp(-z) \ dz$$

$$= \int_{0}^{+\infty} \exp\left(-z \left\{ 1 + \sum_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\lambda(\bar{u}_{ij} - \bar{u}_{ik})) \right\} \right) \ dz$$

$$= -\frac{1}{1 + \sum_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\lambda(\bar{u}_{ij} - \bar{u}_{ik}))} \exp\left(-z \left\{ 1 + \sum_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\lambda(\bar{u}_{ij} - \bar{u}_{ik})) \right\} \right) \Bigg|_{0}^{+\infty}$$

$$= 0 - \left(-\frac{1}{1 + \sum_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(-\lambda(\bar{u}_{ij} - \bar{u}_{ik}))} \right) = \frac{1}{1 + \frac{1}{\exp(\lambda\bar{u}_{ij})} \sum_{\substack{k=1 \\ k\neq j}}^{J_i} \exp(\lambda\bar{u}_{ik})}$$

$$= \frac{\exp(\lambda\bar{u}_{ij})}{\sum_{k=1}^{J_i} \exp(\lambda\bar{u}_{ik})}$$

this completes the proof as we have recovered the expression given in (1.1). □

Some statistical properties associated to estimated gains when errors are Gumbel distributed

Lemma 2. *Assume that the error structure of player $i \in N$ follows (1.2) for all $j = 1, 2, ..., J_i$. Then,*

$$E(\hat{u}_{ij}) = \bar{u}_{ij}$$

$$Var(\hat{u}_{ij}) = \frac{\pi^2}{6}\lambda^{-2}$$

for all $j = 1, 2, ..., J_i$.

Proof. From the pdf $f_i(\varepsilon_{ij}) = \lambda\ \exp(-\exp(-\lambda\varepsilon_{ij} - \gamma))\ \exp(-\lambda\varepsilon_{ij} - \gamma)$, we can compute the expected value of the random shock of player i as follows

$$E(\mathcal{E}_{ij}) = \int_{-\infty}^{+\infty} \varepsilon_{ij}\ \lambda\ \exp(-\exp(-\lambda\varepsilon_{ij} - \gamma))\ \exp(-\lambda\varepsilon_{ij} - \gamma)\ d\varepsilon_{ij}$$

Using the substitution $z = \exp(-\lambda\varepsilon_{ij} - \gamma)$, we obtain

$$\begin{aligned} E(\mathcal{E}_{ij}) &= -\frac{1}{\lambda}\int_0^{+\infty} (ln(z) + \gamma)\ \exp(-z)\ dz \\ &= \frac{1}{\lambda}\left\{ -\int_0^{+\infty} ln(z)\exp(-z)\ dz - \gamma\int_0^{+\infty} \exp(-z)\ dz \right\} \end{aligned}$$

where $-\int_0^{+\infty} ln(z)\exp(-z)\ dz = \gamma \approx 0.5772156$ *is the Euler-Mascheroni constant*

such that we finally get $E(\mathcal{E}_{ij}) = \frac{1}{\lambda}\left\{\gamma - \gamma\right\} = 0$ *and so* $E(\hat{u}_{ij}) = \bar{u}_{ij} + E(\mathcal{E}_{ij}) = \bar{u}_{ij}$.

With regard to the variance, we have the following integral

$$\begin{aligned} Var(\mathcal{E}_{ij}) &= \int_{-\infty}^{+\infty} \varepsilon_{ij}^2\ \lambda\ \exp(-exp(-\lambda\varepsilon_{ij} - \gamma))\ \exp(-\lambda\varepsilon_{ij} - \gamma)\ d\varepsilon_{ij} \\ &= \int_0^{+\infty} \frac{1}{\lambda^2}\left((ln(z))^2 + \gamma^2 + 2\ ln(z)\gamma\right)\ \exp(-z)\ dz \end{aligned}$$

$$= \frac{1}{\lambda^2}\left\{ \int_0^{+\infty} (ln(z))^2 \exp(-z)\, dz + \gamma^2 \int_0^{+\infty} \exp(-z)\, dz + 2\,\gamma \int_0^{+\infty} \exp(-z) ln(z)\, dz \right\}$$

$$= \frac{1}{\lambda^2}\left\{ \gamma^2 + \frac{\pi^2}{6} \;+\; \gamma^2 - 2\,\gamma^2 \right\} = \frac{\pi^2}{6}\lambda^{-2} \quad \Box$$

References

[1] J. S. Mill, *Essays on some unsettled questions of political economy.* JW Parker, 1874.

[2] J. Persky, "Retrospectives: the ethology of homo economicus," *The Journal of Economic Perspectives*, pp. 221–231, 1995.

[3] S. Adam, "An inquiry into the nature and causes of the wealth of nations," *Edwin Cannan's annotated edition*, 1776.

[4] R. J. Herrnstein, "Rational choice theory: Necessary but not sufficient.," *American Psychologist*, vol. 45, no. 3, pp. 356–367, 1990.

[5] H. A. Simon, "From substantive to procedural rationality," in *25 Years of Economic Theory*, pp. 65–86, Springer, 1976.

[6] K. J. Arrow and G. Debreu, "Existence of an equilibrium for a competitive economy," *Econometrica: Journal of the Econometric Society*, pp. 265–290, 1954.

[7] G. Debreu, *Theory of Value.* New York: Wiley, 1959.

[8] J. Savage Leonard, *The foundations of statistics.* New York: Wiley, 1954.

[9] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition).* Princeton university press, 2007.

[10] J. F. Nash *et al.*, "Equilibrium points in n-person games," *Proceedings of the national academy of sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[11] M. Allais, "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine," *Econometrica: Journal of the Econometric Society*, pp. 503–546, 1953.

[12] H. A. Simon, "Rationality in psychology and economics," *The Journal of Business*, pp. S209–S224, 1986.

[13] H. A. Simon, "A behavioral model of rational choice," *The quarterly journal of economics*, pp. 99–118, 1955.

[14] H. A. Simon, "Rational choice and the structure of the environment.," *Psychological review*, vol. 63, no. 2, pp. 129–138, 1956.

[15] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *science*, vol. 185, no. 4157, pp. 1124–1131, 1974.

[16] A. Tversky and D. Kahneman, "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment.," *Psychological review*, vol. 90, no. 4, pp. 293–315, 1983.

[17] A. Tversky and D. Kahneman, "Rational choice and the framing of decisions," *Journal of business*, pp. S251–S278, 1986.

[18] C. Camerer, *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press, 2003.

[19] O. E. Williamson, "The economics of organization: The transaction cost approach," *American journal of sociology*, pp. 548–577, 1981.

[20] R. W. Rosenthal, "A bounded-rationality approach to the study of noncooperative games," *International Journal of Game Theory*, vol. 18, no. 3, pp. 273–292, 1989.

[21] A. Beja, "Imperfect equilibrium," *Games and Economic Behavior*, vol. 4, no. 1, pp. 18–36, 1992.

[22] K.-Y. Chen, "The strategic behavior of rational novices," *Ph.D. dissertation California Institute of Technology*, 1994.

[23] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and economic behavior*, vol. 10, no. 1, pp. 6–38, 1995.

[24] B. W. Rogers, T. R. Palfrey, and C. F. Camerer, "Heterogeneous quantal response equilibrium and cognitive hierarchies," *Journal of Economic Theory*, vol. 144, no. 4, pp. 1440–1467, 2009.

[25] C. F. Camerer, T.-H. Ho, and J.-K. Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, pp. 861–898, 2004.

[26] S. Choi, D. Gale, and S. Kariv, "Social learning in networks: a quantal response equilibrium analysis of experimental data," *Review of Economic Design*, vol. 16, no. 2-3, pp. 135–157, 2012.

[27] B. Lieberman, "Human behavior in a strictly determined $3\times$ 3 matrix game," *Behavioral Science*, vol. 5, no. 4, pp. 317–322, 1960.

[28] B. O'Neill, "Nonmetric test of the minimax theory of two-person zerosum games," *Proceedings of the National Academy of Sciences*, vol. 84, no. 7, pp. 2106–2109, 1987.

[29] A. Rapoport and R. B. Boebel, "Mixed strategies in strictly competitive games: a further test of the minimax hypothesis," *Games and Economic Behavior*, vol. 4, no. 2, pp. 261–283, 1992.

[30] J. Ochs, "Games with unique, mixed strategy equilibria: An experimental study," *Games and Economic Behavior*, vol. 10, no. 1, pp. 202–217, 1995.

[31] R. D. Luce, *Individual Choice Behavior a Theoretical Analysis.* John Wiley and Sons, 1959.

[32] S. P. Anderson, J. K. Goeree, and C. A. Holt, "Rent seeking with bounded rationality: An analysis of the all-pay auction," *Journal of Political Economy*, vol. 106, no. 4, pp. 828–853, 1998.

[33] J. K. Goeree and C. A. Holt, "Asymmetric inequality aversion and noisy behavior in alternating-offer bargaining games," *European Economic Review*, vol. 44, no. 4, pp. 1079–1089, 2000.

[34] G. López-Acevedo, "Quantal response equilibria for posted offer-markets," *Estudios Económicos*, pp. 95–131, 1997.

[35] S. P. Anderson, J. K. Goeree, and C. A. Holt, "Minimum-effort coordination games: Stochastic potential and logit equilibrium," *Games and Economic Behavior*, vol. 34, no. 2, pp. 177–199, 2001.

[36] S. Moinas and S. Pouget, "Rational and irrational bubbles: an experiment," *TSE Working Paper*, vol. 9, 2009.

[37] D. McFadden, "Conditional logit analysis of qualitative choice behavior," *Zarembka P., Frontiers in Econometrics*, pp. 105–142, 1974.

[38] J. K. Goeree, C. A. Holt, and T. R. Palfrey, "Regular quantal response equilibrium," *Experimental Economics*, vol. 8, no. 4, pp. 347–367, 2005.

[39] P. A. Haile, A. Hortaçsu, and G. Kosenok, "On the empirical content of quantal response equilibrium," *The American Economic Review*, vol. 98, no. 1, pp. 180–200, 2008.

[40] J. Marschak, "Binary-choice constraints and random utility indicators," in *Proceedings of a Symposium on Mathematical Methods in the Social Sciences*, vol. 7, pp. 19–38, 1960.

[41] R. D. Luce and P. Suppes, *Preference, utility, and subjective probability.* Wiley, 1965.

[42] R. Woodworth and G. T. Ladd, "Elements of physiological psychology," *Scribners: New York*, 1911.

[43] R. D. McKelvey and T. R. Palfrey, "An experimental study of the centipede game," *Econometrica: Journal of the Econometric Society*, pp. 803–836, 1992.

[44] C.-L. Zhao and H.-J. Huang, "Modeling bounded rationality in congestion games with the quantal response equilibrium," *Procedia-Social and Behavioral Sciences*, vol. 138, pp. 641–648, 2014.

[45] E. Culler and E. Girden, "The learning curve in relation to other psychometric functions," *The American journal of psychology*, pp. 327–349, 1951.