Quantal Response Equilibria with Heterogeneous Agents

Russell Golman[a*]

November 1, 2008

[a] Department of Applied and Interdisciplinary Mathematics, University of Michigan, 2082 East Hall, 530 Church Street, Ann Arbor, MI 48109, USA. Telephone: 734 764-7436. Fax: 734 763-0937. E-mail: rgolman@umich.edu

*I am grateful to Andreas Blass and Scott Page for many helpful conversations.

Abstract

We examine the use of single-agent and representative-agent models to describe the aggregate behavior of heterogeneous quantal responders. We consider heterogeneous quantal response functions arising from a distribution of distributions of payoff shocks. A representative agent would have the average quantal response function. Weakening a standard assumption about the admissible distributions of payoff shocks, we show existence of a representative agent. However, this representative agent does not have a representative distribution of payoff shocks, nor any iid distribution in large enough games. We consider a specific case of heterogeneous logit responders and find that a mis-specified homogenous logit parameter will have downward bias.

KEYWORDS: quantal response equilibria, bounded rationality, representative agent, heterogeneity, logit response

JEL classification codes: C72, C02

1 Introduction

Quantal response equilibrium extends the Nash Equilibrium notion to allow bounded rationality. Players can be seen as making errors while trying to choose optimal strategies, or equivalently, as observing payoffs disturbed by idiosyncratic noise. The result is that players may select any action with positive probability assigned by their quantal response functions.

This paper introduces a general model of quantal response equilibrium with heterogeneous agents. We show that the aggregate behavior of a population of heterogeneous agents can be captured by a representative agent. But, the representative agent may be very different than the actual agents in the population. This illustrates the need to consider heterogeneity and offers insight for how to work around that heterogeneity with representative-agent models. After presenting the representative-agent picture, which allows for arbitrary distributions of payoff noise and applies for all normal form games, we then consider logit responses in the context of a single choice between two pure strategies that is part of a fixed game. We find that in a heterogeneous population of agents, all having their own logit rationality parameters, a mis-specified homogenous logit parameter will always exhibit a downward bias making the population appear to be less rational.

We consider structural quantal response equilibria (QRE) [21, 11] in the context of a population game. In a large population of agents, we should expect heterogeneity of behavior [18, 20]. A population of quantal responders should consist of agents who may have different error rates, or different distributions of payoff noise. In fact, McKelvey, et. al. [22] find experimental evidence for heterogeneous error distributions in trying to fit logit QRE to data on two-by-two asymmetric games. [1]

Prior research into quantal response equilibria with heterogeneous agents has considered a distribution of parameters which parametrize the distributions of payoff noise [24], with particular interest in distributions of logit responders [4]. Here, we model heterogeneous distributions of payoff noise with a functional defined over distribution functions. As we do not assume that distributions of payoff noise take any particular functional forms, this approach allows for more distribution functions than can be described with finitely many parameters.

Our interest is in the behavior of an entire population, and we seek a representative agent whose mixed strategy quantal response always matches the population aggregate. We need representative-agent models because while we believe people really are heterogeneous, we cannot determine each person's quantal response function individually when we fit data. The representative agent is what we can estimate in an experiment.

With weak assumptions on the agents' distributions of payoff noise we prove existence of a representative agent. However, the distribution of payoff disturbances necessary to produce representative choices is not representative of the noise the actual

[1] Further motivation to consider heterogeneity in a population of quantal responders comes from recent findings that models of heterogeneous learners often cannot be adequately approximated by representative-agent models with common parameter values for all [26, 15, 12].

agents observe in their payoffs. We show that in games with enough pure strategies, a representative agent could not have payoff disturbances independent and identically distributed across actions even if the actual agents did. On the other hand, we find that if agents all use regular quantal response functions (as defined by Goeree, et. al. [11]), then the representative agent's quantal response must also be regular. Different roles in asymmetric games will in general have different representative agents.

Much of the QRE literature looks to the logit equilibrium in particular to explain experimental data [5, 8, 10, 6, 1]. Because of the prominence of this logit response specification, we consider a population of heterogeneous logit responders as a special case. Our interest here is how a mis-specified homogenous logit model misrepresents the heterogeneous agents. Because the representative agent for the population is not itself a logit responder, the homogenous model cannot explain equilibrium choice probabilities and payoffs in a choice between more than two actions. When the population has just two pure strategies, we find that the homogenous logit parameter is systematically biased below the average value of the heterogeneous logit parameters. We describe the extent of this bias as it varies with the difference in the two strategies' equilibrium payoffs.

Nash proposed a population game interpretation of equilibrium in his unpublished PhD dissertation [25]. Following his lead, we assume that there is a population of agents for each position of a game. A generic n-player game involves n populations of agents, but if multiple players have identical roles and we adopt the restriction that players in identical roles should play identical population mixed strategies, then these players may be selected from the same population. So, in a totally symmetric game, we may have only a single population of agents. We assume the populations are large, and we are interested in the fraction of a population playing a given strategy. An agent's payoff is the average of his payoffs against all other combinations of agents (or equivalently his expected payoff given random matching).

Population games provide a framework for the use of evolutionary learning dynamics. Learning rules that assume that players noisily best respond often converge to QRE [7, 17, 19, 16, 2]. This paper focuses on the QRE itself and not on any particular learning rule that might lead to it. Population games also describe experimental settings well, as data is accumulated through the randomly matched interactions of many subjects.

This paper is organized as follows. Section 2 introduces the notation in the context of a single population and provides definitions of a QRE and a representative agent. Section 3 contains our general results describing a representative agent. In Section 4, we extend our framework and our results to n-player asymmetric games. Section 5 focuses on logit responders, and section 6 concludes. The Appendix contains proofs omitted from the text.

2 A Single Population

To simplify the presentation, we begin with a single population of agents. The context can be thought of as a symmetric game or alternatively a single player decision subject to incomplete information. In Section 4, we show how to apply these results to general n-player asymmetric games.

Let $S = \{s_1, \ldots, s_J\}$ be the set of pure strategies available to the agents. The collective play of all the agents defines the population mixed strategy x. Formally, $x \in \triangle^{J-1}$, the $(J-1)$-dimensional simplex where $x_j \geq 0$ for all j and $\sum_j x_j = 1$.

A structural QRE arises when agents' utility functions are modified by noise terms, privately observed stochastic payoff disturbances. Denote by π_j the payoff from taking pure strategy s_j. Of course, payoffs are a function of the strategies used by all the players, $\pi_j = \pi_j(x)$, but we omit the function's argument for ease of notation. Also denote the vector $\pi = \pi_1, \ldots, \pi_J$. Formally, $\pi : \triangle^{J-1} \rightarrow \Re^J$. For each pure strategy s_j, agent μ observes a payoff disturbance ϵ_j^μ, making agent μ's disturbed payoff $\pi_j^\mu = \pi_j + \epsilon_j^\mu$. This is the function agents maximize with their choice of strategy in a QRE.

The distribution of payoff disturbances is assumed to be *admissible*, meaning that:

(a1) the disturbances are independent across agents;

(a2) each agent has an absolutely continuous joint distribution of $(\epsilon_1^\mu, \ldots, \epsilon_J^\mu)$, i.e., all marginal densities exist;

(a3) disturbances are unbiased in the sense that they all have mean zero.

Allowing only admissible distributions guarantees the existence of a QRE. Here, we make the additional assumption that for each agent, disturbances are independent and identically distributed (iid) across the set of actions. This assumption could be relaxed, but some such restriction is necessary for the QRE notion to produce falsifiable predictions [14].

When the setup for QRE does not explicitly involve populations of agents, it is assumed that each player has a distribution of payoff disturbances. In the context of a population game, this corresponds to each agent within the population having an identical distribution of disturbances. That is, the convention is to assume homogenous populations. Here, we specifically want to leave open the possibility that agents in the same population have different distributions of payoff shocks. So, we do *not* assume identical distributions of ϵ_j^μ for all μ.

To model heterogeneity in the distributions of payoff disturbances, consider a functional defined over such distributions. Let $P_\epsilon(\cdot)$ be a distribution function for the payoff disturbance to a particular action. Each agent has a distinct P_ϵ, which then applies to ϵ_j^μ for all $1 \leq j \leq J$, i.e., is the same for all actions in that agent's strategy space. Define a functional $F_\epsilon[P_\epsilon]$ that associates to each distribution function P_ϵ a probability mass or density describing the fraction of the population with payoff disturbances distributed by P_ϵ. Technically, we make use of a second functional $I_\epsilon[P_\epsilon]$ that equals 1 to indicate a mass point on P_ϵ and 0 to indicate that $F_\epsilon[P_\epsilon]$ represents a probability density. For this to make sense we require $I_\epsilon[P_\epsilon] = 1$ for only countably many P_ϵ and

$$\sum_{P_\epsilon : I_\epsilon[P_\epsilon]=1} F_\epsilon[P_\epsilon] + \int_{P_\epsilon : I_\epsilon[P_\epsilon]=0} F_\epsilon[P_\epsilon]\, dP_\epsilon = 1.$$

The appropriate measure dP_ϵ depends on the particular form of the heterogeneity.

In this approach, the functional captures a distribution of distributions of payoff shocks in the population. It thus provides a general way to think about heterogeneity of quantal responses. The conventional assumption of a homogenous population can be recaptured, for example, by taking $F_\epsilon[P_\epsilon] = 1$ for a particular P_ϵ and 0 everywhere else.

The quantal response function for each agent returns the agent's likelihood of choosing each strategy given the agent's undisturbed payoffs. Let $Q^\mu_j(\pi)$ be the probability that agent μ selects strategy s_j given the payoffs to each strategy. Formally, for any vector $\pi' = (\pi'_1, \ldots, \pi'_J) \in \Re^J$, define

$$R^\mu_j(\pi') = \{(\varepsilon^\mu_1, \ldots, \varepsilon^\mu_J) \in \Re^J : \pi'_j + \varepsilon^\mu_j \geq \pi'_{j'} + \varepsilon^\mu_{j'} \text{ for all } j' = 1, \ldots, J\}$$

to be the set of realizations of agent μ's joint set of payoff disturbances that would lead to choosing action s_j. Then $Q^\mu_j(\pi) = \text{Prob}\left\{(\epsilon^\mu_1, \ldots, \epsilon^\mu_J) \in R^\mu_j(\pi)\right\}$.

The quantal response functions for all the agents can be aggregated across the population to give the population mixed strategy response to agents' choices at any given population mixed strategy. In a finite population of m agents, the population aggregate quantal response is $Q_j = \frac{1}{m}\sum_{\mu=1}^m Q^\mu_j$ for all j. More generally, the aggregate quantal response in an infinite population is

$$Q_j = \sum_{P_\epsilon : I_\epsilon[P_\epsilon]=1} F_\epsilon[P_\epsilon]\, Q^\mu_j + \int_{P_\epsilon : I_\epsilon[P_\epsilon]=0} F_\epsilon[P_\epsilon]\, Q^\mu_j \, dP_\epsilon \tag{1}$$

where we abuse notation by letting $\mu = \mu(P_\epsilon)$ be an agent with payoff disturbances iid from P_ϵ.

We can now define a quantal response equilibrium and then formally describe a representative agent for this heterogeneous population.

Definition A *quantal response equilibrium* is defined by the fixed point equation $x_j = Q_j(\pi(x))$ for all j.

Definition A *representative agent* has a quantal response function $\hat{Q}(\pi)$ equal to the population aggregate quantal response function:

$$\hat{Q} = (Q_1, \ldots, Q_J). \tag{2}$$

For all games, the population as a whole behaves exactly as if it were homogenously composed of representative agents.

Strategy choices are determined by the differences between disturbed payoffs, so the most relevant variables are the differences between payoff shocks, $\delta^{\mu}_{jj'} = \epsilon^{\mu}_j - \epsilon^{\mu}_{j'}$. These $\delta^{\mu}_{jj'}$ are identically distributed across all $j, j' \neq j$ because the ϵ^{μ}_j are iid across all j. By absolute continuity, the marginal densities exist, and by symmetry, they are even functions. But, there is obviously dependence among these random variables across j and j'. We will consider the $(J-1)$-dimensional random vector $\delta^{\mu}_j = \left(\delta^{\mu}_{1j}, \ldots, \widehat{\delta^{\mu}_{jj}}, \ldots, \delta^{\mu}_{Jj}\right)$ for a particular j, which then determines the value of $\delta^{\mu}_{j'}$ for all other j'. Note that admissibility of the payoff disturbances implies δ^{μ}_j has zero mean because all the ϵ^{μ}_j have zero mean. Let $P_{\delta^{\mu}_{j'}} : \Re^{J-1} \rightarrow [0,1]$ be the joint distribution function of $\delta^{\mu}_{j'}$. Then

$$Q^{\mu}_j(\pi) = P_{\delta^{\mu}_{j'}}(\pi_j - \pi_1, \pi_j - \pi_2, \ldots, \pi_j - \pi_J), \tag{3}$$

naturally omitting $\pi_j - \pi_j$ just as we did δ^{μ}_{jj}.

Heterogeneity in the distributions of payoff shocks leads to heterogeneity in the distributions of the differences between payoff shocks. That is, the functional over P_ϵ induces a functional defined on the joint distributions for δ^{μ}_{j}. Let $P(\cdot)$ be a joint distribution function of δ^{μ}_{j} for some μ and any j in $1, \ldots, J$. Let $\mathcal{P}$ be the set of $P_\epsilon(\cdot)$, distribution functions of ϵ^{μ}_{j}, which give rise to $P(\cdot)$. If there exists a $P_\epsilon \in \mathcal{P}$ such that $I_\epsilon[P_\epsilon] = 1$, then define $I[P] = 1$ and

$$F[P] = \sum_{\{P_\epsilon : I_\epsilon[P_\epsilon]=1, P_\epsilon \in \mathcal{P}\}} F_\epsilon[P_\epsilon].$$

Otherwise, define $I[P] = 0$ and

$$F[P] = \int_{P_\epsilon \in \mathcal{P}} F_\epsilon[P_\epsilon] \, dP_\epsilon.$$

Note that there are joint distribution functions that could not apply to any δ^{μ}_{j} because they do not describe differences between iid random variables, and our definition implies $F = 0$ for these functions.

In an abuse of notation, we will use $\int dP$ as a shorthand for

$$\sum_{P:I[P]=1} + \int_{P:I[P]=0} dP.$$

This notation will be used even when there are mass points and is not meant to suggest their exclusion. It merely reflects our desire not to worry about the particular form the heterogeneity takes.

Our definition of a representative agent can now be translated into a statement about the representative joint distribution of differences between payoff shocks. It means the representative agent has δ_j distributed according to a joint distribution function $\hat{P}(\cdot)$ such that

$$\hat{P} = \int F[P]\, P\, dP. \tag{4}$$

The representative agent's quantal response function can then be found by using $\hat{P}$ in Equation (3). This provides a working definition of a representative agent that is more useful than Equation (2).

Given a functional that describes the heterogeneity of the population, we can use characteristic functions to identify a representative agent. This approach is effective because there is a bijection between distribution functions and characteristic functions. Let $\theta : \Re \to \mathbb{C}$ be the characteristic function of a payoff disturbance ϵ_j with distribution function $P_\epsilon(\cdot)$,

$$\theta(t) = E(e^{it\epsilon_j}).$$

Note that θ is a complex-valued function of a single real variable and $\theta(-t) = \bar{\theta}(t)$. It must be uniformly continuous and non-negative definite and satisfy $\theta(0) = 1$, $|\theta(t)| \leq 1$, properties which can be used to define an arbitrary characteristic function. Take $\phi : \Re^{J-1} \to \mathbb{C}$ to be the characteristic function associated with the joint distribution $P(\cdot)$ of δ_j. We still write $\phi(t)$, now assuming $t = (t_1, \ldots, t_{J-1})$ to be a vector in $\Re^{J-1}$.

We can express ϕ in terms of θ,

$$\begin{aligned}\phi(t) &= E(e^{it\cdot\delta_j})\\ &= E(e^{it_1\delta_{1j}}\cdots e^{it_{J-1}\delta_{Jj}})\\ &= E(e^{it_1(\epsilon_1-\epsilon_j)}\cdots e^{it_{J-1}(\epsilon_J-\epsilon_j)})\\ &= E(e^{it_1\epsilon_1})\cdots E(e^{it_{J-1}\epsilon_J})\cdot E(e^{-i(\sum_{l=1}^{J-1}t_l)\epsilon_j})\\ &= \theta(t_1)\cdots\theta(t_{J-1})\cdot\theta(-\sum_{l=1}^{J-1}t_l). \qquad (5)\end{aligned}$$

In addition to the properties just mentioned, we also know that if $\sum_{l=1}^{J} r_l = 0$, then $\phi(r_1,\ldots,\widehat{r_j},\ldots,r_J)$ is independent of j, because by Equation (5) it has the same expansion in terms of θ for all j. If there are only two actions, $J = 2$, then ϕ is real and positive because P is symmetric. The functional F_ϵ induces a distribution over characteristic functions $\Psi_\epsilon[\theta] = F_\epsilon[P_\epsilon]$, with $\Upsilon_\epsilon[\theta] = I_\epsilon[P_\epsilon]$. Similarly, define $\Psi[\phi] = F[P]$ along with $\Upsilon[\phi] = I[P]$.

Let

$$\hat{\phi}(t) = \int \Psi[\phi]\,\phi(t)\,d\phi. \qquad (6)$$

This representative characteristic function can be constructed by taking the integral pointwise, i.e., independently for every value of t. Fixing the input point t, we know that the functional integral $\int \Psi[\phi]\,\phi(t)\,d\phi$ always converges because $|\phi(t)| \le 1$. (The abuse of notation in this context is the same as for the functional over distribution functions.)

3 A Representative Agent

The first issue to address is whether a representative agent exists. Theorem 1 tells us that there is only one pathological type of heterogeneity for which the population does not have a representative agent. The joint distribution function $\hat{P}(\cdot)$ can be constructed given the functional $F[P]$ describing the heterogeneity in the population, but there is a danger that it is not an admissible distribution function. Specifically, it may fail to have finite mean. A particular consequence of the theorem is the fact that a representative agent is sure to exist whenever only finitely many different distribution functions are in use in the population. Alternatively, relaxing the requirement that distributions of disturbances have zero mean also ensures the existence of a representative agent.

Theorem 1 *Define $\hat{P}(\cdot)$ as in Equation (4). If $\hat{P}(\cdot)$ has finite mean, then a representative agent exists with δ_j distributed by $\hat{P}(\cdot)$ and having characteristic function $\hat{\phi}(t)$.*

Proof It is well known that the maps between distribution functions and characteristic functions are linear. Apply the Levy continuity theorem to Equation (6). This requires $\hat{\phi}(t)$ to be continuous at $t = 0$, which we establish with Lemma 1 in the Appendix. Lemma 2 in the Appendix establishes that the mean of $\hat{P}(\cdot)$ is 0 if it exists, and thus $\hat{P}(\cdot)$ is admissible when this is the case. ∎

Corollary 1 *If $F[P] > 0$ for only finitely many joint distribution functions $P(\cdot)$, then a representative agent exists.*

Proof The only source for divergence of the mean of $\hat{P}(\cdot)$ is the limit that results from $F[P] > 0$ for infinitely many P. All the joint distribution functions in the support of F have zero mean, so a finite linear combination of them also describes a random vector with zero mean. Then Theorem 1 applies. ∎

Taking a closer look at an example $\hat{P}(\cdot)$ that has divergent mean and thus fails to be an admissible joint distribution function offers insight into how such cases arise. For simplicity, assume $J = 2$. The example works just as well with more pure strategies, but the notation becomes cluttered. Partition the set of joint distribution functions $P(\cdot)$ into $\mathcal{P}_y$ such that $P(\cdot) \in \mathcal{P}_y$ implies $P(e^y) \leq 1 - \alpha$ for some fixed positive $\alpha < \frac{1}{2}$. This partition is not uniquely determined, but as long as the $\mathcal{P}_y$ are non-empty, it will do. Consider the functional $F[P]$ where

$$\int_{\mathcal{P}_y} F[P]\, dP = \begin{cases} e^{-y} \text{ for } y \geq 0 \\ \\ 0 \text{ for } y < 0. \end{cases}$$

Then the mean of $\hat{P}(\cdot)$ is divergent because

$$\begin{aligned}
\int_0^\infty \delta\, d\hat{P}(\delta) &= \int_0^\infty \delta \int_0^\infty \int_{\mathcal{P}_y} F[P] P'(\delta)\, dP\, dy\, d\delta \\
&\geq \int_0^\infty \int_{\mathcal{P}_y} F[P] \int_{e^y}^\infty \delta\, P'(\delta)\, d\delta\, dP\, dy \\
&\geq \int_0^\infty \int_{\mathcal{P}_y} F[P] \alpha e^y\, dP\, dy \\
&= \int_0^\infty \alpha\, dy.
\end{aligned}$$

Admissibility requires $\hat{P}(\cdot)$ to have zero mean, but when this fails, we shouldn't conclude that a representative quantal response function does not exist. Instead, we can relax the requirements of admissibility to guarantee that a representative agent always exists. The restriction to zero mean payoff disturbances is not necessary for the existence of a QRE, as fixed point theorems can be applied without it. The desire for unbiased disturbances appears to be aesthetic, and the possible inadmissibility of representative agents is an artifact of the way it is implemented. Consider replacing the zero mean assumption (a3) with the following alternative:

(a3') the Cauchy principal value of the mean of each payoff disturbance is zero[2], and

$$\lim_{\varepsilon \to \infty} \varepsilon \operatorname{Prob}\left\{|\epsilon_j^\mu| \geq \varepsilon\right\} = 0 \text{ for each } \epsilon_j^\mu.$$

Assumption (a3') holds whenever assumption (a3) is satisfied, so this is a weaker condition to impose on the payoff disturbances. Even though the mean of ϵ_j^μ may blow up under assumption (a3'), these disturbances are still unbiased, and their likelihood still decays sufficiently quickly as they get large.

Definition We say payoff disturbances are *weakly admissible* if assumptions (a1), (a2) and (a3') hold.

With just this slight relaxation of admissibility, we always get a representative agent.

Corollary 2 *Allow weakly admissible payoff disturbances. A representative agent exists with δ_j distributed by $\hat{P}(\cdot)$ and having characteristic function $\hat{\phi}(t)$.*

Proof Lemma 2 shows that $\hat{P}(\cdot)$ always satisfies the weak admissibility assumption

[2]The Cauchy principal value of an improper integral $\int_{-\infty}^{\infty} f(t)\,dt$ is defined as $\lim_{T\to\infty} \int_{-T}^{T} f(t)\,dt$.

(a3'). In turn, there exists a joint distribution of $(\epsilon_1, \dots, \epsilon_J)$ that satisfies (a3') and is consistent with δ_j being distributed by $\hat{P}(\cdot)$. ∎

We have defined a representative agent with the property that the agent's choice of strategy is representative of the population as a whole. We now show that this is not equivalent to having representative noise in the underlying payoffs. We say $\hat{P}_\epsilon(\cdot)$ is a *representative distribution of payoff shocks* if it is a (weakly) admissible distribution function and

$$\hat{P}_\epsilon = \int F_\epsilon[P_\epsilon]\, P_\epsilon \, dP_\epsilon. \tag{7}$$

By applying the Levy continuity theorem here too, we find that a representative distribution of payoff shocks has characteristic function $\hat{\theta}(t) = \int \Psi_\epsilon[\theta]\, \theta(t)\, d\theta$. With this groundwork in place, we are ready for Theorem 2, which says that a representative quantal response function does not arise from a representative distribution of payoff shocks.

Theorem 2 *A representative agent has a representative distribution of payoff shocks if and only if the population is homogenous.*

Proof Let Θ be the set of characteristic functions of ϵ_j that give rise to a given $\phi(\cdot)$. Using Equation (5), $\Theta = \{\theta : \phi(t) = \left(\prod_{l=1}^{J-1} \theta(t_l)\right) \cdot \theta(-\sum_{l=1}^{J-1} t_l)\}$. From the relationships between the functionals, we have

$$\begin{aligned} \Psi[\phi] = & \int_\Theta \Psi_\epsilon[\theta]\, d\theta & \text{if } \Upsilon[\phi] = 0 \\ & \sum_{\{\theta \in \Theta : \Upsilon_\epsilon[\theta] = 1\}} \Psi_\epsilon[\theta] & \text{if } \Upsilon[\phi] = 1. \end{aligned}$$

We can then express a representative agent's characteristic function for δ_j as

$$\hat{\phi}(t) = \int \Psi_\epsilon[\theta] \left(\prod_{l=1}^{J-1} \theta(t_l) \right) \theta(-\sum_{l=1}^{J-1} t_l) \, d\theta.$$

But

$$\int \Psi_\epsilon[\theta] \left(\prod_{l=1}^{J-1} \theta(t_l) \right) \theta(-\sum_{l=1}^{J-1} t_l) \, d\theta \neq \left(\prod_{l=1}^{J-1} \int \Psi_\epsilon[\theta] \, \theta(t_l) \, d\theta \right) \cdot \int \Psi_\epsilon[\theta] \, \theta(-\sum_{l=1}^{J-1} t_l) \, d\theta \tag{8}$$

unless for each t_l, $\theta(t_l)$ is the same for all θ in the support of Ψ_ϵ. Since t_l is an arbitrary variable, this would mean there could only be one function in the support of Ψ_ϵ, i.e., no heterogeneity of distributions of payoff shocks in the population. ∎

In light of the fact that a representative agent for a heterogeneous population does not have a representative distribution of payoff shocks, the question arises as to what distribution of payoff shocks could actually produce a representative agent. According to the next result, if there are enough actions and there is heterogeneity of the $\delta_{jj'}$, then the representative agent cannot arise from any distribution of payoff shocks that is iid across the set of actions. Theorem 3 says that if there are just two actions, there is an iid distribution of payoff shocks (possibly many such distributions) that generates the representative agent. But, if there are at least four actions, assuming heterogeneity of the $\delta_{jj'}$, it is impossible for an iid distribution of payoff shocks to generate the representative agent.[3]

[3]Examples indicate that when there are three actions, the representative agent usually cannot arise from iid shocks, but we cannot rule out special cases of heterogeneity for which the representative agent is compatible with iid disturbances.

Theorem 3 *Given a representative agent, if $J = 2$, there exists a distribution of payoff shocks iid across all actions and each with characteristic function $\hat{\vartheta}(\cdot)$ such that*

$$\hat{\phi}(t) = \left(\prod_{l=1}^{J-1} \hat{\vartheta}(t_l)\right) \cdot \hat{\vartheta}(-\sum_{l=1}^{J-1} t_l). \tag{9}$$

But, when $J \geq 4$, there is no $\hat{\vartheta}(\cdot)$ that satisfies Equation (9) unless every $P_\epsilon(\cdot)$ in the support of F_ϵ gives the same distribution of the $\delta_{jj'}$.

Proof When $J = 2$, we must find a $\hat{\vartheta}(\cdot)$ such that $\hat{\phi}(t_1) = \hat{\vartheta}(t_1) \cdot \hat{\vartheta}(-t_1)$. Recall $J = 2$ implies that all $\phi(\cdot)$ are real and positive, and hence so is $\hat{\phi}$. It suffices to take $\hat{\vartheta}(t_1) = \hat{\vartheta}(-t_1) = \sqrt{\hat{\phi}(t_1)}$.

Now consider $J \geq 4$. Given that individual agents do have payoff shocks that are iid across all actions, any $\phi(\cdot)$ in the population can be expressed in terms of $\theta(\cdot)$ with Equation (5). Specifically, $\phi(a, -a, a, 0, \cdots, 0) = (\theta(a)\theta(-a))^2$. Similarly, $\phi(a, 0, \cdots, 0) = \theta(a)\theta(-a)$. Thus,

$$\phi(a, -a, a, 0, \cdots, 0) = (\phi(a, 0, \cdots, 0))^2 .$$

But

$$\int \Psi[\phi]\, \phi(a, -a, a, 0, \cdots, 0)\, d\phi \neq \left(\int \Psi[\phi]\, \phi(a, 0, \cdots, 0)\, d\phi\right)^2$$

unless there is no variance of $\theta(a)\theta(-a)$ in the population. Note that $\delta_{jj'}$ has characteristic function $\theta(t)\theta(-t)$. Thus, if there are two distribution functions in the support of $F_\epsilon[P_\epsilon]$ that give different distributions of $\delta_{jj'}$, then for some a,

$$\hat{\phi}(a, -a, a, 0, \cdots, 0) \neq \left(\hat{\phi}(a, 0, \cdots, 0)\right)^2 .$$

This would mean $\hat{\phi}(\cdot)$ could not be expressed as $\left(\prod_{l=1}^{J-1} \hat{\vartheta}(t_l)\right) \cdot \hat{\vartheta}(-\sum_{l=1}^{J-1} t_l)$ for any $\hat{\vartheta}(\cdot)$. ∎

Theorem 3 sounds a cautionary note that even if we believe all agents have noise in their payoffs that is iid across their actions, heterogeneity of the agents leads the population as a whole to behave as if payoff disturbances were not iid across actions.

We desired agents with payoff noise iid across actions because this assumption imposes restrictions on behavior that can be tested empirically. Although it turns out the representative agent may not have payoff noise iid across actions, the representative agent notion still has empirical content because some properties are inherited from the underlying agents.

Goeree, et. al. [11] introduce four axioms which define a *regular quantal response function* $Q^{\mu} : \Re^J \to \triangle^{J-1}$ without reference to payoff noise:

(A1) *Interiority*: $Q_j^{\mu}(\pi) > 0$ for all $j = 1, \ldots, J$ and for all $\pi \in \Re^J$.

(A2) *Continuity*: $Q_j^{\mu}(\pi)$ is a continuous and differentiable function for all $\pi \in \Re^J$.

(A3) *Responsiveness*: $\frac{\partial Q_j^{\mu}(\pi)}{\partial \pi_j} > 0$ for all $j = 1, \ldots, J$ and for all $\pi \in \Re^J$.

(A4) *Monotonicity*: $\pi_j > \pi_{j'}$ implies $Q_j^{\mu}(\pi) > Q_{j'}^{\mu}(\pi)$, for all $j, j' = 1, \ldots, J$.

They argue that all quantal response functions obey Continuity and weakly obey Responsiveness. If the density of payoff disturbances has full support, then Interiority and Responsiveness are strictly satisfied. When payoff disturbances are iid across actions, then the quantal response function obeys Monotonicity as well.

We now show that any regularity property that holds for the underlying agents in the population also holds for the representative agent.

Theorem 4 *If a regularity axiom* $\{$*(A1), (A2), (A3), or (A4)*$\}$ *applies to* Q^{μ} *for all* μ *(i.e., for* $\mu = \mu(P)$ *whenever* $P(\cdot)$ *is in the support of* F*), then that axiom applies to the representative agent's quantal response function* $\hat{Q}$*.*

Proof Continuity holds for all quantal response functions as a result of the admissibility assumption (a2) that distributions of payoff noise must be absolutely continuous [11]. Interiority, Responsiveness, and Monotonicity each follow from Equations (1) and (2), which define a representative agent. Essentially, we just use the fact that an integral (or sum) must be positive if the integrand (summand) is always positive. For Responsiveness, we pass the partial derivative inside the integral and sum in Equation (1). For Monotonicity, we express $\hat{Q}_j(\pi) - \hat{Q}_{j'}(\pi)$ using Equation (1) and then pair up terms to form a single sum and integral. ∎

Theorem 4 tells us that in our framework, the representative agent's quantal response function always satisfies Monotonicity. It is this Monotonicity property that carries empirical content. In principle, subjects in an experiment could violate Monotonicity and choose actions with lower payoffs more often than actions with higher payoffs. This would be inconsistent with the predicted behavior of the representative agent.

4 Asymmetric Games

All of these results, initially presented in the context of a single population, apply to general asymmetric games. Consider a normal form game with n populations of agents. The strategy sets may differ across players, so we let $S_i = \{s_{i1}, \ldots, s_{iJ_i}\}$ be the set of pure strategies available to agents in population i. Now $x = x_1 \times \cdots \times x_n$ denotes the mixed strategy profile across all n populations in the game, with each mixed strategy vector $x_i = (x_{i1}, \ldots, x_{iJ_i}) \in \triangle^{J_i - 1}$.

The vector $\pi_i = \pi_{i1}, \ldots, \pi_{iJ_i}$ denotes the payoff to agents in population i from their available pure strategies. Now, $\pi_i : \triangle^{J_1 - 1} \times \cdots \times \triangle^{J_n - 1} \rightarrow \Re^{J_i}$. Agent μ observes payoff disturbances ϵ^{μ}_{ij} for each strategy. The variables $\delta^{\mu}_{ijj'} = \epsilon^{\mu}_{ij} - \epsilon^{\mu}_{ij'}$ and $\delta^{\mu}_{ij} = \left(\delta^{\mu}_{i1j}, \ldots, \widehat{\delta^{\mu}_{ijj}}, \ldots, \delta^{\mu}_{iJ_ij}\right)$ are similarly defined for each population. The quantal response functions $Q^{\mu}_{ij}(\pi_i)$ depend on the payoffs in population i, which in turn depend on the entire n-population mixed strategy profile x. Thus, Equation (3) applies for each i. The functionals also have to be indexed for the population so that $F^i_{\epsilon}[P_{\epsilon}]$ describes the fraction of population i with payoff disturbances distributed by P_{ϵ} and each F^i_{ϵ} induces a $F^i[P]$, $\Psi^i_{\epsilon}[\theta]$, and $\Psi^i[\phi]$. Equation (1) now applies for each i, making $(Q_{i1}, \ldots, Q_{iJ_i})$ the i^{th} population aggregate quantal response. The fixed point equation defining a quantal response equilibrium becomes $x_{ij} = Q_{ij}\left(\pi_i(x)\right)$ for all $i \in 1, \ldots, n$ and all $j \in 1, \ldots, J_i$.

Theorem 1 now describes the existence of a representative agent for population i with δ_{ij} distributed by the joint distribution function $\hat{P}^i(\cdot)$ and having characteristic function $\hat{\phi}^i(t)$. For each i, these representative functions are given by Equations (4) and (6), just as before. And Theorems 2, 3, and 4 apply to the representative agent from any given population.

However, while obtaining representative agents for each role in the game, we caution that there is no reason to assume the existence of a single representative agent the same for all players of an asymmetric game. Such an assumption would deny heterogeneity across the different roles of the game. And given the fact that a representative agent does not have a representative distribution of payoff shocks and that different players may have different strategy sets, it's not clear exactly what is meant by a single representative agent for all players of an asymmetric game. The problem is that we want to have a representative quantal response function $\hat{Q}_{ij}(\cdot)$ that is independent of i, for each fixed j, but this does not make sense when the set of actions j depends on i.

Here, we say that there is a single representative agent for all players if there exist $\hat{P}^i$ such that $J_{i_1} \leq J_{i_2}$ implies that for all $(\varpi_1, \ldots, \varpi_{J_{i_2}})$ and any $j \leq J_{i_1}$,

$$\hat{Q}_{i_1 j}(\varpi_1, \ldots, \varpi_{J_{i_1}}) = \lim_{\varpi_{J_{i_1}+1} \to -\infty} \cdots \lim_{\varpi_{J_{i_2}} \to -\infty} \hat{Q}_{i_2 j}(\varpi_1, \ldots, \varpi_{J_{i_2}}).$$

With this definition, we single out the representative agent from a population i that maximizes J_i. We can think of this agent with δ_{ij} distributed by $\hat{P}^i(\cdot)$ as representative of all players in all roles by assuming that when playing a role with too few possible actions, the agent imagines there are additional actions with infinitely negative payoffs. In the particular case that all players have the same number of actions,

$J_i = J$ for all i, a single representative agent for all players would have differences in payoff shocks jointly distributed by a $\hat{P}$ that satisfies $\hat{P} = \hat{P}^i$ for all i. The representative agent for each population would have to be the same. There are plenty of QRE which are incompatible with identical representative agents for all populations, as the following game illustrates.

Asymmetric Matching Pennies

	Left	Right
Up	9, -1	-1, 1
Down	-1, 1	1, -1

Goeree, et. al. [11] analyze the asymmetric matching pennies game shown above and find the set of possible QRE is the rectangle $\frac{1}{6} < p < \frac{1}{2}$, $\frac{1}{2} < q < 1$, where p is the probability the column player chooses left and q is the probability the row player chooses up. However, given the restriction that representative row and column players have the same quantal response function, a QRE must satisfy the additional constraint $q < 3p$ if and only if $q > 1 - p$. This is because $q < 3p$ means that $\pi_{2R} - \pi_{2L} < \pi_{1U} - \pi_{1D}$, i.e., the cost of an error is higher for the row player, and must lead to relatively fewer errors by the row player, $q > 1 - p$. The converse holds equivalently. Note also that if both row and column players use identical logit responses, the set of possible QRE is reduced to a curve extending from the center of the strategy space to the Nash Equilibrium $(p = \frac{1}{6}, q = \frac{1}{2})$ [21]. In summary, unlike the existence of a representative agent for each population, we do not necessarily have a single representative agent for all players.

5 Logit Responders

Most of the literature on QRE has assumed that the payoff disturbances are all independently drawn from an extreme value distribution, which generates tractable logit response functions:

$$x_{ij}(\pi_i) = \frac{e^{\lambda \pi_{ij}}}{\sum_{l=1}^{J_i} e^{\lambda \pi_{il}}} \tag{10}$$

where the parameter λ can be interpreted as the agents' level of rationality. As λ goes to infinity, agents best respond perfectly, producing a Nash Equilibrium. Conversely, as λ tends to zero, agents play the uniform mixed strategy, choosing each action with the same probability without regard to payoffs.

We now assume agents' quantal response functions take this logit form, but we preserve heterogeneity in the populations by allowing the agents to have their own individual rationality parameters. Thus,

$$Q_{ij}^{\mu}(\pi_i) = \frac{e^{\lambda_\mu \pi_{ij}}}{\sum_{l=1}^{J_i} e^{\lambda_\mu \pi_{il}}}. \tag{11}$$

For the purposes of this section, it suffices to consider finite populations of agents, so

$$x_{ij} = \frac{1}{m_i} \sum_{\mu=1}^{m_i} Q_{ij}^{\mu} \tag{12}$$

for all i and j.

It would be straightforward to apply the results from Section 3 and identify representative agents. In a truly heterogeneous population, i.e., with logit parameters not all degenerate, the representative agent will not be a logit responder. In this section, we see what happens when a theorist tries to force a homogenous logit equilibrium model on a population that is actually heterogeneous. Because the homogenous logit equilibrium is a mis-specified model of the populations we're assuming, the value of the single logit parameter will vary with the game being considered. But, a single logit parameter value can explain any particular choice probabilities between two actions if payoff monotonicity is preserved (i.e., if choice probabilities are increasing in the payoffs). For this reason, we restrict attention to a population with two pure strategies taking part in a (possibly larger) fixed game.

We identify a downward bias in the single logit parameter determined by the mis-specified homogenous model as compared with the average of the true logit parameters in use. Thus, the population seems to behave less rationally if the modeler believes the agents are all alike when in fact they each have their own levels of rationality. This bias is exacerbated as the magnitude of the difference in payoffs between the two actions grows.

First, we present a formula relating the logit parameter of the homogenous model to the true logit parameters and the equilibrium payoffs in the heterogeneous model. Let $J_i = 2$ in a particular population i. Fix equilibrium choice probabilities and payoffs in accordance with Equations (11) and (12), and denote them x_i^* and π_i^* respectively in population i. Assume a game in which the equilibrium payoffs to the two actions in population i are not equal, $\pi_{i1}^* \neq \pi_{i2}^*$. Denote by λ the logit parameter of the homogenous model describing behavior in population i. We can use Equation (10) to express λ in terms of the choice probabilities and payoffs in population i. The ratio of the choice probabilities is $\frac{x^*_{i1}}{x^*_{i2}} = e^{\lambda(\pi^*_{i1}-\pi^*_{i2})}$. Thus,

$$\lambda = \frac{1}{\pi^*_{i1} - \pi^*_{i2}} \ln\left(\frac{x^*_{i1}}{x^*_{i2}}\right). \tag{13}$$

Equation (13) could also be derived as the maximum likelihood estimate of the homogenous logit parameter given data on the equilibrium choice probabilities and payoffs in population i. With sufficiently many observations of choice probabilities and payoffs, this data should accurately reflect the equilibrium satisfying Equations (11) and (12). These equations, (11) and (12), give us the actual ratio of choice probabilities in population i. We plug into Equation (11) to get

$$Q^{\mu}_{i1}(\pi^*_i) = \frac{e^{\lambda_\mu \Delta\pi}}{e^{\lambda_\mu \Delta\pi} + 1}$$

and

$$Q^{\mu}_{i2}(\pi^*_i) = \frac{1}{e^{\lambda_\mu \Delta\pi} + 1},$$

where we let $\Delta\pi = \pi^*_{i1} - \pi^*_{i2}$ to simplify the notation. Then, using Equation (12),

$$\frac{x^*_{i1}}{x^*_{i2}} = \frac{\frac{1}{m_i}\sum_{\mu=1}^{m_i} \frac{e^{\lambda_\mu \Delta\pi}}{e^{\lambda_\mu \Delta\pi}+1}}{\frac{1}{m_i}\sum_{\mu=1}^{m_i} \frac{1}{e^{\lambda_\mu \Delta\pi}+1}}.$$

Finally, we obtain our desired formula:

$$\lambda = \frac{1}{\Delta\pi} \ln\left(\frac{\sum_{\mu=1}^{m_i} \frac{e^{\lambda_\mu \Delta\pi}}{e^{\lambda_\mu \Delta\pi}+1}}{\sum_{\mu=1}^{m_i} \frac{1}{e^{\lambda_\mu \Delta\pi}+1}}\right). \tag{14}$$

Observe that λ depends both on the heterogeneous logit parameters $\{\lambda_\mu\}$ and the equilibrium payoff difference $\Delta\pi$. We sometimes refer to the function given by Equation (14) as $\lambda(\{\lambda_\mu\}, \Delta\pi)$.

Our next result helps us interpret this formula. Theorem 5 says that this homogenous logit parameter is always less than the average of the heterogeneous logit parameters actually used by the agents. Moreover, the size of this bias in the homogenous model depends on the equilibrium payoffs. When the magnitude of the difference in payoffs between the two actions gets large, the homogenous logit parameter approaches the smallest of the heterogeneous logit parameters in the population. In this limit, the population behaves like its single most irrational agent. On the other hand, when the magnitude of the payoff difference gets small, the homogenous logit parameter approaches the average of the agents' true logit parameters.

Theorem 5 *Consider a quantal response equilibrium in accordance with Equations (11) and (12) such that population i has two actions with different equilibrium payoffs, i.e., $J_i = 2$, and $\Delta\pi \neq 0$. Let $\bar{\lambda} = \frac{1}{m_i}\sum_{\mu=1}^{m_i} \lambda_\mu$ be the average of the heterogeneous logit parameters used by the agents in population i, and let λ be the homogenous logit parameter that explains the population's choice probabilities for these particular payoffs. Then*

$$\lambda \leq \bar{\lambda} \tag{15}$$

with equality if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_{m_i}$.[4] *Additionally,*

$$\lim_{\Delta\pi \to \pm\infty} \lambda = \min\{\lambda_\mu\} \tag{16}$$

[4]The convention of using the parameter λ to represent a player's rationality is somewhat arbitrary in the sense that a modeler could just as well have defined $\kappa = e^{\lambda}$ to be the rationality parameter. Proposition 1 in the Appendix establishes an inequality analogous to (15), showing downward bias for such an alternative rationality parameter.

and

$$\lim_{\Delta\pi \to 0} \lambda = \bar{\lambda}. \tag{17}$$

Proof Equation (14) gives the exact value of λ. We twice apply Jensen's Inequality to pieces of this expression in order to derive (15).

Without loss of generality, assume action 1 has the higher equilibrium payoff so that $\Delta\pi > 0$. Then $\frac{1}{e^{\xi\,\Delta\pi}+1}$ is a concave up function of ξ. Applying Jensen's Inequality to this function,

$$\frac{1}{m_i} \sum_{\mu=1}^{m_i} \frac{1}{e^{\lambda_\mu \Delta\pi} + 1} \geq \frac{1}{e^{\bar{\lambda}\,\Delta\pi} + 1} \tag{18}$$

with equality if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_{m_i}$. Similarly, $\frac{e^{\xi\,\Delta\pi}}{e^{\xi\,\Delta\pi}+1}$, being equivalent to $1 - \frac{1}{e^{\xi\,\Delta\pi}+1}$, is a concave down function of ξ. So Jensen's Inequality implies

$$\frac{1}{m_i} \sum_{\mu=1}^{m_i} \frac{e^{\lambda_\mu \Delta\pi}}{e^{\lambda_\mu \Delta\pi} + 1} \leq \frac{e^{\bar{\lambda}\,\Delta\pi}}{e^{\bar{\lambda}\,\Delta\pi} + 1} \tag{19}$$

with equality if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_{m_i}$. When we plug into Equation (14), the denominators on the right-hand sides of (18) and (19) cancel, giving us

$$\begin{aligned} \lambda &\leq \frac{1}{\Delta\pi} \ln\left(e^{\bar{\lambda}\,\Delta\pi}\right) \\ &= \bar{\lambda}. \end{aligned}$$

Here again, equality holds exactly when $\lambda_1 = \lambda_2 = \ldots = \lambda_{m_i}$.

We prove the limits in (16) and (17) in the Appendix. ∎

Theorem 5 describes a downward bias in the determination of a homogenous logit parameter when agents are really heterogeneous. The less rational agents seem to leave a larger mark on the aggregate population behavior. This bias gets worse when one action's equilibrium payoff gets much larger than the other's. Conversely, the bias disappears as the payoff difference tends to zero.

Our formula for λ, Equation (14), also allows us to ask whether a determination of the homogenous logit parameter from data on a choice between two actions restricts the set of possible logit parameters for members of the population. The next result says it very well may. A large value of the homogenous logit parameter imposes a minimum possible value on the set of heterogeneous parameters. Conversely, a small homogenous logit parameter precludes any individual agent from having too large a value. For intermediate homogenous logit parameters, however, we cannot rule out any parameters for a single agent. Naturally, these bounds depend on the population size and are much less restrictive for a large population.

Theorem 6 *Retain the context of Theorem 5. If* $e^{\lambda\,|\Delta\pi|} > 2m_i - 1$, *then*

$$\min\{\lambda_\mu\} \geq \frac{1}{|\Delta\pi|} \ln\left(\frac{1}{m_i}\left(e^{\lambda\,|\Delta\pi|} - (m_i - 1)\right)\right).$$

If $e^{\lambda\,|\Delta\pi|} < \frac{m_i+1}{m_i-1}$, *then*

$$\max\{\lambda_\mu\} \leq \frac{1}{|\Delta\pi|} \ln\left(\frac{(m_i+1)e^{\lambda\,|\Delta\pi|} - (m_i - 1)}{m_i + 1 - (m_i - 1)e^{\lambda\,|\Delta\pi|}}\right).$$

Proof See Appendix.

Homogenous logit parameters are estimated in much of the experimental literature on two-by-two games, although often with data pooled across many populations and many games. Theorem 6 applies to a homogenous logit parameter calculated for a single population in a particular game. If we believe that agents use logit responses, but are heterogeneous in their levels of rationality, this theorem translates a mis-specified homogenous logit parameter into restrictions on the set of possible logit parameters in a finite population.

To illustrate these results, we can compare a homogenous logit model fit to data in a two-by-two symmetric game to compatible heterogeneous logit models featuring two types of responders – one with a high rationality parameter and the other with a low one. To make the example as simple as possible, we assume exactly half the agents are of each type (though with the data coming from an experiment on 214 subjects, we have no reason to actually believe there are just two types). We consider Guyer and Rapoport's "No Conflict" game, Game #6 in their series of experiments [13]. The payoff matrix is:

No Conflict

	A_2	B_2
A_1	4, 4	2, 3
B_1	3, 2	1, 1

The players have a dominant strategy choosing action A. Guyer and Rapoport observe action A played 90% of the time. Choosing to model this as a homogenous logit equilibrium, we have an equilibrium payoff difference $\Delta\pi = 1$ (as the payoff to A happens to always exceed the payoff to B by one), and thus $\lambda = \ln(9)$ in accordance with Equation 13.[5]

Plugging $\Delta\pi = 1$ and $\lambda = \ln(9)$ into Equation 14 produces an equation implicitly relating λ_1 and λ_2. Figure 1 shows possible values of these heterogeneous logit parameters. Pairs of λ_1 and λ_2 values are determined by fixed x-values in the graph. Larger x-values correspond to greater dispersion in the heterogeneous logit parameter values, but the scaling along this axis is arbitrary. We can see that the average of λ_1 and λ_2 always exceeds $\ln(9)$, and the lower value is bounded below by $\ln(4)$ while the higher value may be arbitrarily large. Guyer and Rapoport's data thus puts a bound on how irrational the low-type agents can be, and they only approach this bound if the other agents are hyper-rational.

Because a homogenous logit model is mis-specified in the presence of heterogeneity, estimates of a single rationality parameter do not translate across different game environments. Theorems 5 and 6 imply the following result, which tells us that an estimate of a homogenous rationality parameter in a particular game environment places no restriction on such an estimate in an alternative game environment, even with a working assumption that agents' rationality levels are fixed across games. Theorem 7 states that the set of heterogeneous logit parameters that is consistent with a given logit equilibrium in any one game could in some other game give rise to behavior consistent with any other homogenous logit parameter.

[5]Goeree and Holt [9] estimate a homogenous logit parameter from data pooled across 37 games, including this one, from Guyer and Rapoport's study. We obtain a different value of the homogenous logit parameter because we use data from just this one game.

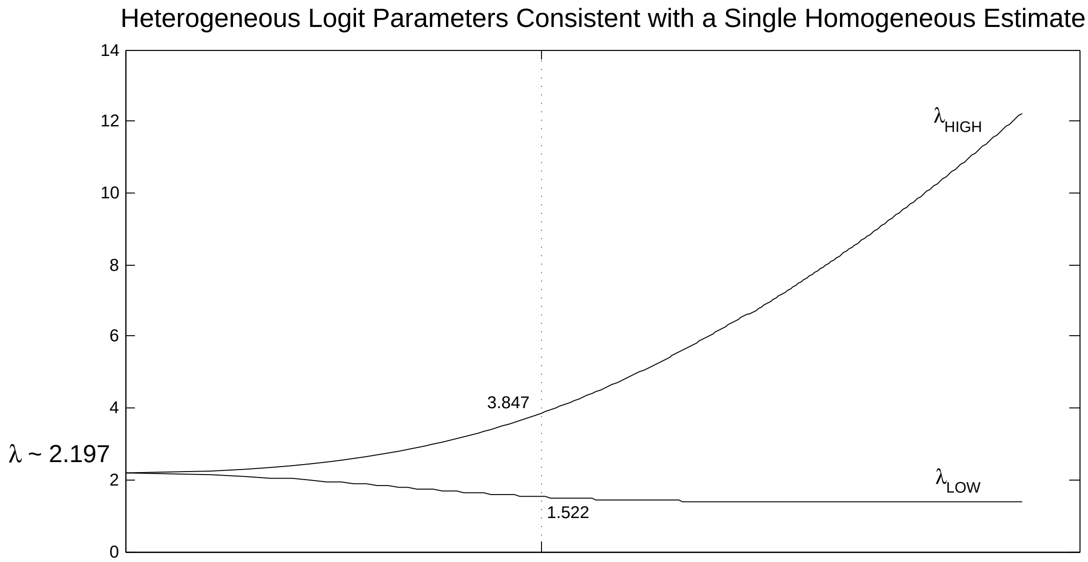

Figure 1: Possible values of a pair of logit parameters (determined at any fixed x-value) that would be consistent with a homogeneous $\lambda = \ln(9)$, when $\Delta\pi = 1$. These values fit data from Guyer and Rapoport's (1972) "No Conflict" game.

Theorem 7 *Consider normal form games for which population i has two actions, $J_i = 2$. For any logit equilibrium with population i having payoff difference $\Delta\pi^* \neq 0$ and rationality parameter $\lambda^* > 0$ in such a game Γ, and any alternative value $\lambda' > 0$, there exists a set of heterogeneous logit parameters $\{\lambda'_\mu\}$ that are consistent with the homogenous logit model applied to population i in Γ,*

$$\lambda\left(\{\lambda'_\mu\}, \Delta\pi^*\right) = \lambda^*, \tag{20}$$

and there exists a game Γ' with a heterogeneous logit equilibrium in which population i has payoff difference $\Delta\pi' \neq 0$, such that

$$\lambda\left(\{\lambda'_\mu\}, \Delta\pi'\right) = \lambda'. \tag{21}$$

Proof See Appendix.

Recall that $\lambda\left(\{\lambda_\mu\}, \Delta\pi\right)$ gives the homogenous logit parameter that produces the same equilibrium choice probabilities as the heterogeneous logit parameters $\{\lambda_\mu\}$ when the equilibrium payoff difference is $\Delta\pi$. Thus, Equation (20) means that any estimate of a homogenous rationality parameter in a given game environment can be explained by some set of heterogeneous logit parameters, and Equation (21) means that these heterogeneous logit parameters could be consistent with any other homogenous parameter in an alternative game environment. We should not expect mis-specified parameter estimates to accurately describe behavior across all games.

6 Discussion

We have proposed a model of heterogeneous populations playing quantal response equilibria. The paper contributes general results that apply to quantal response equilibria without specification of their functional form as well as particular results that are specific to the logit response model.

We have paid extra attention to the logit specification because it is so commonly employed in practice. The representative agent for a population of heterogeneous logit responders is not another logit responder. In the case of heterogeneous logit responders choosing between two pure strategies, we have obtained a formula (Equation 14) relating a mis-specified homogenous logit parameter to the actual heterogeneous parameters in the population. Maximum likelihood estimation could be used to fit a homogenous logit parameter to the behavior of heterogeneous agents choosing between any number of pure strategies, but a closed form solution is not generally possible. Our formula provides insights in two directions. It tells us that the homogenous model is biased towards less rationality, as the homogenous logit parameter is always less than the average of the heterogeneous ones. It also allows us to bound the possible values of the true logit parameters if we have a mis-specified homogenous model already in place.

These results are applicable to experimental work in which a homogenous logit model has been fit to data. One particular extension is to explicitly model the existence of clueless players by giving some fraction of the agents a logit parameter of zero. This would address the common problem of some subjects not understanding the game they are playing [3].

Working with a general model that does not assume that quantal responses take any particular functional forms, we have found that representative agents exist for heterogeneous populations if we allow weakly admissible payoff disturbances. A representative agent chooses strategies in the same proportions as the entire population, but does not have payoff disturbances distributed in the same proportions as the population. In games with many pure strategies, representative behavior cannot arise from any iid distribution of disturbances.

This impossibility of having a representative agent with disturbances iid across actions stems from the fact that averaging probability distributions almost never preserves independence. Thus, if we believe populations of agents are heterogeneous, but desire representative-agent models, we must be willing to consider noise terms that are jointly dependent across actions. Our findings support the use of regular quantal response functions. Regular quantal response equilibrium does generate falsifiable predictions and is consistent with the representative-agent framework.

Appendix

Lemma 1 *$\hat{\phi}(t)$ is continuous at $t = 0$.*

Proof Recall that $|\phi(t)| \leq 1$ and $\phi(0) = 1$ for all ϕ and thus for $\hat{\phi}$ as well. We will show for all $h > 0$, there exists $k > 0$ such that $\|t\| < k$ implies $\operatorname{Re}\left\{\hat{\phi}(t)\right\} > 1 - h$. Let $\mathcal{K}$ be a compact subset of characteristic functions ϕ such that $\int_{\mathcal{K}} \Psi[\phi]\, d\phi > 1 - \frac{h}{4}$. Because all the ϕ are continuous at $t = 0$, we can choose $k[\phi] > 0$ such that $\operatorname{Re}\{\phi(t)\} > 1 - \frac{h}{2}$ for all $\|t\| < k[\phi]$. Then take $k = \min_{\phi \in \mathcal{K}} k[\phi]$, and $k > 0$ because the minimum is taken over a compact space and the extreme value theorem applies. We then obtain for all $\|t\| < k$,

$$\begin{aligned} \operatorname{Re}\left\{\hat{\phi}(t)\right\} &= \int_{\phi \in \mathcal{K}} \operatorname{Re}\{\phi(t)\}\, \Psi[\phi]\, d\phi + \int_{\phi \notin \mathcal{K}} \operatorname{Re}\{\phi(t)\}\, \Psi[\phi]\, d\phi \\ &> \left(1 - \frac{h}{2}\right)\left(1 - \frac{h}{4}\right) + (-1)\left(\frac{h}{4}\right) \\ &= 1 - h + \frac{h^2}{8} \;>\; 1 - h. \qquad \blacksquare \end{aligned}$$

Lemma 2 *The Cauchy principal value of the mean of $\hat{P}(\cdot)$ is 0. Additionally, if the random vector $(\bar{\delta}_{1j}, \ldots, \widehat{\bar{\delta}_{jj}}, \ldots, \bar{\delta}_{Jj})$ is distributed according to $\hat{P}(\cdot)$, then*

$$\lim_{\gamma \to \infty} \gamma \operatorname{Prob}\left\{|\bar{\delta}_{j'j}| \geq \gamma\right\} = 0 \textit{ for all } j'.$$

Proof A property of characteristic functions is $\left[\frac{\partial}{\partial t_{j'}}\phi(t)\right]_{t=0}$ exists if and only if:

(i) $\mathrm{PV}\left\langle \delta_{j'j} \right\rangle$ exists and

(ii) $\lim_{\gamma\to\infty}\gamma\,\mathrm{Prob}\left\{|\delta_{j'j}|\geq\gamma\right\}=0$,

and when these conditions are satisfied, $\left[\frac{\partial}{\partial t_{j'}}\phi(t)\right]_{t=0} = i\,\mathrm{PV}\left\langle \delta_{j'j} \right\rangle$ [27, 23]. So, it suffices to show $\left[\frac{\partial}{\partial t_{j'}}\hat{\phi}(t)\right]_{t=0} = 0$ for all j'. Differentiability of $\hat{\phi}(t)$ follows from the differentiability of all ϕ in the support of Ψ, using an argument completely analogous to the proof of continuity of $\hat{\phi}(t)$, Lemma 1. Thus, $\left[\frac{\partial}{\partial t_{j'}}\hat{\phi}(t)\right]_{t=0} = \int \Psi[\phi]\left[\frac{\partial\phi}{\partial t_{j'}}\right]_{t=0} d\phi$. For all ϕ in the support of Ψ, all j', $\left[\frac{\partial\phi}{\partial t_{j'}}\right]_{t=0} = 0$ because $\mathrm{PV}\left\langle \delta_{j'j} \right\rangle = 0$ and $\lim_{\gamma\to\infty}\gamma\,\mathrm{Prob}\left\{|\delta_{j'j}|\geq\gamma\right\}=0$. Each $\delta_{j'j}$ must satisfy these two conditions because the underlying ϵ_j and $\epsilon_{j'}$ are required to obey them by assumption (a3) or (a3'). ∎

Proposition 1 *Retain the context of Theorem 5. Let $f:\Re_+\to\Re$ be a monotonically increasing function, and $g:\Re\to\Re_+$ be its inverse, $g=f^{-1}$. Denote $\kappa=f(\lambda)$ and $\kappa_\mu=f(\lambda_\mu)$ for all μ. Let $\bar{\kappa}=\frac{1}{m_i}\sum_{\mu=1}^{m_i}\kappa_\mu$. If*

$$\frac{g''(\xi)}{\left(g'(\xi)\right)^2} < \Delta\pi\left(\frac{e^{g(\xi)\Delta\pi}-1}{e^{g(\xi)\Delta\pi}+1}\right) \text{ for all } \xi\in\left[\min\{\kappa_\mu\},\max\{\kappa_\mu\}\right], \tag{22}$$

then $\kappa\leq\bar{\kappa}$ with equality if and only if $\kappa_1=\kappa_2=\ldots=\kappa_{m_i}$.[6]

[6]Note that the hypothesis is satisfied when f is an exponential function, i.e., for $\kappa=e^\lambda$, because in this case $g(\xi)=\ln(\xi)$ and g'' is always negative whereas the right-hand side of Inequality (22) is always positive.

Proof It is straightforward, albeit tedious, to take a second derivative of $\frac{1}{e^{g(\xi)\Delta\pi}+1}$ and obtain Inequality (22) as the condition implying that this function is concave up (assuming once again $\Delta\pi > 0$ without loss of generality). By the logic used to prove Theorem 5, $\lambda \leq g(\bar{\kappa})$ with equality if and only if $\kappa_1 = \kappa_2 = \ldots = \kappa_{m_i}$. Because f is monotonically increasing, we can apply it to both sides of this inequality to obtain $\kappa \leq \bar{\kappa}$. ∎

Completing the Proof of Theorem 5.

To obtain

$$\lim_{\Delta\pi \to \pm\infty} \lambda = \min\{\lambda_\mu\},$$

we take the limit as $\Delta\pi$ goes to ∞. By symmetry, the result then holds when $\Delta\pi$ goes to $-\infty$ as well. First, we use algebra in Equation (14) to come up with a new expression for λ:

$$\lambda = \frac{1}{\Delta\pi} \ln \left(\frac{\sum_{\mu=1}^{m_i} e^{\lambda_\mu \Delta\pi} \prod_{\omega \neq \mu} \left(e^{\lambda_\omega \Delta\pi} + 1\right)}{\sum_{\mu=1}^{m_i} \prod_{\omega \neq \mu} \left(e^{\lambda_\omega \Delta\pi} + 1\right)} \right).$$

In the limit of $\Delta\pi$ going to ∞,

$$\begin{aligned} e^{\lambda_\mu \Delta\pi} \prod_{\omega \neq \mu} \left(e^{\lambda_\omega \Delta\pi} + 1\right) &\to e^{\lambda_\mu \Delta\pi} \prod_{\omega \neq \mu} e^{\lambda_\omega \Delta\pi} \\ &= \prod_{\omega} e^{\lambda_\omega \Delta\pi} \end{aligned}$$

and

$$\sum_{\mu=1}^{m_i} \prod_{\omega \neq \mu} \left(e^{\lambda_\omega \Delta\pi} + 1\right) \to \prod_{\omega \neq \arg\min\{\lambda_\mu\}} e^{\lambda_\omega \Delta\pi}.$$

Thus,

$$\lim_{\Delta\pi\to\infty}\lambda = \lim_{\Delta\pi\to\infty}\frac{1}{\Delta\pi}\ln\left(\frac{\sum_{\mu=1}^{m_i}\prod_{\omega}e^{\lambda_\omega\Delta\pi}}{\prod_{\omega\neq\arg\min\{\lambda_\mu\}}e^{\lambda_\omega\Delta\pi}}\right)$$
$$= \lim_{\Delta\pi\to\infty}\frac{1}{\Delta\pi}\ln\left(m_i\, e^{\min\{\lambda_\mu\}\Delta\pi}\right)$$
$$= \lim_{\Delta\pi\to\infty}\frac{\min\{\lambda_\mu\}\Delta\pi + \ln(m_i)}{\Delta\pi}$$
$$= \min\{\lambda_\mu\}.$$

To obtain

$$\lim_{\Delta\pi\to 0}\lambda = \bar{\lambda},$$

we apply l'Hospital's Rule to the expression for λ given in Equation (14). We have

$$\frac{d}{d\Delta\pi}\left[\ln\left(\frac{\sum_{\mu=1}^{m_i}\frac{e^{\lambda_\mu\Delta\pi}}{e^{\lambda_\mu\Delta\pi}+1}}{\sum_{\mu=1}^{m_i}\frac{1}{e^{\lambda_\mu\Delta\pi}+1}}\right)\right] =$$
$$\left(\frac{\sum_{\mu=1}^{m_i}\frac{1}{e^{\lambda_\mu\Delta\pi}+1}}{\sum_{\mu=1}^{m_i}\frac{e^{\lambda_\mu\Delta\pi}}{e^{\lambda_\mu\Delta\pi}+1}}\right)\frac{\left(\sum_{\mu=1}^{m_i}\frac{1}{e^{\lambda_\mu\Delta\pi}+1}\right)\left(\sum_{\mu=1}^{m_i}\frac{\lambda_\mu e^{\lambda_\mu\Delta\pi}}{\left(e^{\lambda_\mu\Delta\pi}+1\right)^2}\right) - \left(\sum_{\mu=1}^{m_i}\frac{e^{\lambda_\mu\Delta\pi}}{e^{\lambda_\mu\Delta\pi}+1}\right)\left(\sum_{\mu=1}^{m_i}\frac{-\lambda_\mu e^{\lambda_\mu\Delta\pi}}{\left(e^{\lambda_\mu\Delta\pi}+1\right)^2}\right)}{\left(\sum_{\mu=1}^{m_i}\frac{1}{e^{\lambda_\mu\Delta\pi}+1}\right)^2}.$$

So

$$\frac{d}{d\Delta\pi}\left[\ln\left(\frac{\sum_{\mu=1}^{m_i}\frac{e^{\lambda_\mu\Delta\pi}}{e^{\lambda_\mu\Delta\pi}+1}}{\sum_{\mu=1}^{m_i}\frac{1}{e^{\lambda_\mu\Delta\pi}+1}}\right)\right]_{\Delta\pi=0} = (1)\frac{\left(\frac{m_i}{2}\right)\left(\sum_{\mu=1}^{m_i}\frac{\lambda_\mu}{4}\right) - \left(\frac{m_i}{2}\right)\left(\sum_{\mu=1}^{m_i}\frac{-\lambda_\mu}{4}\right)}{\left(\frac{m_i}{2}\right)^2}$$
$$= \frac{\sum_{\mu=1}^{m_i}\lambda_\mu}{m_i}$$
$$= \bar{\lambda}.$$

The denominator in (14) is $\Delta\pi$, so its derivative is 1. Thus,

$$\lim_{\Delta\pi\to 0}\lambda = \bar{\lambda}. \quad \blacksquare$$

Proof of Theorem 6.

Without loss of generality, assume $\Delta\pi > 0$.

To obtain

$$\min\{\lambda_\mu\} \geq \frac{1}{\Delta\pi} \ln\left(\frac{1}{m_i}\left(e^{\lambda\,\Delta\pi} - (m_i - 1)\right)\right),$$

we make use of the following inequalities: $\frac{e^{\lambda_\mu \Delta\pi}}{e^{\lambda_\mu \Delta\pi}+1} \leq 1$ and $\frac{1}{e^{\lambda_\mu \Delta\pi}+1} \geq 0$ for all $\mu \neq \arg\min\{\lambda_\omega\}$. Applying these inequalities to our formula for λ in Equation (14), we get

$$\begin{aligned}
\lambda &\leq \frac{1}{\Delta\pi} \ln\left(\frac{\frac{e^{\min\{\lambda_\mu\}\Delta\pi}}{e^{\min\{\lambda_\mu\}\Delta\pi}+1} + m_i - 1}{\frac{1}{e^{\min\{\lambda_\mu\}\Delta\pi}+1}}\right) \\
&= \frac{1}{\Delta\pi} \ln\left(e^{\min\{\lambda_\mu\}\Delta\pi} + (m_i - 1)(e^{\min\{\lambda_\mu\}\Delta\pi} + 1)\right) \\
&= \frac{1}{\Delta\pi} \ln\left(m_i\, e^{\min\{\lambda_\mu\}\Delta\pi} + m_i - 1\right).
\end{aligned}$$

So

$$e^{\lambda\,\Delta\pi} \leq m_i\, e^{\min\{\lambda_\mu\}\Delta\pi} + m_i - 1,$$

and thus,

$$\frac{1}{\Delta\pi} \ln\left(\frac{1}{m_i}\left(e^{\lambda\,\Delta\pi} - (m_i - 1)\right)\right) \leq \min\{\lambda_\mu\}.$$

Note that this bound is meaningful only if $e^{\lambda\,\Delta\pi} > 2m_i - 1$.

To obtain

$$\max\{\lambda_\mu\} \leq \frac{1}{\Delta\pi} \ln\left(\frac{(m_i+1)e^{\lambda\,\Delta\pi} - (m_i - 1)}{m_i + 1 - (m_i - 1)e^{\lambda\,\Delta\pi}}\right),$$

we follow a similar approach using the fact that $\frac{e^{\lambda_\mu \Delta\pi}}{e^{\lambda_\mu \Delta\pi}+1} \geq \frac{1}{2}$ and $\frac{1}{e^{\lambda_\mu \Delta\pi}+1} \leq \frac{1}{2}$ for all

$\mu \neq \arg\max\{\lambda_\omega\}$. Putting these inequalities into Equation (14) produces

$$\lambda \geq \frac{1}{\Delta\pi} \ln \left(\frac{\frac{e^{\max\{\lambda_\mu\}\Delta\pi}}{e^{\max\{\lambda_\mu\}\Delta\pi}+1} + (m_i - 1)\frac{1}{2}}{\frac{1}{e^{\max\{\lambda_\mu\}\Delta\pi}+1} + (m_i - 1)\frac{1}{2}} \right)$$

$$= \frac{1}{\Delta\pi} \ln \left(\frac{2e^{\max\{\lambda_\mu\}\Delta\pi} + (m_i - 1)(e^{\max\{\lambda_\mu\}\Delta\pi} + 1)}{2 + (m_i - 1)(e^{\max\{\lambda_\mu\}\Delta\pi} + 1)} \right)$$

$$= \frac{1}{\Delta\pi} \ln \left(\frac{(m_i + 1)e^{\max\{\lambda_\mu\}\Delta\pi} + m_i - 1}{(m_i - 1)e^{\max\{\lambda_\mu\}\Delta\pi} + m_i + 1} \right).$$

So

$$e^{\lambda\,\Delta\pi} \geq \frac{(m_i + 1)e^{\max\{\lambda_\mu\}\Delta\pi} + m_i - 1}{(m_i - 1)e^{\max\{\lambda_\mu\}\Delta\pi} + m_i + 1},$$

and thus,

$$(m_i + 1)e^{\lambda\,\Delta\pi} - (m_i - 1) \geq (m_i + 1)e^{\max\{\lambda_\mu\}\Delta\pi} - (m_i - 1)e^{\lambda\,\Delta\pi}\, e^{\max\{\lambda_\mu\}\Delta\pi},$$

and finally,

$$\frac{1}{\Delta\pi} \ln \left(\frac{(m_i + 1)e^{\lambda\,\Delta\pi} - (m_i - 1)}{m_i + 1 - (m_i - 1)e^{\lambda\,\Delta\pi}} \right) \geq \max\{\lambda_\mu\}.$$

This bound is meaningful only if $e^{\lambda\,\Delta\pi} < \frac{m_i+1}{m_i-1}$. ∎

Proof of Theorem 7.

Choose $m_i \in \mathbb{N}$ such that $m_i > \frac{e^{\lambda^*\Delta\pi^*}+1}{2}$ and $m_i > \frac{e^{\lambda^*\Delta\pi^*}+1}{e^{\lambda^*\Delta\pi^*}-1}$. This ensures that neither of the bounds in Theorem 6 apply. Thus, we can take $\lambda'_1 = 0$ and $\lambda'_{m_i} > m_i\,\lambda'$ and still be able to choose the remaining $\{\lambda'_\mu\}$, for $\mu = 2 \ldots m_i - 1$, such that Equation (20) holds. That means these heterogeneous logit parameters will be consistent with the homogenous logit model with rationality parameter λ^* and equilibrium payoff difference $\Delta\pi^*$. We have specifically chosen λ'_1 and λ'_{m_i} so that

$\lambda'_1 < \lambda' < \overline{\lambda'}$. Thus, noting the limits we take in Theorem 5, we establish that $\lambda\left(\{\lambda'_\mu\}, \Delta\pi\right)$ is above λ' when $\Delta\pi \approx 0$ and is below λ' when $\Delta\pi$ is large. Because $\lambda\left(\{\lambda'_\mu\}, \Delta\pi\right)$ is continuous in $\Delta\pi$, there is some $\Delta\pi'$ for which $\lambda\left(\{\lambda'_\mu\}, \Delta\pi'\right) = \lambda'$. ∎

References

[1] Anderson, S., Goeree, J., and Holt, C. (2002) The Logit Equilibrium: A Perspective on Intuitive Behavioral Anomalies. *Southern Economic Journal* 69(1), 21-47.

[2] Anderson, S., Goeree, J., and Holt, C. (2004) Noisy Directional Learning and the Logit Equilibrium. *Scandinavian Journal of Economics* 106(3), 581-602.

[3] Camerer, C. (2003) *Behavioral Game Theory.* Princeton University Press, Princeton, NJ.

[4] Camerer, C., Palfrey, T., and Rogers, B. (2007) Heterogeneous Quantal Response Equilibrium and Cognitive Hierarchies. Working Paper, California Institute of Technology, http://www.hss.caltech.edu/~trp/hqre20070315-jet.pdf

[5] Capra, M., Goeree, J., Gomez, R., and Holt, C. (1999) Anomalous Behavior in a Traveler's Dilemma? *American Economic Review* 89(3), 678-690.

[6] Fey, M., McKelvey, R., and Palfrey, T. (1996) An Experimental Study of Constant-Sum Centipede Games. *International Journal of Game Theory* 25(3), 269-287.

[7] Fudenberg, D., and Levine, D. (1998) *The Theory of Learning in Games*. MIT Press, Cambridge, MA.

[8] Goeree, J., and Holt, C. (1999) Stochastic Game Theory: For Playing Games, Not Just for Doing Theory. *Proceedings of the National Academy of Sciences* 96(19), 10564-10567.

[9] Goeree, J., and Holt, C. (2004) A Model of Noisy Introspection. *Games and Economic Behavior* 46(2), 365-382.

[10] Goeree, J., Holt, C., and Palfrey, T. (2002) Quantal Response Equilibrium and Overbidding in Private-Value Auctions. *Journal of Economic Theory* 104(1), 247-272.

[11] Goeree, J., Holt, C., and Palfrey, T. (2005) Regular Quantal Response Equilibrium. *Experimental Economics* 8, 347-367.

[12] Golman, R. (2008) Why Learning Doesn't Add Up: Equilibrium Selection and Compositions of Learning Rules. Working Paper, University of Michigan.

[13] Guyer, M., and Rapoport, A. (1972) 2x2 Games Played Once. *The Journal of Conflict Resolution* 16(3), 409-431.

[14] Haile, P., Hortacsu, A., and Kosenok, G. (2008) On the Empirical Content of Quantal Response Equilibrium. *American Economic Review* 98(1), 180-200.

[15] Ho, T., Wang, X., and Camerer, C. (2008) Individual Differences in EWA Learning with Partial Payoff Information. *The Economic Journal* 118, 37-59.

[16] Hofbauer, J., and Hopkins, E. (2005) Learning in Perturbed Asymmetric Games. *Games and Economic Behavior* 52(1), 133-152.

[17] Hofbauer, J., and Sandholm, W.H. (2002) On the Global Convergence of Stochastic Fictitious Play. *Econometrica* 70, 2265-2294.

[18] Hommes, C. (2006) Heterogeneous Agent Models in Economics and Finance. In: Judd, K., Tesfatsion, L. (eds) Handbook of Computational Economics, vol 2, Agent-Based Computational Economics. North Holland, Amsterdam.

[19] Hopkins, E. (1999) A Note on Best Response Dynamics. *Games and Economic Behavior* 29, 138-150.

[20] Kirman, A. (2006) Heterogeneity in Economics. *Journal of Economic Interaction and Coordination* 1, 89-117.

[21] McKelvey, R., and Palfrey, T. (1995) Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10(1), 6-38.

[22] McKelvey, R., Palfrey, T., and Weber, R. (2000) The effects of payoff magnitude and heterogeneity on behavior in 2x2 games with unique mixed strategy equilibria. *Journal of Economic Behavior & Organization* 42, 523-548.

[23] Pitman, E. J. G. (1956) On the Derivatives of a Characteristic Function at the Origin. *The Annals of Mathematical Statistics* 27(4), 1156-1160.

[24] Tsakas, E. (2005) Mixed Quantal Response Equilibria for Normal Form Games. Paper presented at the International Conference on Game Theory.

[25] Weibull, J. (1994) The Mass-Action interpretation of Nash equilibrium. Paper presented at Nobel symposium in honor of John Nash.

[26] Wilcox, N. (2006) Theories of Learning in Games and Heterogeneity Bias. *Econometrica* 74, 1271-1292.

[27] Wolfe, S. (1973) On the Local Behavior of Characteristic Functions. *The Annals of Probability* 1(5), 862-866.