The Quantal Response Equilibrium

Objective: Understand what the Quantal Response Equilibrium (QRE) is and how it can be applied.

A pure strategy Nash equilibrium makes sharp predictions about how people will play a game. Players pick a strategy and play it 100% of the time. A mixed strategy equilibrium prediction is also sharp in the sense that an individual randomly chooses a strategy based on some fixed probability that depends only on opponents' payoffs. This type of behavior is the exception rather than the rule in lab experiments. In the lab, people often deviate from the sharp predictions of the Nash equilibrium. Furthermore, these deviations are sensitive to variations in a player's own payoffs as well as opponents' payoffs in unpredicted ways.

The QRE was developed as a probabilistic extension of the Nash equilibrium that may help explain why people might systematically deviate. The basic idea is that people are likely but not certain to pick the strategy with the highest expected payoff due to unobservable random factors (e.g. unobserved risk preferences). Therefore, strategy choices in the QRE are probabilistic rather than deterministic. But this is also true for a mixed strategy Nash equilibrium. However, the QRE is distinct from a mixed strategy Nash equilibrium because it depends on a player's own payoffs as well as his opponents.

It should also be noted that the QRE is founded on well established empirical principles, the same principles that have been used to develop statistical models of discrete choice. This is in stark contrast to most other equilibrium concepts that are founded primarily in mathematical theory and the ingenuity of the theorist. The QRE is designed for and well suited to the empirical analysis of gaming behavior.

We will confine our treatment to finite normal form games. Let $N = \{1, 2,\ldots, n\}$ be the set of players. The ith player has the set of pure strategies $S^i = \{s_1^i, s_2^i,\ldots, s_{J(i)}^i\}$ where $J(i)$ is the number of pure strategies. A strategy profile is defined as $s = \{s^1,s^2,\ldots,s^n\}$ where $s^i \in S^i\ \forall\ i \in N$. $S = \times_{i\in N}S^i$ is the set of all possible strategy profiles. The function $u^i(s) \in \Re$ represents the ith player's payoff from strategy profile s. The collection of these payoffs for all players is $u(s) = \{u^1(s), u^2(s),\ldots, u^n(s)\}$.

Define Δ^i as a $J(i)$ dimensional simplex and $\Delta = \times_{i\in N} \Delta^i$. Let $p^i = (p_1^i,p_2^i,\ldots,p_{J(i)}^i) \in \Delta^i$ represent a probability distribution over player i's strategies such that $\sum_{j=1}^{J(i)}p_j^i = 1$ and $p_j^i \geq 0\ \forall\ j = 1, 2,\ldots, J(i)$. Let $p = \{p^1, p^2,\ldots, p^n\} \in \Delta$. Define $p^i(s^i)$ as the probability player i chooses strategy s^i and $p(s) = \Pi_{i=1}^{n}p^i(s^i)$ as the probability of strategy profile $s \in S$ given p. Player i's expected payoff is $Eu^i(p) = \sum_{s\in S} p(s)u^i(s)$.

Given this specification of the game, $p' = \{p^{i\prime},p^{-i\prime}\}$ is a Nash equilibrium if for all $i \in N$ and all $p^i \in \Delta^i$, $Eu^i(p^{i\prime},p^{-i\prime}) \geq Eu^i(p^i,p^{-i\prime})$ where $p^{-i\prime}$ is p' exclusive of $p^{i\prime}$.

Now define $Eu_j^i(p) = Eu^i(s_j^i,p^{-i}) + \varepsilon_j^i$ where ε_j^i is a random error. Define $\varepsilon^i = (\varepsilon_1^i,\varepsilon_2^i,\ldots,\varepsilon_{J(i)}^i)$ with joint distribution $f^i(\varepsilon^i)$ such that $E(\varepsilon^i) = 0$ and assume the marginal distribution $f_j^i(\varepsilon_j^i)$ exists for all $j \in J(i)$ and $i \in N$. Let $\varepsilon = (\varepsilon^1,\varepsilon^2,\ldots,\varepsilon^n)$ and $f = (f^1,f^2,\ldots,f^n)$.

Assume player i chooses the strategy j when $Eu_j^i(p) \geq Eu_k^i(p)\ \forall\ k = 1,2,\ldots,J(i)$. That is, a player's strategy choice is the strategy with the highest expected payoff, subject to some error, given the probability of his opponents' strategy choices. It is important to note that players know their optimal strategy which may or may not be pure. In the event that these strategies are pure, it is natural to question why we have all these probabilities floating around. What is important to realize is that the ith player knows ε^i, but not ε^k for $i \neq k$. Instead, we assume player i only knows the distribution $f^k(\varepsilon^k)$, which means k's strategy choice is random from the perspective of i. However, k's strategy choice is not uniformly random because it also depends on his payoffs and lack of knowledge of j's specific strategy choice.

Now define $R_j^i(p) = \{\varepsilon^i | Eu_j^i(p) \geq Eu_k^i(p)\ \forall\ k = 1,2,\ldots,J(i)\}$ and $\sigma_j^i(p) = \int_{R_j^i(p)} f(e)de$. $R_j^i(p)$ captures the region of possible errors that make strategy j player i's optimal strategy, while $\sigma_j^i(p)$ reflects the probability the realized errors fall in this region. The QRE is then any $\pi \in \Delta$ such that $\pi_j^i = \sigma_j^i(\pi)\ \forall\ j = 1, 2,\ldots, J(i)$ and $i \in N$ where π_j^i is the probability player i chooses strategy j.

To find the QRE, we need to find a set of probabilities for each player's strategies that are mutually consistent with player's choosing those strategies subject to some error. McKelvey and Palfrey go on to show that in finite normal form games a QRE exists under very general conditions (the proof is reminiscent of the mixed strategy Nash equilibrium proof). There could be more than one however.

Assuming the ε_j^i's are identically and independently distributed (iid) extreme value (Weibull), the QRE implies the logit function $\pi_j^i = \dfrac{e^{l Eu^i\left(s_j^i, p^{-i}\right)}}{\sum_{k=1}^{J(i)} e^{l Eu^i\left(s_k^i, p^{-i}\right)}}$ where λ is a parameter that is inversely related to the dispersion or variance of error. For $\lambda = 0$, probabilities are uniform. As λ approaches ∞, the dispersion of error approaches 0 and the QRE approaches a Nash equilibrium.

To get a better grasp on what is required for a QRE, let us work with this logit specification and the simple game described in Table 1.

Table 1

		Player 2: s_1^2	Player 2: s_2^2
Player 1	s_1^1	u_{11}^1 ; u_{11}^2	u_{12}^1 ; u_{12}^2
	s_2^1	u_{21}^1 ; u_{21}^2	u_{22}^1 ; u_{22}^2

Note that player 1's payoffs are in the lower left-hand corner of the cell, while player 2's are in the upper right-hand corner. Since each player has only two strategies and probabilities must sum to 1, we can define $\pi_1^i = \pi^i$ and $\pi_2^i = 1 - \pi^i$ for i = 1, 2. We then have two equations and two unknowns for which we can solve:

(1) $$p^1 = \frac{e^{l\left(p^2 u_{11}^1 + \left(1-p^2\right)u_{12}^1\right)}}{e^{l\left(p^2 u_{11}^1 + \left(1-p^2\right)u_{12}^1\right)} + e^{l\left(p^2 u_{21}^1 + \left(1-p^2\right)u_{22}^1\right)}}$$, and

(2) $$p^2 = \frac{e^{l\left(p^1 u_{11}^2 + (1-p_1)u_{21}^2\right)}}{e^{l\left(p^1 u_{11}^2 + \left(1-p^1\right)u_{21}^2\right)} + e^{l\left(p^1 u_{12}^2 + \left(1-p^1\right)u_{22}^2\right)}}.$$

Unfortunately, this system of equations does not have a nice closed form solution, but that's O.K. because there are lots of computer algorithms we can use to solve them given some more details about the game.

For example, let us look at the Prisoner's Dilemma game in Table 2.

Table 2

		Prisoner 2: Deny	Prisoner 2: Confess
Prisoner 1	Deny	10 / 10	0 / 15
	Confess	15 / 0	5 / 5

Figure 1 shows the QRE equilibrium values of π^j, the probability Prisoner j chooses to Deny the allegations, as the value of λ increases. Since the game is symmetric, $\pi^j = \pi^1 = \pi^2$. For $\lambda = 0$, Deny and Confess are played with equal probability. As λ increases, the probability of Deny decreases monotonically approaching 0, the Nash equilibrium.

Experimental investigations using one shot Prisoner's Dilemma games do find that people sometimes choose Deny, which low values of λ support. Repeated Prisoner's Dilemma games can produce this type of cooperation also but with probabilities well in excess of 0.5, which suggests there is probably something else going on. Later we will see other explanations for cooperation in the one-shot and repeated prisoner's dilemma game: reciprocal behavior in one-shot games and reputation effects in repeated games.

Let us also look at the Matching Pennies games reported in Goeree and Holt. Table 3 provides the general structure of the game.

Table 3

		Player 2: Left	Player 2: Right
Player 1	Top	40 x	80 40
	Bottom	80 40	40 80

The treatments they report consist of a symmetric game with x = 80 and two asymmetric games with x = 320 and x = 44. For all versions of this game, there is a unique mixed strategy Nash equilibrium. In the symmetric version, both players mix with equal probabilities. For the asymmetric games, Player 1 should mix with equal probabilities, while player 2 should choose Left with probability 0.125 when x = 320 and 0.909 when x = 44.

For x = 80, the results they report are consistent with the mixed strategy Nash prediction. Player 1 chooses Top 48% of the time. Player 2 chooses Left 48% of the time.

For x = 320, the results for Player 2 choosing Left are not damning, 16% versus 12.5%, but the results for Player 1 choosing Top are miserable, 96% versus 50%.

For x = 44, the results for Player 2 choosing Left are not terrible, 80% versus 90.9%, but for Player 1 choosing Top, the results are again miserable, 8% versus 50%.

How does the QRE perform? Figure 2 shows the QRE probabilities for Top as λ increases from 0. Figure 3 shows the QRE probabilities for Left as λ increases. When x = 80 in the symmetric game, the QRE probability for Top and Left is 0.5 regardless of λ, which supports the results above. When x = 320, the QRE probability for Top initially increases with λ, but then declines approaching the Nash prediction. For Left, the QRE probability gradually decreases and approaches the Nash. When x = 44, the QRE probability for Top initially decreases, but then begins to increase. For Left, it steadily increases.

For low to moderate values of λ, the QRE does better than the Nash for Player 2 because it predicts Left slightly more often when x = 320 and slightly less often when x = 44. Both these results are consistent with observation. For Player 1, compared to the Nash equilibrium, the QRE predicts Top more often for x = 320 and less often for x = 44, which is again consistent with observation. It should also be noted however that regardless of λ the QRE prediction for x = 320 falls short of the 96% of observed Top play. For x = 44, the QRE never comes close to the 8% observed Top play regardless of λ. Therefore, while it is fair to say the QRE does better than the Nash, there still maybe something else going on here.

I mentioned previously that the QRE was well suited for empirical analysis. Let us see how. Suppose we have N pairs of subjects play the game in Table 2. Denote the ith pair of subjects strategy choices as y_i where $y_i \in$ {(Top, Left), (Top, Right), (Bottom, Left), (Bottom, Right)}. The probability of y_i can be written as

(3) $$\Pr(y_i) = \begin{cases} p^1 p^2, & for\ y_i = (Top, Left) \\ p^1(1-p^2), & for\ y_i = (Top, Right) \\ (1-p^1)p^2, & for\ y_i = (Bottom, Left) \\ (1-p^1)(1-p^2), & for\ y_i = (Bottom, Right) \end{cases}.$$

Let **y** = $\{y_1,\ldots,y_N\}$, so the probability of **y** is

(4) $$L = \Pi_{i=1}^{N} \Pr(y_i).$$

Using equations (1) and (2) we can solve for l: $\pi^1(l)$ and $\pi^2(l)$. Substituting this solution into equation (3) makes equation (4) dependent only on λ. We can identify the $\lambda \geq 0$ that maximizes the likelihood. Of course, all this has to be done numerically since $\pi^1(l)$ and $\pi^2(l)$ do not have a closed form solution.

McKelvey and Palfrey (1995) do this for a variety of games, in order to explore a variety of questions. For example, they compare the estimated probabilities for each strategy to the implied Nash probabilities. They explore if play is simply random by testing if $\lambda = 0$. They check for learning by comparing parameter estimates for λ as subjects gain experience playing the game. Presumably, as subjects learn, the dispersion of error will decrease. So, estimates of λ should be higher for later periods of an experiment that has subjects play the same game repeatedly. They reject Nash play and random play and find mixed support for learning.

Figure 1: QRE for Prisoner's Dilemma

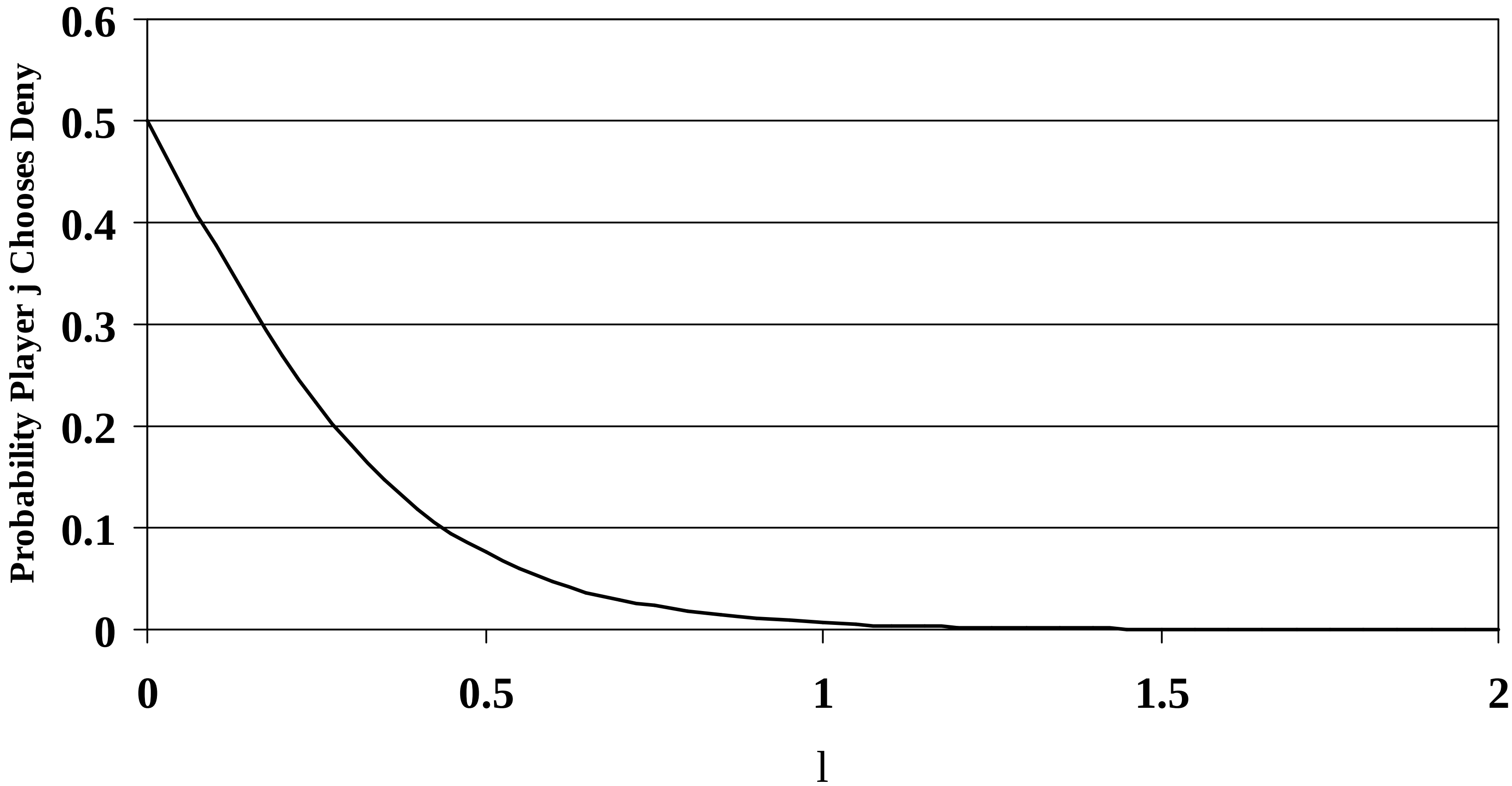

Figure 2: Player 1's QRE for Matching Pennies

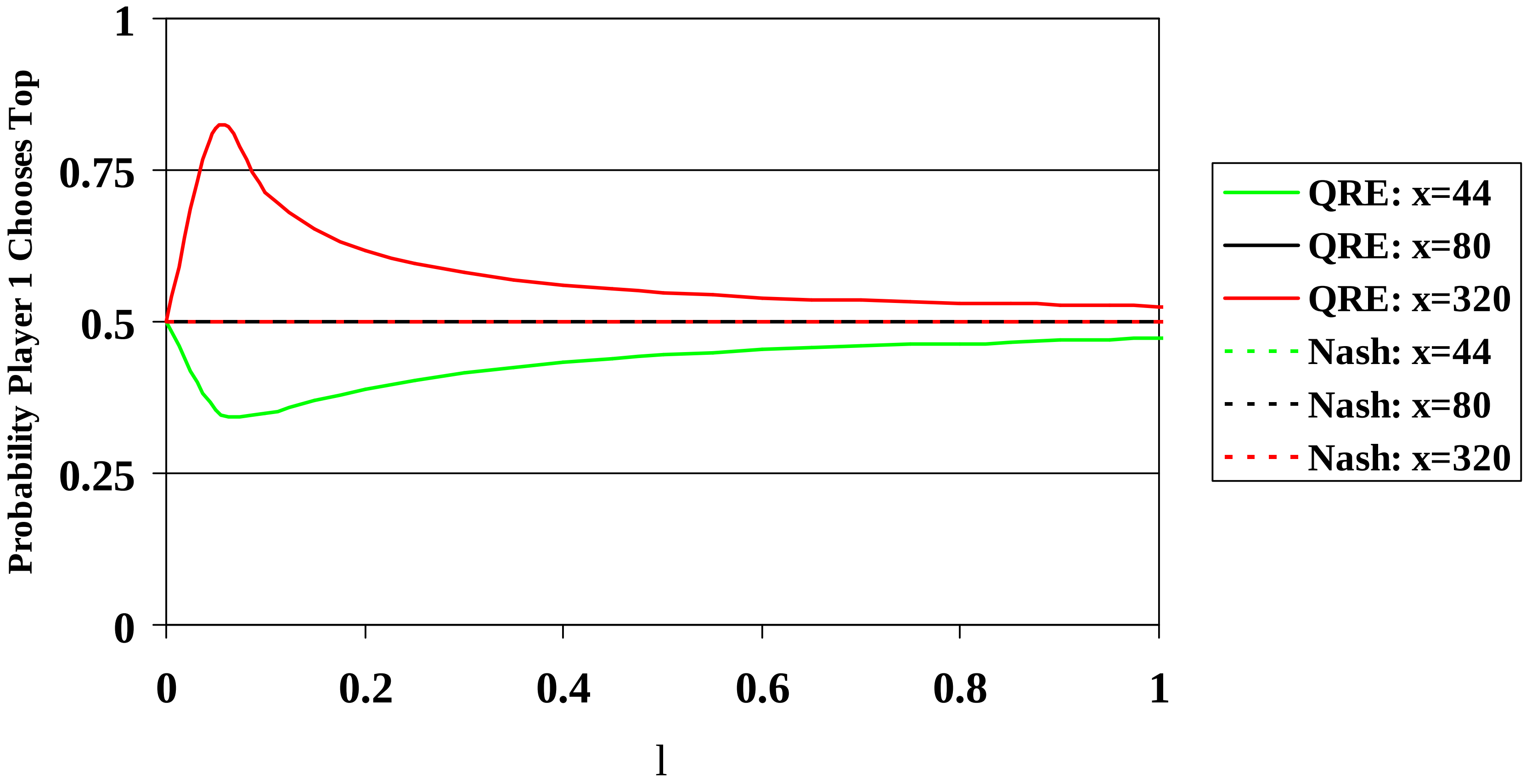

Figure 3: Player 2's QRE and Nash for Matching Pennies