Quantal Response Equilibrium with Non-Monotone Probabilities: A Dynamic Approach

Suren Basov[1]

Department of Economics, University of Melbourne

Abstract In this paper I will give an example of a population game and of a locally improving stochastic learning process such that the quantal response equilibrium assigns to the different strategies the probabilities that are non-monotone in the payoffs. Moreover, if the initial state probabilities are payoff-monotone, then the learning can be shown the shrink mistakes in one direction and exacerbate them in the other direction.

Keywords: quantal response equilibrium, dynamic probabilistic choice models

JEL classification numbers: C7

---

[1]Suren Basov, Department of Economics, The University of Melbourne, Victoria 3010, Australia.

e-mail: s.basov@cupid.ecom.unimelb.edu.au Tel. (61 3)-8344-7154, Fax (61 3)-8344-6899

1 INTRODUCTION

Past decades have witnessed growing empirical evidence that calls into question the utility maximization paradigm. For a description of systematic errors made by experimental subjects, see Arkes and Hammond (1986), Hogarth (1980), Kahneman, Slovic, and Tversky (1982), and Nisbett and Ross (1980), and the survey papers by Payne, Bettman, and Johnson (1982), and by Pitz and Sachs (1984).

To capture this type of behavior, Luce (1959) introduced a probabilistic choice model, currently known as the *logit* model. This and similar probabilistic choice models have already found their application in economics. See, for example, McKelvey and Palfrey (1995, 1998), Chen, Friedman, Thisse (1997), , Offerman, Schram, Sonnemans (1998), and Anderson, Goeree, and Halt (1998, 2001).

It is typical in this kind of models to impose some intuitive restrictions on the probability density of choices and study the distributions that satisfy such restrictions. One of the most natural assumptions to make is payoff monotonicity: the choice probabilities are increasing in the payoffs. McKelvey and Palfrey (1995) applied the idea of probabilistic choice to develop a new equilibrium concept: *quantal response equilibrium*. Assumptions made

2

in McKelvey and Palfrey (1995) on the decision process guarantee that the payoff monotonicity will be satisfied for the equilibrium choice probabilities. However, the interpretation of McKelvey and Palfrey (1995) of quantal response equilibrium requires that the players should be able to calculate their expected payoff, possibly with some mistake. The same is true about all other papers cited in the previous paragraph.

One might argue that the ability of the players to calculate their expected payoff does not square well with the assumption of bounded rationality, which motivated introduction of probabilistic choice models in the first place.[2] The problem arises because the approach taken in these papers is *static*. An alternative would be to require that the individuals adjust their choices based on some simple rule. The adjustment rule will give rise to a stochastic process on the choice set. Then the probability distribution of choices of the static approach will correspond to the steady state of this stochastic process. I call this approach *dynamic*. Unfortunately, the assumptions usually imposed on the probability density of choices need not be satisfied by the steady state density of the dynamic approach even for reasonable adjustment rules.

---

[2]McKelvey and Palfrey (1995) discuss an alternative view that the players are fully rational but have an additive payoff disturbance associated with each pure strategy. Most authors, however, take a view that quantal response is associated with bounded rationality.

In this paper I take a dynamic point of view and give an example of a population game and of a locally improving stochastic learning process such that the quantal response equilibrium assigns to the different strategies the probabilities that are non-monotone in the payoffs. Moreover, if the initial state probabilities are payoff-monotone, then the learning can be shown the shrink mistakes in one direction and exacerbate them in the other direction. Basov (2002) argues that the situation described in this example is generic.

2   A POPULATION GAME

Assume a population consists of a continuum of individuals. Each individual selects a vector $(x, y) \in R^2$. The payoff to the individual is given by

$$u(x, y) = -\frac{1}{2}(x - \alpha X)^2 - \frac{1}{2}(y - \beta Y)^2, \tag{1}$$

where $\alpha \neq 1$, $\beta \neq 1$, and X and Y are population means of x and y.

Nash Equilibrium: It is straightforward to find the unique Nash equilib-

rium in this game. Indeed, the best response of an individual is given by

$$
\begin{cases}
x = \alpha X \\
y = \beta Y
\end{cases}.
\tag{2}
$$

Taking expectations one concludes that $X = Y = 0$ and (2) implies that $x = y = 0$ in the Nash equilibrium.

Logit Equilibrium: The logit equilibrium is a version of quantal response equilibrium when deviations of the perceived payoffs from the realized ones has the extreme value distribution. The population density of choices is given by

$$
f(x, y) = \frac{1}{\pi \sigma^2} \exp(-\frac{(x - \alpha X)^2 + (y - \beta Y)^2}{\sigma^2}).
\tag{3}
$$

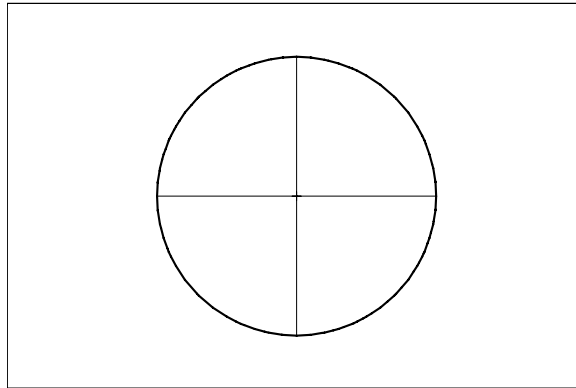Using it to calculate the population means one gets

$$
\begin{cases}
X = \alpha X \\
Y = \beta Y
\end{cases}.
\tag{4}
$$

Therefore, $X = Y = 0$ and the equilibrium choice density is

$$
f(x, y) = \frac{1}{\pi \sigma^2} \exp(-\frac{x^2 + y^2}{\sigma^2}).
\tag{5}
$$

5

Note that the logit equilibrium choice is unbiased with respect to the Nash equilibrium choice and the choice probability is an increasing function of the payoff. The equiprobability curves, i.e. the curves along which the probability density of choices is constant, are circumferences with the center at the origin.

Figure 1

An equiprobability curve for logit equilibrium.

3   A DYNAMIC MODEL

In the previous Section I described a population game and found its Nash and Logit equilibria. Note that both equilibrium notions used in the previous

Section are static, since they do not rely on any explicit learning dynamics. In this Section I am going to assume instead that the players adjust their choices gradually. The expected adjustment vector will be equal to the gradient of their utility. I will also assume that the adjustment is subject to a random mistake. I will call the steady state of this model a *dynamic quantal response equilibrium*. Formally, assume that

$$
\begin{cases}
dx = -(x - \alpha X)dt + \sigma_1 dW_1 \\
dy = -(y - \beta Y)dt + \sigma_2 dW_2
\end{cases}
, \tag{6}
$$

where W_1 and W_2 are independent standard Wiener processes. Then the probability density of choices satisfies the following partial differential equation (Ito, 1992)

$$
\frac{\partial f}{\partial t} - \frac{\partial}{\partial x}((x - \alpha X)f) - \frac{\partial}{\partial y}((y - \alpha Y)f) = \frac{\sigma_1^2}{2}\frac{\partial^2 f}{\partial x^2} + \frac{\sigma_2^2}{2}\frac{\partial^2 f}{\partial y^2}. \tag{7}
$$

The stationary solution of the equation (7) is given by

$$
f(x, y) = \frac{1}{\pi \sigma_1 \sigma_2} \exp(-\frac{(x - \alpha X)^2}{\sigma_1^2} - \frac{(y - \beta Y)^2}{\sigma_2^2}). \tag{8}
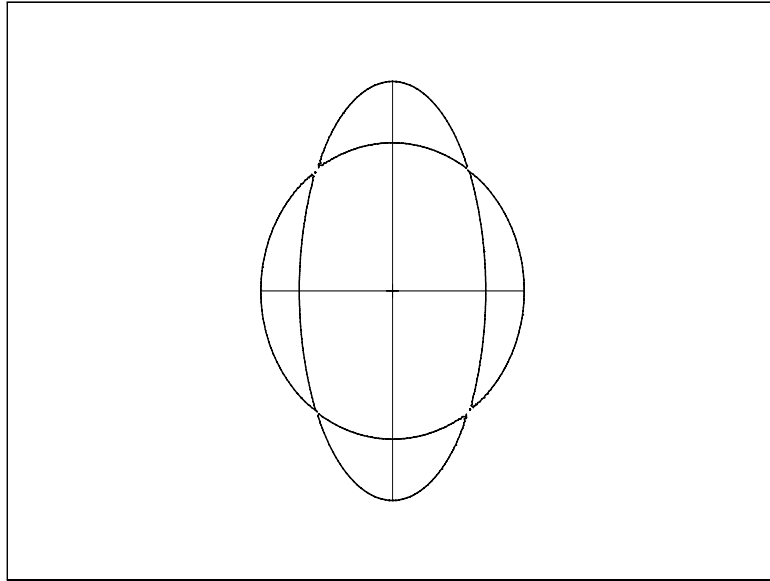$$

Using it to calculate the population means one gets

$$\begin{cases} \overset{\cdot}{X} = \alpha X \\ \overset{\cdot}{Y} = \beta Y \end{cases}. \tag{9}$$

Therefore, $X = Y = 0$ and the equilibrium choice density is

$$f(x, y) = \frac{1}{\pi \sigma_1 \sigma_2} \exp(-\frac{x^2}{\sigma_1^2} - \frac{y^2}{\sigma_2^2}). \tag{10}$$

Note that the dynamic quantal response equilibrium choice is unbiased with respect to the Nash equilibrium choice. It is, however, not monotone in payoffs, unless $\sigma_1 = \sigma_2$. To see this geometrically, note the the equiprobability curves are now ellipses. Figure 2 depicts an equiprobability and an indifference curve corresponding to the equilibrium of the model.

Figure 2

An indifference curve (circumference) and an equiprobability curve (ellipse)

in the dynamic quantal response equilibrium.

Let point A belong to the part of the ellipse above the circumference, while point B to the part of the circle left of the ellipse, then $u(A) > u(B)$ (since A is above the indifference curve, while B is below it), while $f(A) < f(B)$ (since A is below the equiprobability curve, while B is above it).

The above example demonstrates that the dynamic quantal response equilibrium can result in the steady state choice probabilities that are non-monotone on payoffs. Moreover, suppose that the initial density of choices is

given by

$$f(x, y; t = 0) = \frac{1}{\pi\sigma^2} \exp(-\frac{x^2 + y^2}{\sigma^2}). \qquad (11)$$

Note that the initial choices are unbiased with respect to the Nash equilibrium choices and their density is payoff monotone. Now assume that the players adjust their choices according to rule (7) with σ_1 and σ_2 such that $\sigma_1 < \sigma_2$ and $\sigma_1^2 + \sigma_2^2 = 2\sigma^2$. Then the equilibrium steady state density will be given by (10). Our assumptions on σ_1 and σ_2 insure that the total variability of the choices in the dynamic quantal response equilibrium remains the same as in the initial state:

$$Var(x; t = \infty) + Var(y; t = \infty) = Var(x; t = 0) + Var(y; t = 0). \qquad (12)$$

However,

$$Var(x; t = \infty) < Var(x; t = 0), \ Var(y; t = \infty) > Var(y; t = 0). \qquad (13)$$

Therefore, learning shrinks mistakes in one direction, while exacerbating them in the other. If one interprets stochastic terms in (6) as experimentation, this result will imply that learning will exacerbate mistake in direction in

which players experiment too much. If one assumes furthermore that people tend to experiment more aggressively in new unfamiliar environments, this results will imply that learning can lower payoffs exactly in the situations when there is a lot to learn.

4   CONCLUSIONS

In this paper I have shown by an example that the probability density of a strategy in a dynamic quantal response equilibrium need not be monotone in payoffs, even for a reasonable (locally improving) adjustment rule. In a recent paper Basov (2002) argued that this situation is generic. Moreover, for some initial conditions learning can lead to diminishing of mistakes in one direction and exacerbation in the other. This calls for a critical re-examination of the conclusions obtained from static equilibrium models and calls for an explicit modelling of a dynamic adjustment process

REFERENCES

Anderson S. P., Goeree, J. K., Holt C. A, 1998, Rent seeking with bounded rationality: An analysis of all-pay auction. Journal of Political Economy, 106, 828-853.

Anderson S. P., Goeree, J. K., Holt C. A., 2001, Minimum-effort coordination games: stochastic potential and logit equilibrium. Games and Economic Behavior, 34, 177-199.

Arkes, H. R., Hammond, K. R., 1986, Judgment and decision making: an interdisciplinary reader. Cambridge: Cambridge U. Press.

Basov, S., 2002, Bounded rationality: Static versus Dynamic Approaches, The University of Melbourne, Department of Economics Working Paper 864.

Chen, H. C., Friedman, J. W., Thisse, J. F., 1997, Boundedly rational Nash equilibrium: a probabilistic choice approach, Games and Economic Behavior, 18, 32-54.

Hogarth, R.: Judgment and choice: Psychology of decision, 1980, New York: Wiley.

Ito, S., 1992, Diffusion equation. American Mathematical Society.

Kahneman, D., Slovic P., Tversky A., 1982, Judgment under uncertainty:

heuristic and biases. Cambridge: Cambridge U. Press.

Luce, R. D., 1959, Individual choice behavior, New York: Wiley.

McKelvey, R. D., Palfrey, T. R., 1995, Quantal response equilibria for normal form games. Games and Economic Behavior, 10, 6-38.

McKelvey, R. D., Palfrey, T. R., 1998, Quantal response equilibria for extensive form games. Experimental Economics, 1, 9-41.

Nisbett, R., Ross L., 1980, Human inference: Strategies and shortcomings in the social judgment. Englewood Cliffs: Prentice-Hall.

Offerman, T., Schram, A., Sonnemans, J., 1998, Quantal response models in step-level public good games. European Journal of Political Economy, 14, 89-100.

Payne, J. W., Bettman, J. R., Johnson E. J., 1992, Behavioral decision research: a constructive processing perspective. Annual Review of Psychology **43**, 87-131.

Pitz, G., Sachs N. J.: Judgment and decision, 1984, Theory and application. Annual Review of Psychology, 35,139-63.