A note on ecological rationality, Bayesian learning and rational expectations

Francisco M. Torralba

1 Introduction

Economic theory is mostly based on an unrealistic view of decision making. Agents are assumed to be fully rational maximizers with vast computational capabilities and resources. This view has *not* been inspired by empirical evidence. It is rather based on a theoretical apparatus which is appealing because of its predictive power, its mathematical tractability and its elegance. Once this apparatus is abandoned, the modeler faces so many choices with regard to how much knowledge and what kind of cognitive abilities to attribute to agents in his models.

In this note I will consider a model of bounded rationality: a simple rule based on frugal observation and well adapted to the environment where it must be used. Bounded rationality, as defined by Gigerenzer (2002), Gigerenzer and Selten (2002) and Selten (2002), is not a general theory of decision making. Instead, models of bounded rationality consist of simple rules that function well under the particular state of available information and cognitive ability available to the user of those rules. Hence, every agent carries a box of simple tools (rules), each of them adapted to a particular problem, rather than an all-purpose sophisticated computer (a general approach to decision making).

In the environment I have created, based to some extent in Ellison and Fudenberg (1993) , agents must choose between two alternatives with one period of anticipation. The payoff to one of the alternatives is uncertain at the time they make their choice, but it is related only to the payoff of the same alternative at that time, which they can observe. Several possibilities can be considered with regard to the way One option for the modeler is to make the agents expected-payoff maximizers and provide them with the machinery (knowledge of the structure of the uncertainty and of the relevant probabilities, as well as certain cognitive abilities) to calculate expected payoffs. The outcome of such a decision process would be rational, in the sense of Sargent (1993). Maintaining the assumption that agents are expected-payoff maximizers, another option for the modeler is to make agents solve their problem under the constraint that they do not know the relevant probabilities, but they know how to calculate them. The problem is that even if the computations involved were simple (and in other environments they are not), they require vast resources, in terms of memory and time required to calculate accurate estimates. Yet another option is to provide the agents with fast and frugal decision rules that do not involve optimization. If these rules are well adapted to the environment, they will lead to outcomes that mimic those obtained under the rational paradigm. My decision rule is based, at least in part, on the experience of other agents, a process that has been called social learning.

In this note I do not attempt to investigate the process by which heuristic rules are adopted, dropped and modified. Instead, I take a decision problem to be solved in a specfic environment, and compare the outcomes provided by fully rational agents, Bayesian learners and boundedly rational agents who follow a heuristic rule. I will argue that, for a broad range of scenarios, simple heuristics lead to outcomes that are observationally equivalent to those generated by fully rational decision rules. Moreover, I show that a simple heuristic rule can beat Bayesian learning with regard to how close its outcome is to the rational outcome, on average.

2 Description of the model

Consider a population of identical workers, who are initially distributed among two sectors: A and B. I will denote the fraction of workers in sector A at time t by x_t. (The fraction of the population of workers in sector B will then be $1-x_t$). Each period a fraction α of the workers is allowed to switch from one sector to the other at no cost [1]. The workers who are allowed to switch observe the current payoff of sector A, y_t, and the current proportion of workers in sector A, x_t. If they choose sector A, they will receive the wage y_{t+1} starting in period $t+1$ and until next time they are allowed to switch (therefore their payoff is uncertain at the time they have to choose a sector but constant thereafter). For all t, y_{t+1} is governed by an n-state, first-order Markov process with transition probability matrix Π and takes values from the set $Y = \{y_{1,...,}y_n\}$. If the worker chooses sector B she receives the wage $w_{t+1}(x_t) = \exp(\gamma(x_t - b))$, starting in period $t+1$ and until next time she is allowed to switch (so the payoff in B varies with time but it is not uncertain at the time a worker chooses the sector; moreover, as in sector A, the worker receives a constant payoff until next time she is summoned to revise her choice of sector). I assume that $\gamma > 0$, reflecting the fact that the level of wage offers in sector B becomes lower as the proportion of workers in that sector increases.

[1] Ellison and Fudenberg (1993) review some empirical literature on inertia in processes of technology adoption.

2.1 Rational expectations: the benchmark environment

As a benchmark environment, suppose that all the workers know the true stochastic process governing y_{t+1}, conditional on y_t, i.e. they have rational expectations as defined by Sargent (1993). Then their decision criterion for the choice of sectors is to choose sector A if

$$E_t y_{t+1} = \sum_{j=1}^{n} \pi_{ij} y_j \geq w_{t+1}(x_t) \tag{1}$$

where I use the conventional notation for Markov processes: π_{ij} is the ijth element of the matrix Π and represents the probability of $y_{t+1} = y_j$, given that $y_t = y_i$.Under this decision rule and given the knowledge of the random process the evolution of the fraction in workers in sector A is

$$x_{t+1} = \begin{cases} (1-\alpha)x_t + \alpha & \text{w.p. } p(x_t) \\ (1-\alpha)x_t & \text{w.p } 1 - p(x_t) \end{cases} \tag{2}$$

where p, the probability of an upward movement of x, can be calculated as

$$p(x_t) = \sum_{i=1}^{n} \pi_i \left\{ \Pr\left[\sum_{j=1}^{n} \pi_{ij} y_j \geq w_{t+1}(x_t) \right] \right\} \tag{3}$$

where π_i is the unconditional probability of visiting state i.

The probability of visiting each of the three states is given by the stationary unconditional probability distribution, π, which is the solution to $\pi' = \pi'\Pi$, or $(I - \Pi')\pi = 0$. Thus the unconditional probability of visiting state i is the ith element of the eigenvector (normalized to satisfy $\sum_i \pi_i = 1$) associated with a unit eigenvalue of Π'. The fact that Π has nonnegative elements and that the sum of the elements in any row is 1 guarantees the existence of at least one unit eigenvalue. Moreover, if all the elements in Π are strictly positive there exists a unique unit eigenvalue. (See Ljungqvist and Sargent (2000)).

The average value of the sequence $\{x_t\}_{t=1}^{\infty}$, $\bar{x}$, is determined by the following equation

$$\bar{x} = p(\bar{x}) \left[(1-\alpha)\bar{x} + \alpha \right] + (1 - p(\bar{x})) \left[(1-\alpha)\bar{x} \right]$$

or

$$\bar{x} = p(\bar{x}) \tag{4}$$

Proposition 1 *There exists a solution to (4). Moreover, if all the elements in* Π *are strictly positive the solution is unique.*

Proof. For a given x_t, $\Pr\left[\sum_{j=1}^{n} \pi_{ij} y_j \geq w_{t+1}(x_t)\right]$ only takes the values 0 or 1. Moreover, by construction of the unconditional probability distribution, $\pi_i \in [0,1]$, for all i. Therefore, $p(x_t) \in [0,1]$, for all x_t. On the other hand, since $w_{t+1}(x_t)$ is a strictly increasing function of x_t, and since the elements of Π do not depend on x_t, $p(x_t)$ is a decreasing function of x_t. By the mean value theorem, for a given unconditional probability distribution, there exists a unique $\bar{x}$ such that equation (4) holds. Finally, if all the elements in Π are strictly positive there exists a unique unit eigenvalue and hence a unique unconditional probability distribution. ■

Before finishing this section, let us remark that $p(x_t)$ will be a step function, each step being determined by the level of x_t such that $\sum_{j=1}^{n} \pi_{ij} y_j = w_{t+1}(x_t)$, for some i. Then graphically, in the positive orthant of the $(x, p(x))$-plane, the steady state corresponds to the intersection of that decreasing step function with the 45-degree line as drawn in Figure 1, in the Appendix.

2.2 Bayesian learning: a slow convergence to the rational expectations steady state

Suppose now that, instead of knowing the true probability of each state of the world, the workers have to learn about them. In particular, each worker will count and remember the number of times he has observed a transition from any of the n states of the world to any of the n states, which she observes only whenever she is selected at random to revise her choice of sector, as in the previous subsection. For each state $i = 1, 2, ..., n$, I let $\tau_t^r(ij)$ denote the cumulative number of times that worker r has observed a transition from state i to state j at time t. Each time a worker is allowed to switch from one sector to the other, she will use her estimated transition matrix for the Markov process, whose ijth element will be $\hat{\pi}_{ij,t}^r \equiv \tau_t^r(ij) / \sum_i \sum_j \tau_t^r(ij)$, to replace π_{ij} in equation (1). Then she will choose a sector using that decision rule, observe y_{t+1} and update her estimated transition matrix. (I assume that workers observe y_{t+1} no matter whether they choose sector A or B).

This learning process is sure to make the transition probabilities estimated *by each worker* converge to the true transition probalities, and thus in the limit the steady state of x under Bayesian learning will converge to the rational expectations steady state. However, that convergence can be extremely slow for two reasons. First of all, for a finite duration of T periods of the process, a worker observes a transition of y an average of just αT times. Then a transition from state i to state j will be observed by a worker an average of $(\alpha T)\pi_i \pi_{ij}$ times. Given that α, π_i and π_{ij} are all between 0 and 1, that number can be quite small unless workers are allowed to "live" for a very large number of periods. Slow learning, consistent as it may be, means no learning at all for short observed histories of the stochastic variables that one wants to learn about.

2.3 Social learning and simple heuristics

Consider now a learning process that incorporates a limited history of y and the current popularity of sector A as a source of information. In particular, workers will be assumed to construct an indicator that combines the proportion of workers in sector A and the difference between the current payoff of sector A and the payoff they will receive in sector B if they choose that sector. A worker chooses sector A if

$$m(2x_t - 1) + (1 - m)(y_t - w_{t+1}(x_t)) \geq 0 \quad (5)$$

Under this rule m is the amount of popularity weighting. Notice that when $x_t = 1/2$ both occupations are equally popular and popularity conveys no information. In that case workers choose the sector with the highest payoff at time t (which may not coincide with the sector with they highest payoff at the time they start working, i.e. $t + 1$). The case $m = 0$ corresponds to a simple heuristic rule that leads to choose whichever sector that offers the highest payoff, regardless of the expected payoff during the time that the worker will spend there; $m = 1$ corresponds to the case where only popularity matters. The introduction of popularity in the decision rule may cause the evolution of x_t to be characterized by herding, specially for high values of m, for which the excess payoff may become irrelevant to determine the sign of the left hand side of expression (4).

3 Some experiments

In this section I present Matlab simulations of the evolution of x_t for particular values of the parameters of this model under each of the decision rules. For each of the three decision rules, $n = 3$, $y_1 = -1$, $y_2 = 1$, $y_3 = 3$, $\gamma = 1/3$, $b = 2/3$ and $\alpha = 0.01$. The true probability transition matrix is $\begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.4 & 0.3 & 0.3 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$.

Under rational expectations, the expected values of y_{t+1}, conditional on each of the three possible realizations of the payoff in sector A, are $E[y_{t+1}|y_t = -1] = -0.2$, $E[y_{t+1}|y_t = 1] = 0.8$ and $E[y_{t+1}|y_t = 3] = 1$, respectively. Workers will choose sector B over A whenever they observe $y_t = -1$ since, by construction, the payoff in sector B cannot be negative. Whenever any of the other two states of the world is observed, the choice of sector will depend on the value of x_t.

For the example at hand, the stationary unconditional distribution is $[\pi_1\ \pi_2\ \pi_3] \simeq [0.4685\ 0.2785\ 0.2532]$. To construct $p(x)$ we need to determine the states for which the expected payoff in sector A exceeds the promised payoff in sector B, and add up the stationary unconditional probabilities of all such states, for each value of x. Since $w_{t+1} = \exp(1/3(x_t - 2/3))$ is positive and strictly increasing for $x_t \in [0, 1]$, we just need to find two cutoff values of x, which are the solutions to

$$w = \exp(1/3(x - 2/3)) = 1 = E[y_{t+1}|y_t = 3]$$

and

$$w = \exp(1/3(x - 2/3)) = 0.8 = E[y_{t+1}|y_t = 1]$$

respectively. The first cutoff value is $x^* = 2/3$ and the second one is $x^{**} = -0.003$. For all $x > x^*$, the payoff in sector B exceeds that in sector A and thus $p(x) = 0$; for $x^{**} < x \leq x^*$, the payoff in sector A (weakly) exceeds that in sector B and thus $p(x) = \pi_3$. (The case where $x \leq x^{**}$ is irrelevant here since x^{**} is negative). Hence the only steady state value of the process is $\bar{x} = \pi_3 = 0.2532$.

For the chosen values for the parameters of the problem, workers join sector A only if they observe the realization of state 3 at the time they are summoned to update their choice of sector. For any other realization of y, the proportion of workers in sector A will drop. A realization of a history of x under rational expectations is presented in Figure 2.

To simulate a history of x under Bayesian learning it is necessary to provide the workers with some initial estimate of the probability transition matrix. (Remember that they will estimate that matrix using only their own experience). I used a population of 100 workers and assumed that only one of them can revise their choice of sector at a time (this corresponds to $\alpha = 0.01$, as in the rational expectations simulation). I provided all the workers with a history of length 10 of y and let them estimate transition probabilities from there. The workers will use these estimates the first time they are allowed to revise their choice of sector and they will update them thereafter. All the workers were given the same initial history, but the updating of the estimated probabilities was made individual by individual, based on their experience. A realization of a history of x under Bayesian learning and rational expectations, for the same realized history of y, is presented in Figure 3.

A numerical analysis of equation (5) yielded the result that, for $x_t \leq 0.3947$, a social weight $m \in [0.3, 0.69]$ produces a path of x that is observationally equivalent to one generated by a population of agents with rational expectations. Figure 4 depicts the histories generated in each of the three cases presented, for a social weight $m = 0.4$. Notice that only two lines can be observed, because the green line (corresponding to the evolution of x when agents use simple heuristics) is superimposed on the blue line (corresponding to the rational expectations case).

For m outside that interval the paths with rational expectations and simple heuristics do not coincide because, for at least one state of the world, the decision rules for each of those cases yield differing choices of sector. Figures 5 and 6 depict the histories of x for $m = 0.2$ and $m = 0.8$, respectively.

For $m = 0.2$, equation (5) is satisfied for $x_t > 0.21$, whereas it is never satisfied under rational expectations. Therefore, when agents use simple heuristics, they join sector A whenever they observe states 2 or 3, which yields an expected value of x_t of 0.5317. For $m = 0.8$, equation (5) is never satisfied if x_t drops below 0.24 and, consequently a worker always chooses sector B when that lower bound is crossed, leading causes x_t to "escape" to 0.

4 Concluding comments

In the paper that might follow this note I would like to take on the problem of defining a general class of heuristic rules, perhaps based on psychological evidence of human behavior. Next I would analyze the adaptation of rules to the environment where they have to be used. Such an analysis will be carried out in the context of a specific problem, perhaps a version of the problem considered in this note.

References

[1] Ellison, G. and Fudenberg, D. (1993) "Rules of Thumb for Social Learning", Journal of Political Economy 101(4), pp. 612-642.

[2] Gigerenzer, G. (2002) "The Adaptive Toolbox" in: Gigerenzer, G. and Selten, R. eds. Bounded Rationality: the Adaptive Toolbox. Cambridge, MA: The MIT Press.

[3] Gigerenzer, G. and Selten, R. (2002) "Rethinking Rationality" in: Gigerenzer, G. and Selten, R. eds. Bounded Rationality: the Adaptive Toolbox. Cambridge, MA: The MIT Press.

[4] Ljungqvist, L. and Sargent, T.J. (2000) Recusive Macroeconomic Theory. Cambridge, MA: The MIT Press, pp. 3-5.

[5] Sargent, T. J. (1995) Bounded Rationality in Macroeconomics. Oxford: Clarendon Press.

[6] Selten, R. (2002) "What is Bounded Rationality" in: Gigerenzer, G. and Selten, R. eds. Bounded Rationality: the Adaptive Toolbox. Cambridge, MA: The MIT Press.