Bounded Rationality in Macroeconomics: Comments

Amaresh Das

College of business, Southern University at New Orleans, New Orleans, LA 70056, USA

Abstract: The paper is a journey into the dilemmas faced by economists attempting to justify and expand the scope of the theory in the direction of bounded rationality. The Robbins-Monro Algorithm that introduces stochastic approximation could be applied to more rational decisions making. Some of the methodological issues regarding the construction of new models on hypothesis other than that of rationality are discussed.

Key words: Bounded rationality • Stochastic approximation • Cumulative pay-off matching

INTRODUCTION

To judge from what gets published, the rational expectations revolution in macroeconomics has been an unqualified success. Even those who call themselves new Keynesians now assume rational expectations as an uncontrolled matter of routine. But, the limitations of this new conventional assumption have not been unnoticed among more thoughtful macroeconomists, including the former leaders of the revolution who have long pointed out that rational expectations makes sense only in recurrent situations where people have had a chance to learn about and adjust to their environment. Furthermore, it is also widely recognized now that multiplicity of rational expectations equilibria is a pervasive phenomenon and in such situations rational expectations may be of little use in determining what will happen even if environment is stable. Aside from problems of transition and multiplicity, assuming rational expectations amounts to assuming away the very coordination problems that once defined the discipline of macroeconomics, because it presupposes that people have adapted fully and costlessly not just to exogenous factors their environment but also to each other.

When rationality fails, there is a mismatch between the decision-making environment and the choices of the decision maker. The macroeconomists refer to this mismatch as “bounded rationality showing through” (Simon 1998). This conception has an important implication. In structured situations, at least, we may conceive of any decision as having two components: environmental demands (seen by the individual as incentives, positive or negative) and bounds on adaptability in the given decision-making situation. Ideally, an analysis based on rational choice should be able to specify what the incentives are and to predict decisions based on those incentives. What cannot be explained is either random error (even the most rational of us may make an occasional mistake, but these are not systematic) or bounded rationality showing through. Standard statistical techniques give us the tools to distinguish systematic from random factors, so in principle it should be possible to distinguish the rational, adaptive portion of a decision from bounds on rationality.

Bounded rationality is a school of thought about decision making that developed from dissatisfaction with the “comprehensively rational” economic and decision theory models of choice. Those models assume that preferences are defined over outcomes that those outcomes are known and fixed and that decision makers maximize their net benefits, or utilities, by choosing the alternative that yields the highest level of benefits (discounted by costs) [1,2]. These considerations have inspired an emergent literature in macroeconomics that replaces rational expectations with a more adaptive or evolutionary or even a ‘complex’ approach to the formation of beliefs.

Obviously, such an approach may ask for a rather limiting perspective for the analysis of long-run economic dynamics. What matters for the long run are inherently uncertain phenomena of which radical technological changes are major ingredients [3-5]. The direction of technological change and the outcome of the R&D process are impossible to predict in an accurate way and
...hence the perfect foresight or rational expectations assumptions associated with full rationality models are hard to maintain. Just how difficult it is to think of fully rational agents in connection to long-run technological trends becomes clear from a quotation from Katz and Philips provided by Lus and Soete [6]. “The general view prior to 1950 was that there was no commercial demand for computers. Thomas J. Watson Senior [the executive of IBM] felt that the one SSEC machine which was on display at IBM’s New York offices ‘could solve all the scientific problems in the world involving scientific calculations’. He saw no commercial possibilities”. In 2001, it would of course be easy to ridicule this prediction [7-9].

Although one cannot prove the point with anecdotes as the above one, it is clear that ‘bounded rationality’ in the sense of Simon [10] is a much more realistic concept for the basis of long-run economic models than the representative and fully rational agent. Although most political scientists are aware of Simon’s contributions, many fail to appreciate that bounded rationality was the first, and because of its ripple effects in some disciplines, the most important idea (even academic school of thought) that political science has ever exported A brief retelling of the tale is in order. As an undergraduate at the University of Chicago, Simon returned to his native Milwaukee in 1935 to observe budgeting in the city’s recreation department. He wrote:

I came as a gift-bearing Greek, fresh from an intermediate price theory course taught by the grandfather of Chicago-School neoclassical laissez-faire economics, Henry Simons. My economics training showed me how to budget rationally. Simply compare the marginal utility of a proposed expenditure with its marginal cost and approve it only if the utility exceeds the cost. However, what I saw in Milwaukee didn’t seem to be an application of this rule. I saw a lot of bargaining, of reference back to last year’s budget and incremental changes in it. If the word “marginal” was ever spoken, I missed it. Moreover, which participants would support which items was quite predictable. I could see a clear connection between people’s positions on budget matters and the values and beliefs that prevailed in their sub organizations.

I brought back to my friends and teachers in economics two gifts, which I ultimately called “organizational identification” and “bounded rationality.” (Simon [10]).

This has given rise to a class of economic models that is often referred to as ‘evolutionary’, or ‘neo-Schumpeterian’ (e.g., Nelson and Winter, [8]). Key concepts in these models are indeed bounded rationality, heterogeneity (of agents’ technological capabilities) and disequilibrium dynamics. Contrary to the full rationality assumption in neoclassical models, this approach leaves room for different mental models and thus, different actions and outcomes between agents dealing with a similar environment and similar goals.

---

1. The subjective expected-utility variant of rational choice integrates risk and uncertainty into the model by associating a probability distribution, estimated by the decision maker, with outcomes. The decision maker maximizes expected utility. Choices among competing goals are handled by indifference curves—generally postulated to be smooth (twice differentiable)—that specify substitutability among goals.

2. A major implication of the approach is that behavior is determined by the mix of incentives facing the decision-maker. A second implication is that adjustment to these incentives is instantaneous; true maximizers have no learning curves.

3. The argument that technology is largely considered as an exogenous factor also holds for most practical approaches to macroeconomic modeling in the Keynesian tradition. For example, this was exactly the point that Ericsson et al [3] raised against Thirlwall’s [12] concept of balance of payments restricted growth.

4. The evolutionary dynamics have an inherently large level of uncertainty with respect to their outcome. Normally, fully developed theoretical evolutionary models underline a complex interaction between ‘chance and necessity’ for a discussion in the context of evolutionary models of economic growth see Thirlwall [12]. Although the exact mix between random factors (chance) and deterministic factors (necessity) is still a matter of debate between evolutionists, it is clearer than in every evolutionary process, there is a considerable degree of randomness. Contingencies may direct biological evolution into completely unanticipated directions, and this phenomenon may have its counterparts in economic evolution. The topic of evolutionary/learning models deserves a complete and separate series of papers.

5. Our comment paper mostly talks about models in which decision makers deliberate decisions by applying procedures that guide their reasoning about what to do and probably also about how to decide. In contrast, evolutionary models treat agents as automata, merely responding to changing environments, without deliberating about their decisions.

6. The idea can be found in the opening chapter to Wealth of Nations. ‘Men are much more likely to discover easier and readier methods of attaining any object than when the whole attention of their minds is directed towards that single object than when it is dissipated among a variety of things, (Smith, 1776, BK 1, Ch 1, see Spall [11] surveyed in in much the same form to the beginning of the 20th century.

---

1003
principle is that people do not have a deeper understanding of the economy than do econometricians, as most rational expectations theories assume but instead they behave just like econometricians. That is, they theorize, estimate and draw statistical inferences from a limited set of data. Thomas Sargent, one of the leaders of the rational-expectations revolution surveys this emergent literature, (see Sargent [9]. Thomas Sargent goes beyond ‘what I know about these techniques’ by identifying a common formal structure underlying a seemingly diverse literature. The unifying structure can take the form of the Robbins – Munro algorithm (see below) for stochastic approximation. From a computational point of view, decision procedures are encoded in algorithms. That is, the effective rationality of an agent is determined by its computational intelligence. Everything else being equal, an agent that has better algorithms could make "more rational" (more optimal) decisions than one that has poorer heuristics and algorithms.

The Robbins –Monroe Algorithm: The algorithm is a method of solving a broad class of problems in statistical estimation. In this method, each successive estimate is obtained in the form of a correction of the preceding estimate, the correction being based solely on new observations. The principal features responsible for the popularity of stochastic approximation in both theoretical and applied work are the method’s nonparametric character (its applicability when the information available on the object of observation is scanty) and its recursive nature (simplicity of recalculation of the estimate when new observation results are obtained). Stochastic approximation is used in many other applied problems in control theory and in problems of engineering, biology and medicine.\footnote{Aside from introducing the adaptive stochastic approach Spall [11] presents practical implementation guidance, asymptotic theory, and a nontrivial numerical evaluation. Also included is a discussion and numerical analysis comparing the adaptive stochastic approach with the iterate-averaging approach to stochastic approximation.}

Let \( M(x) \) be a given function and \( \alpha \) be a given constant such that the equation

\[
M(x) = \alpha \tag{1}
\]

has a unique root \( x = \theta \). There are many methods for determining the value of \( \theta \) by successive approximation.

With such method we begin by choosing one or more values of \( x_1, x_2, \ldots, x_n \) more or less arbitrarily and then successively obtain new values of \( x_n \) as certain functions of the previously obtained, \( x_1, x_2, \ldots, x_{n-1} \), the values \( M'(x_1), \ldots, M'(x_{n-1}) \) and possibly those of the derivatives \( M''(x_1), \ldots, M''(x_{n-1}) \) etc.

If

\[
\lim_{n \to \infty} x_n = \theta
\]

irrespective of the arbitrarily initial values \( x_1, \ldots, x_n \), then the method is effective for the particular function \( M(x) \) and value of \( M(x) \). The speed of the convergence in (2) and the ease with which the \( x \) can be computed determine the practical utility of the method.

A Stochastic Generalization: We can consider a stochastic generalization of the problem in which the nature of the function \( M(x) \) is unknown to the experimenter. Instead we suppose that to each value \( x \) corresponds a random variable \( y = y(x) \) with the distribution function \( P_r[y(x)] = H(y|x) \) such that

\[
M(x) = \int_{-\infty}^{\infty} y \, dH(y|x) \tag{3}
\]

is the expected value of \( Y \) for the given \( x \). Neither the exact nature of \( H(y|x) \) nor that of \( M(x) \) is known to the experimenter, but, it is assumed that equation (1) has a unique root \( \theta \) by making successive observations on \( Y \) at levels \( x_1, \ldots, x_n \) determined sequentially in accordance with some definite experimental procedure. If (2) holds in probability irrespective of any arbitrary initial values \( x_1, \ldots, x_n \), we shall, in conformity with usual statistical terminology, call the procedure consistent for the given \( H(y|x) \) and value \( \alpha \). A sharpening of the result is:

**Theorem 1:** Let the following conditions be satisfied

\[
\inf_{x} (x - \theta) \, R(x - \alpha) \geq \theta > 0 \\
\varepsilon < |x - \theta| < \varepsilon^{-1} \text{ for each } \varepsilon < 0
\]

There exists a positive number \( d \) such that for all \( x \)

\[
E y^2(x) \leq d(1 + x^2)
\]

\[
\varepsilon < |x - \theta| < \varepsilon^{-1} \text{ for each } \varepsilon < 0
\]
Robbins and Monro (see Chang [2]) provide a particular procedure for estimating $\theta$ which is considered under certain restrictions on the nature of $H(y|x)$. These restrictions are severe and could no doubt be lightened considerably. But they are often satisfied in practice. No claim is made that the procedure has any optimum properties (i.e., that it is ‘efficient’) but the results indicate at least that the subject of stochastic approximation is likely to be useful and is worthy of further study. [8] Stochastic approximation is being used by engineers specifically to solve problems under limited knowledge which is precisely the situation posited by theories of bounded rationality. Many of the techniques used by econometricians and all of the adaptive schemes have been used in the literature, can be expressed as variants of the Robbins – Munro algorithm whose asymptotic behavior is closely related to that of the more tractable ordinary differential equation: $d\alpha/dt = f(x, \alpha)$

It may not be necessary to analyze in detail the process of reasoning underlying bounded rationality. If we believe that we will choose an action that gets them "close" to the optimum, then we can use the notion of ‘epsilon-optimization’, that means you choose your actions so that the payoff is within epsilon of the optimum. If we define the optimum (best possible) payoff as $U^*$ then the set of epsilon-optimizing options $S(\epsilon)$ can be defined as all those options $S$ such that: $U(s) > U^* - \epsilon$.

Although adaptive methods are useful for selecting among equilibria they do not provide a solid foundation for studying real time behavior, mainly because they require the modeler to make seemingly arbitrary assumptions concerning how people solve their inference problems – how much they know, what class of learning algorithms they use, to what extent they parameterize their inference problems and so on.

**Recursive Statistical Procedure:** Typically, a sequence of estimates is obtained by means of some recursive statistical procedure. The $n$th estimate is some function of the $(n-1)$th estimate and of some new observational data and the aim is to study convergence and other qualitative properties (convergence rate) of the algorithm. A methodology has been presented for the derivation of the asymptotic properties of recursive stochastic algorithms. It has been discussed how to use directly results dealing with stochastic approximation techniques, well-known inequalities, the Lyapunov approach and the martingale theory. In general, the problem of determining whether a given stochastic algorithm will converge or not requires a great deal of ingenuity and resourcefulness. It has been made easier by presenting a methodology, the complete analysis of two recursive algorithms and many direct applications of the standard inequalities, well-known lemmas and theorems.

Actually, very simplistically, Thomas Sargent [9] observes that the first order conditions for stochastic estimation and optimization take exactly the form of solving a differential equation of this form:

$$a_t = a_{t-1} + \gamma Q(\gamma, a_{t-1})$$

were the ‘gain’ behaves asymptotically as $\frac{1}{t}$. Stochastic approximation was designed by engineers specifically to solve such problems under limited knowledge which is precisely the situation posited by theories of bounded rationality. He goes on to show that many of the techniques used by econometricians and all of the adaptive schemes that have been used in the literature surveyed, can be expressed as variants of the Munro-Robbins algorithm, whose asymptotic behavior is closely related to that of the more tractable ordinary differential equation:

$$da / dt = E Q(x, \alpha)$$

So in the end Sargent concludes that although adaptive methods are useful for selecting among equilibria, they do not provide a solid foundation for studying real time behavior, mainly because they require the modeler to make many seemingly arbitrary assumptions concerning how people solve their inference problems – how much they know, what class of real algorithms they use, to what extent they parameterize their inference problems and so on. While Sargent [9] provides lots of evidence to this effect, the problem with the conclusion is its implicit suggestion that rational expectations theory does provide such a foundation. In truth, if bounded rationality does not provide a foundation, then *a fortiori*, neither does rational expectations theory, which derives what power it has as

---

8We suppose that $H(y|x)$ is, for every $x$, a distribution function in $y$ and that there exists a positive constant $C$ such that $Pr[Y(x) > C] = \int_{-\infty}^{\infty} (y|x) = 1$. It follows in particular that for every $x$ the expected value $M(x)$ defined by (3) exists and is finite. We suppose, moreover, that there exist finite constants, such that $M(x) \leq \theta$ for $x < \theta$, $M(x) \geq \theta$ for $x > \theta$. Whether $M(\theta) = \alpha$ is immaterial.
a parable for understanding real economies form the idea that people’s intelligent opportunities will lead an economy to a rationality to model.\textsuperscript{9}

It is certainly true that adaptive control theory leaves many questions unanswered. But to criticize a concept for not having solved the age-old problem of inference says more about one’s priors than about the concept or its literature. If macroeconomics seem lost in a ‘wilderness’, to quote Thomas Sargent, once they depart from rational expectations, it is not for lack of techniques, but for lack of good substantive hypotheses concerning how economies function. Surely the thing to do when lost in the real world is to start searching for clues to guide one’s way, not to retreat to the make-believe world of rational expectations and perfect coordination. Even if one believes like Kenneth Arrow that there is no general principle that prevents the creation of an economic theory based on other hypotheses than that of rationality, the only way to prove the power of including the procedural aspects of decision-making in specific economic theories by actually doing so. This is the challenge facing the scholars of rationality. ‘Adoptive heuristics’ like algorithms might help.

One of the simplest examples of a learning heuristics is a cumulative payoff matching, in which the subject pays actions next period with probabilities proportional to their cumulative payoffs to date. Specifically, consider a finite stage game \( G \) that is played infinitely often, when all payoffs are assumed to be positive. Let \( a_{ij}(t) \) denote the cumulative payoff to players \( i \) over all those periods \( 0 \leq t' \leq t \) when he played action \( j \), including some initial propensity \( a_{ij} \geq 0 \). The cumulative payoff matching rule stipulates that in period \( t + 1 \), player \( i \) chooses action \( j \) with probability

\[
p_{ij}(t+1) = \frac{a_{ij}(t)}{\sum_{k} a_{ik}(t)^{10,11}}
\]

**CONCLUSION**

Some models of human behavior in the social sciences assume that humans can be reasonably approximated or described as "rational" entities (see for example rational choice theory). Many economic models assume that people are on average rational and can in large enough quantities be approximated to act according to their preferences. The concept of bounded rationality revises this assumption to account for the fact that perfectly rational decisions are often not feasible in practice because of the finite computational resources available for making them. Without full rationality one can devise stochastic heuristics that can cause period-by-period behaviors to come close to equilibrium – most of the time (without necessarily converging to equilibrium). There may, however, an infinite number of ‘plausible’ models that can explain social phenomena. Without the assumption of bounded rationality we will be left with a strong sense of arbitrariness. Although the paper did not contribute much to the discussion and settling of the issue, the author thinks it is worth pursuing. The following are examples of basic intuitions that await proper explanation:

- Advertising is an activity that is supposed to influence an economic agent’s decisions not only by supplying information and changing preferences but also by influencing the way decisions are made

\textsuperscript{9}The purpose of this Economic Journal Feature is to critically examine and to contribute to the burgeoning multi literature on markets as Complex Adaptive Systems (CAS). Three economists, Robert Axtell, Steven Durlauf and Arthur Robson who have distinguished themselves as pioneers in different aspects of how the thesis of evolutionary complexity pertains to market environments have contributed to this special issue. The recent Merkosee et al \cite{7} over view aims to use the wider literature on complex systems to provide a conceptual framework within which to discuss the issues raised for Economics in the above contributions and elsewhere. In particular, some assessment will be made on the extent to which modern complex systems theory and its application to markets as CAS constitutes a paradigm shift from more mainstream economic analysis.

\textsuperscript{10}Notice that the distribution has full support given the initial propensities are positive. The idea was first proposed by the psychologist Herrnstein \cite{5} to explain certain types of animal behavior and falls under the general rubric of reinforcement learning, Bush and Mosteller \cite{1}. Reinforcement models differ in various details that materially affect their theoretical behavior as well as their empirical plausibility.

\textsuperscript{11}Classical statistical hypothesis is a heuristic for learning Nash equilibrium of the stage game. Moreover, if the players adapt hypotheses that condition on history, they can learn complex equilibrium of the repeated game, including forms of sub game perfect equilibrium. One can show the following: given any \( \theta \), if the response parameter \( \theta \) is sufficiently large, the test tolerance is sufficiently small ( given \( \theta \) ), and the amount of data collected \( s \) is sufficiently large ( given \( \theta \) ), then the players' period-by-period behaviors constitute an - equilibrium of the stage game \( G \) at least \( \theta \) - of the time. ((Foster and Young \cite{4}, Lus and Sornette \cite{6-12}).
Decisions makers are not equally capable of analyzing a situation when the information available to all of them is the same. The difference in their economic success can be attributed to these difference.

Many social institutions, like standard contracts and legal procedures, exist and are structured as they are, in order to simplify decision making.

REFERENCES