Bounded-Rationality Models: Tasks to Become Intellectually Competitive

RONALD M. HARSTAD AND REINHARD SELTEN

Research in experimental economics has cogently challenged the fundamental precept of neoclassical economics that economic agents optimize. The last two decades have seen elaboration of boundedly rational models that try to move away from the optimization approach, in ways consistent with experimental findings. Nonetheless, the collection of alternative models has made little headway supplanting the dominant paradigm. We delineate key ways in which neoclassical microeconomics holds continuing and compelling advantages over bounded-rationality models, and suggest, via a few examples, the sorts of further, difficult pushes that would be needed to redress this state of affairs. Closer collaboration between theoretic modeling and experiments is clearly seen to be necessary. (JEL B40, C72, C90, D01, D21)

1. Introduction

Economics is witnessing the solid beginnings of a revolution in microeconomic theory. Research has demonstrated a need for boundedly rational models of firms and markets, and their potential for ascendancy into legitimate intellectual competition with neoclassical theory.\(^1\)

The neoclassical research program, by which we mean the broad body of models and inferences based on the premise that all economic decision making reduces to (rational)

\(^1\)The alternative to neoclassical economics, which we oversimplify by discussing in the singular, is intended to be broadly described as modeling and observing behavior outside the paradigm of optimization. This has the effect of treating many who describe themselves as “behavioral economists” as falling outside this alternative (in the neoclassical camp, to the extent that they consider their models fully rational), which we realize may not match self-descriptions (cf., for example, Koszegi and Rabin 2006, Crawford, Costa-Gomes, and Iriberri 2013, and their references). The critique we outline here is directed at optimization as a methodology rather than at the particular arguments of the posited objective function.
optimization of profits or utility, is still—and appropriately still—the ruling mainstream of the discipline’s research, scholarly publication, and microeconomic policy analysis. It has produced an elegant, parsimonious, imposing, imperial structure that is by no means without predictive success.

Yet it can no longer be firmly defended as an appropriate first approximation to reality. Increasingly persuasive evidence has accumulated that the behavioral assumptions underlying the traditional optimization approach are incorrect, and can point analysis in seriously wrong directions. As neoclassical theory has been more clearly elaborated, its serious deficiencies have become increasingly visible.

Bounded-rationality models of economic decisionmaking are not novel, as will be discussed shortly. However, generations of scholars proposing boundedly rational approaches to economic analysis have not yet led us to a coherent body of behavior-based (as opposed to optimization-based) theory that can offer a legitimate challenge to the primacy of neoclassical theory in economics scholarship, teaching, and policy analysis.

This essay lays out a vision of the interrelated tasks necessary to reach that position. It is intended much more as a welcome mat than as a critical commentary on the current state of affairs. In particular, we urge readers not to expect a survey of the theoretical literature on bounded rationality, nor to feel that their favorite (or their own) approach is in any way denigrated by our omission of explicit discussion of it. We regard as uncontroversial the empirical proposition that bounded-rationality models have not reached a position to challenge the central role of neoclassical models in, say, graduate education, or in the funding decisions by economics divisions of national science foundations. Below we cite authors and books when we feel the need to make a point about the particular contribution of that author, book, or paper, or to illustrate the particular function that remains to be fulfilled. We intend no aspersion on any author or work because it leaves some of the tasks articulated below unfinished or yet to be addressed. It is our fervent hope that both scholars who have worked on boundedly rational and behavioral approaches and those new to the field will be drawn to the tasks remaining and that the interrelationships between these tasks will be more fully visualized, drawn out, and applied.

The ultimate task facing bounded-rationality microeconomics is to develop a coherent approach to a range of topics as wide as that addressed by neoclassical microeconomics. However, our illustrative focus is on the thinner side of markets, that is, on the tasks facing a boundedly rational theory of the firm and of the markets in which firms compete with each other, usually imperfectly.

2. The Challenger Weighs In: A Short Background

Behavior-based criticism of optimization by consumers dates back at least four score years to Mayer’s (1932) criticism that the decisions of a consumer spending her budget on hundreds of commodities could not possibly be guided by utility maximization. (As will become apparent in section 7 below, the modern and more compelling version of this criticism focuses on the dimensionality of the problems neoclassical economics assume.

Selection of topics, papers, and authors to cite has been based on closest congruence to the particular point being made, our familiarity with relevant details of the paper, and simplicity of presentation given congruence.

A useful introduction to this side of behavioral economics is Armstrong and Huck (2010).
The perception that neoclassical microeconomics can blithely set aside behavioral criticisms so long as no blatant incoherencies are replicably exhibited has as long a history.

Criticism of the idea that firms can be modeled as optimizers began earlier. For example, Pigou and Robertson (1924) decry concepts like marginal cost and marginal productivity as “empty boxes” without any counterpart in the reality of the firm. Hall and Hitch’s (1939) work on full-cost pricing also proffered such arguments. Neoclassicists merely shrugged off the arguments: Machlup’s (1946) defense of optimization pointed to the driver who unconsciously performs quite complex calculations when overtaking another car on the road.

The first to address the tasks of preparing to challenge neoclassical economics was Herbert A. Simon, beginning in 1955. Simon was not satisfied with merely criticizing (though some of his best barbs remain a discomfiting part of our discourse today), but made initial and seminal strides toward providing an alternative theory. He saw decision making as a process involving a search for alternatives, satisficing, and aspiration adaptation. His pursuit of this approach led him to become one of the founders of two fields of academic endeavor: artificial intelligence and cognitive psychology.

Newell and Simon’s book, Human Problem Solving (1972), is still an important contributor to our understanding of human decision making. Simon’s ideas led to A Behavioral Theory of the Firm (Cyert and March 1963). This literature pushed far enough to develop computer simulations capable of out-of-sample predictions. However, the results were not readily generalizable, even if they were in some respects descriptively superior. Nelson and Winter, in their book on evolutionary economics (1982), proceed from the assumption of boundedly rational behavior. They recognize the need for analytical results and take steps in this direction.

Simon and his followers produced many interesting results. But what is still lacking is a body of systematic theory that is as orderly and teachable as the neoclassical theory of the firm. Next, we describe how more recent developments have exposed neoclassical theory in ways less easily eluded, and then offer thoughts about how to overcome this state of affairs, on what is needed to achieve a competitive challenge to the neoclassical paradigm.

3. Experiments Target Attacks More Tightly

Though perhaps 90 percent of experimental economics papers are less than two decades old, the field began its slow (and for years nearly invisible) ascent (separately in Germany and the United States) a couple of years after Simon began publishing models of boundedly rational behavior. German efforts began, and remained, connected with notions of bounded rationality. In contrast, early American experiments focused on demonstrating the power of neoclassical theory to predict behavior in markets, often to an even greater extent than suggested by mainstream theory (for example, markets converging to competitive equilibrium with only a few buyers and sellers, Smith 1962, 498

4 By the 1990s, it became clear that, should some consumer choices or some net outputs of firms be integer-valued (buying single-paned or double-paned windows, handling repetitive transport needs via owned or via public transport, choosing whether a retail location with largely transient business will be run by the corporation or by a franchisee), the optimization envisioned is NP-complete, thus computationally unmanageable.

5 Accident frequency challenges the adeptness of the Machlupian driver’s optimization. A case can be made that good drivers primarily depend on such rules of thumb as the two-second following distance rule for dry pavement, or the flow-pattern simplification that on an expressway away from exits, each car in your rear-view mirror is either one that you passed or one that will pass you.
1965). Early research was also concerned with developing replicable methods for isolating system effects and the role of incentives, as well as with illustrating the capacity of economic incentives to instill behavior consistent with optimization.

The methodology of experimental economics is capable of isolating impacts of incentives, institutions, and other economic forces on behavior, and is capable of satisfying the assumptions of economic theory, in ways that psychological experiments seldom achieve. It was inevitable that experimental observations would combine with internal elaborations of neoclassical theory to make the deficiencies of the optimization-based approach so directly visible that shrugging them aside is no longer an option.

Perhaps the earliest important example of this phenomenon was Lichtenstein and Slovic (1971), who first reported the phenomenon of preference reversal. Grether and Plott (1979) suspected that their results were due to an inadequate methodology. They introduce their report by saying: “Taken at face value the data are simply inconsistent with preference theory and have broad implications about research priorities within economics. The inconsistency is deeper than the mere lack of transitivity or even stochastic transitivity. It suggests that no optimization principles of any sort lie behind even the simplest of human choices and that the uniformities of human choice behavior that lie behind market behavior may be of a completely different sort from those generally accepted. This paper reports the results of a series of experiments designed to discredit the psychologists’ works as applied to economics” (623). Indeed, Grether and Plott found it easy to discredit nine experiments by psychologists reporting preference reversals; the original Lichtenstein and Slovic experiments, for example, had mis-specified incentives. Many of the psychology experiments were mere surveys of subjects, with no incentives at all to express true preferences or to state higher values for lotteries valued more highly.

However, it was straightforward to carefully apply the methodology of laboratory economics in order to obtain observations that could not be discredited along any similar grounds. Grether and Plott did, and to their express surprise, they found stronger tendencies toward preference reversals than Lichtenstein and Slovic. They provide nine reasons why neoclassical theory could discard the psychological results, but all nine burdens of proof are met by their careful design. This example is important per se, and also because it demonstrates the capacity of economic experimentation to obtain observations of behavior inconsistent with neoclassical theory in situations to which the theory putatively applies.

The simple categorization we specify as {neoclassical and thus optimization} versus {bounded rationality} does not readily characterize either of the other two viewpoints in this forum. Rabin (this forum) discusses approaches such as Gul (1991) and Koszegi and Rabin (2006) that model agents who optimize via objective functions differing from neoclassical utility functions in ways that might not be invalidated by evidence of preference reversals. Rabin also discusses approaches such as Levine (1998) and Fehr and Schmidt (1999) that model agents with utility functions that incorporate interpersonal comparisons and would yield predictions inconsistent with preference reversals. Crawford (this forum) discusses level-k.

---

6This phenomenon is a consistent preference for a less risky lottery over a riskier lottery, but consistently valuing the less preferred lottery higher.

7In several dimensions, the sharp ability of experimental economic design to isolate particular forces within and aspects of microeconomic theory closely relate to the parsimony of a general theory focused on optimization.
models that focus on behavior in games. The sole experimental game data on preference reversal we are aware of is Bohm (1994), in which a subject exhibiting preference reversal behavior could not be consistent with optimization models, including level-$k$.

4 Combined Attacks

We cite two more examples before moving on to describe upcoming tasks. In the first, elaboration of the theory, empirical phenomena, and laboratory experiments play complementary roles. This example begins with Milgrom and Stokey (1982) who derived a static-model theorem implying that the total volume of purely speculative trade in an economy with neoclassical, rational traders is zero, and with Tirole (1982) who proved that speculative bubbles cannot exist even in a fully dynamic rational-expectations equilibrium. While it is difficult on most financial markets precisely to delineate those trades that are purely speculative, it is clear that on a typical business day the no-trade prediction misses the mark by trillions of dollars.

Speculative bubbles in financial and real-estate markets have caused observers to question the relevance of neoclassical predictions. For example, Shiller (2000) reported on the difficulties reconciling neoclassical theory with bubbles observed in the prior decade in several major western economies. Nonetheless, models of rational bubbles have been produced, and some have argued that the ability of neoclassical theory to predict trading in markets withstands this evidence, and that no particular evidence of irrational individual behavior has arisen.

Enter experiments, initially Smith, Suchanek, and Williams (1988). Their experimental setup has been replicated directly, with variants to check for robustness, by at least a dozen researchers. It is so simple as to make the irrationality of transactions prices painfully transparent: an abstract asset is traded, which has a finite life and a common and commonly known distribution of dividends in each period of its life; at no time in these experiments does any trader possess any private information of any relevance. Speculative bubbles occur in which the asset comes to be traded at “impossible” prices—prices exceeding its value even if the highest possible dividends were to be drawn every period. Inevitably, the bubbles crash. To our knowledge, given the absence of private information, this bubble-then-crash data is at odds with behavioral economics and level-$k$ models just as it is at odds with mainstream neoclassical models.

Allen, Morris, and Postlewaite (1993) and Abreu and Brunnermeier (2003), et al. argue that subjects in these bubble experiments may be behaving rationally, but expecting that other subjects will be irrational. That is, it can be rational to buy an asset at any price if one is convinced that someone else is sufficiently likely to buy it from you at an even higher price. These characterizations depend critically on the presence of irrational traders in the market.

However, these explanations are amenable to further experimentation. Lei, Noussair,

---

8 Bohm’s evidence, however, is contrary: in a small sample, preference reversals are notably less common when the preferences are not over abstract lotteries, but over real goods of uncertain value (used cars).

9 Philips (1998) reports a study of a particular commodities futures market, the Brent fifteen-day forward market, where 95 percent of trading volume can be identified as purely speculative.

10 His book echoes and supports concerns he and others had raised over two decades.

11 Several experimenters report decreases in the severity of bubbles as repetitions of the same finite-period life of an identically parametrized security give subjects more and more experience, presumably mainly experience with the inevitable crash. However, Dufwenberg, Lindqvist, and Moore (2005) find little evidence of bubbles when (a) some traders have substantial experience in the experiments with bubbles and crashes, and (b) inexperienced traders are added to the market with awareness that they will be trading with experienced traders.
and Plott (2001) create an experimental design that parallels Smith, Suchanek, and Williams, except without the possibility of resale. Subjects are assigned to be either buyers or sellers, and once an asset is transacted, it remains in the buyer’s hands until the end of its life. Thus, there can be no purchase based on an assumption of profitable resale. (Notice that this type of evidence is only available via controlled experiments.) Bubbles nonetheless persisted, with significant numbers of purchases at prices where purchase could not possibly be profitable. No model based on optimization can explain the bubbles that arise in the Lei, Noussair, and Plott experiments.

The second example begins with an empirical observation. Petroleum engineers Capen, Clapp, and Campbell (1971) observe that the riskiest endeavor an integrated oil company undertakes is offshore exploration. Neoclassical economics clearly suggests then, that this endeavor should reward greater risk-taking with greater average profitability. However, the record from the first auctions of offshore oil tracts in 1954 up to their writing in 1971 showed the reverse, an industry-wide average profitability of zero in offshore explorations. They ascribed the unprofitability to the winner’s curse: the net value of oil recoverable from an offshore tract is hugely uncertain at the time of bidding, and will be independent of the identity of the winning bidder, so a bidder whose estimate of the recoverable value is overly optimistic will be disproportionately likely to win.

Models of such common-value auctions by Rothkopf (1969), Wilson (1977), and Milgrom and Weber (1982) characterize equilibrium bidding. Of course, since bidding zero would guarantee against losses, the equilibrium bid makes a strong enough winner’s-curse correction to yield positive expected profit. Neoclassical economics naturally extends its domain to include auctions that combine competition with large asset-value uncertainties; indeed, equilibrium expected profit is typically increasing in the amount of the uncertainty.

However, experiments have disproved the predictions of common-value auction theory. Kagel and Levin (1986) and Kagel, Levin, and Harstad (1995) find the winner’s curse to be a strong, persistent phenomenon in laboratory experiments using graduate business and statistics students, despite clear and immediate feedback. With enough experience in four-bidder auctions, subjects come to earn positive profits, although they earn only a small fraction of equilibrium profits. The learning generating positive profits, though, is situation-specific; an announced switch from four to seven bidders—when subjects’ 250 periods of experience includes over 100 in seven-bidder auctions—yields a return to negative profits. Given normal patterns of promotion in the corporate world, these subjects were presumably operating with more experience and feedback than seen in most common-value bidding situations.

Nor does auction theory fare well in the simplest possible strategic situation: second-price, private-values auctions. These auctions are incentive-compatible in that your bid will affect whether you win, but not how

---

12 Specifically, in a common-value auction, an asset of unknown but common value is sold to bidders each of whom privately observes their own estimate of asset value. In the “mineral rights” (minor) simplification, bidders’ estimates are independent conditional on the unknown common value.

13 Dyer, Kagel, and Levin (1989) find that highly experienced professional bidders in major building-construction auctions, when brought into the lab, commit the winner’s curse to an extent statistically indistinguishable from experienced graduate students.

14 Second-price rules specify that the highest bidder wins and pays a price equal to the highest losing bid; private-values specifies that each bidder is privately informed of a certain value to him of the auctioned asset, and knows the distributions from which rival bidders’ private values are drawn. Both definitions stem from Vickrey (1961).
much you pay if you win. Whether you bid $50 or $60 is irrelevant if the highest rival bid falls below $50 or above $60; given that it matters, bidding $50 is dominated if your asset value is at least $60, bidding $60 is dominated if your asset value is at most $50. Harstad (2000) finds that, even with 90 rounds of bidding and feedback, subjects systematically bid above the dominant strategy of bidding one’s value—natural but complete feedback does not appear to generate subject awareness of the dominance argument.

Can we argue that even though auction subjects may not be acting like homo economicus, they are nonetheless behaving as if they were optimizing? Eyster and Rabin (2005) offer a theory of “cursed equilibrium” in which a bidder is autonomously unable to fully comprehend the relationship between their rivals’ private information and their rivals’ actions, but is otherwise fully rational. Their theory predicts that subjects in second-price, private-values auctions adopt the dominant strategy, but that in common-value auctions subjects may perceive committing the winner’s curse as optimal. The reason is that they only partly perceive a rival bid below theirs as suggesting that the rival observed private information suggesting a lower asset value. Similarly, level-$k$ analyses (cf., Nagel, 1995, Crawford in this forum, Crawford, Costa-Gomes, and Iriberri 2013, section 5.2, and references there) posit a distribution of rival levels (which can be based on the experimental data), and find the experimentally observed degree of bidding aggressiveness in common-value auctions may be as if level-$k$ optimized. Of course, neither of these models can support their predictions as ex post equilibria. Indeed, in the experiments, these solution concepts require that subjects stick to their models (and still project positive profit conditional on winning with the same bid functions) in the face of immediate and complete feedback showing persistent negative profits.\footnote{In Kagel, Levin, and Harstad (1995), this includes subjects with 250 periods of common-value bidding with the same underlying statistical distributions and continual losses when there are more than three rival bidders; in Albers (1994), this includes subjects with 500 periods of observations of repeatedly cycling into and out of negative profits in five-bidder common-value English auctions. Both sets of experiments feature an onscreen end-of-period report containing all bids, showing for each bid that subject’s private information, and publicly calculating the winner’s profit or loss.}

5. Needs in Order to Challenge Effectively

The type of evidence that has arisen against the neoclassical approach, as experimental economics has gone from infancy to intellectual adolescence, is sharp and incisive. It creates a situation where bounded rationality, or other models not based on optimization, has a welcome mat out, like never before. The dominant paradigm is in transparent trouble. Why, then, does neoclassical economics maintain its dominant position in top journals and on graduate reading lists at top departments? Clearly the answer is that the tasks of making an alternative a serious contender remain to be accomplished. We next describe some of the advantages of the neoclassical approach that make it difficult to challenge, let alone supplant its preeminence. None of the aspects we mention in this section are completely novel, but understanding their cumulative effect is aided by seeing them collected in a few paragraphs.

\footnote{In offshore oil auctions (Capen, Clapp, and Campbell 1971), celebrity-authored publishing (Harstad, private discussions with publishing executives), and several other venues (Smith 1990), participants suggest that the alternative to winning (common-value) auctions somewhat regularly is a negative-profit alternative, a belief more readily maintained in the presence of continual winner’s curse results in the auctions.}
Harstad and Selten: Bounded-Rationality Models

Range of applicability: A single basic set of optimization tools allows neoclassical economics to be applied (with facile reinventing) to a wide variety of topics, ranging from the extremely serious to the less so: arms control negotiations, international trade, competition between political candidates, anthropology and evolution, housing location and cultural and economic segregation, crime and the wages of foot soldiers working for drug cartels, the economics of matching medical residents with positions, of marriage, of extramarital affairs.

Of course, this capacity of neoclassical economics to reign imperial over a multitude of topics and disciplines stems partly from its degree of abstraction, and partly from its related degree of rigor. The theory of satisficing, with adaptation of aspirations, as initiated by Simon, and elaborated, for example, in Sauermann and Selten (1962) and Selten (1998), asserts a wide range of applicability and also proffers a significant degree of abstraction. But it has so far been unable to categorically state the requirements, the assumptions on primitives, for some problem to fall within its grasp. In contrast, the neoclassical theorist confidently asks: “Can we assume the decision maker behaves as if he has a preference order? Yes? Good, can we treat this order as quasi-convex? You’re not sure what that means in your context? Oh, don’t worry. Can we assume that feasible sets are convex, and vary upper-hemicontinuously with key exogenous parameters? Perhaps, you say? Have I got a deal for you!” Of course the speech doesn’t always contain any nine-syllable words, but a more-or-less simple set of questions like this has allowed neoclassical economists to cite theorems stating that their methods are relevant to the problem at hand. Bounded-rationality models have yet to be anywhere near as specific about the claimed boundaries of their applicability.

Coherence, internal structure, and teachability: Kim Border tells the story of the first class session he taught as a new assistant professor at CalTech, in a principles course. He went over the basics, discussed the role of markets and of individual decision making by firms and households, and gave a couple examples illustrating interrelations across markets. At the end of the hour, a student asked, “Couldn’t a couple of fixed points clear up all this confusion?”

Neoclassical economics attains a huge advantage via its simplicity and unified internal structure. It serves to guide attempts at elaboration, more or less standardizing the process of refereeing and publication. And it exposes the methods of “charlatan” economists (those who would claim general economic benefits for governmental policies that favorably affect their purse strings or those of their employer or client without bothering to check the soundness of their conclusions). Most of all, from the undergraduate masses to the professors-in-training, it greatly facilitates the closely related processes of teaching and training students to “think like an economist.”

Isolation of economic forces and definitiveness: The tools of neoclassical economics allow its practitioners to identify and isolate key economic forces. Though there are many examples, the cleanest and among the most important is comparative statics. It is possible to parameterize an optimization problem (e.g., net output decisions as a function of exogenous levels of capital rental rate and wage), and then to simply report the derivatives of the solution with respect to the parameters as indications of the impact of these particular forces in isolation from all others (higher equipment rental, or lower delivery service costs).

Not only does optimization allow practitioners to promptly and readily isolate the roles of forces and parameters, and thus to offer simple answers to questions, but the form and definitiveness of the answers draws attention to the conclusion of some adjustment process (reoptimization or reequilibration) and away from the details of the adjustments that presumably reach these ends. Nonoptimizing behavioral approaches are hard-pressed to do as well. Being about process, boundedly rational approaches will find it more difficult to produce predictions about the isolated roles of individual economic forces that do not get bogged down in issues of process.

Solution concepts and stationarity: Optimization that smoothly adjusts to parameters lends itself to characterization via general equilibria or Nash equilibria or some refinement thereof. This approach cleanly allows assumptions about individual actors—convexity of preferred sets and technologies—to yield analyses of the whole economy or game under study and the interactions therein. Decades of working with equilibria lend a sense that the system remains stationary until underlying parameters change, as when firms become able to accelerate depreciation deductions from taxes.

A particular need, then, for boundedly rational models, is for stationarity concepts, that will be well-defined independent of a particular application. Indeed, ideally a single stationarity concept will be applicable to as wide a variety of models as possible. It might be a rest point or basin of attraction of some system of equations that is generally accepted as indicative of a behavioral dynamics. If it is a set-valued concept, then there will be analytical disadvantages: it is more difficult to describe the way a set responds to a change in parameters than the way a vector responds to parametric changes, and this will limit its appeal and usefulness.

Streamlining and focusing empirical and econometric studies: Neoclassical microeconomics pervades not only core micro classes but also graduate study in econometrics and empirical research methods. The behavior generating databases is assumed to result from optimizing; this streamlines and lends structure to regression analyses; many a regression equation has been justified by a simple appeal to the envelope theorem. An empirical estimation of the impact of a minimum-wage increase on unemployment will be far more complicated if jobseekers are satisficing aspirations adapters and hiring and firing actions begin as qualitative decisions.

Efficiency conclusions and minimal value judgments: Microeconomic theory is able to reach conclusions about allocative efficiency because it assumes that all underlying activity stems from parametrized optimization. We have all seen neoclassical analyses and policy papers point to the allocative efficiency of some position, perhaps cite consumer sovereignty, and suggest that the value judgments taken in recommending a position are truly minimal, merely those of Pareto comparisons.

6. A Critical Dichotomy

One aspect of the internal structure of neoclassical theory deserves special emphasis: by virtue of reducing all decisions to optimizations, neoclassical theory has rendered all decision-making procedures to be quantitative procedures.

Qualitative decision making is clearly a critical part of the behavior economists seek to explain. Consumers have come to expect airfares and the prices at which stocks and mutual fund shares can be purchased to change frequently, but not the prices of most goods and services. In most industries, a firm that has been producing a product or service

18Of course, this presents nearly a litmus test: of earth’s seven billion people, how many are noneconomists who believe that all their own decision making is quantitative?
for some time is not constantly reevaluating its price; it is aware that changing the price incurs adjustment costs and imposes similar costs on retailers and final demanders (and that price changes may provoke reconsideration of habitual orders and purchases that are dangerous to the product’s sales). So price change consideration typically begins as a qualitative decision. Would no price change be preferable to a small price change? How long has it been since the last price change? How likely is it that conditions might change nonnegligibly in the near future, so that delaying a price change might mean that the decision is affected by updated information? Should retailers be consulted or given advance notice? Should a not-small price change be accompanied by a temporary increase in advertising, and, if so accompanied, will it still be profit enhancing? How will we justify the price change in the media, and in response to queries and complaints from retailers and final demanders? With the right answers to these questions, the “how much” quantitative decision may eventually arise, but it may be the tenth or twelfth step in the decision process. It is similarly folly to imagine that the decisions of governmental bodies are quantitative, even when the Federal Reserve is considering whether to adjust interest rates.

Developing a coherent model in which specified categories of decisions are quantitative and the rest qualitative (or the reverse, specifying which decisions are qualitative) is a difficult hurdle that has yet to be directly approached, and which is unlikely to be surmounted in any single leap.

Among the most important reasons why this dichotomy is so critical is the need for and usefulness of aggregation. Fallacies of aggregation (avoiding crowds by going to a further beach or commuting earlier than the height of rush hour, both of which will fail to avoid crowds if everyone attempts the avoidance behavior) can be discerned and explained because the ease with which quantitative decisions aggregate vastly (perhaps incomprehensibly) exceeds that with which qualitative decisions might be seen to have coherent aggregate effects. Aggregation is known to be incomplete in neoclassical micro (cf. Debreu 1971 on possible multiplicity of general equilibria and Sonnenschein 1972 on the loss of characterizations when net demand correspondences are aggregated), but the state of affairs is well beyond what might currently be hoped for in aggregating bounded-rationality models.

7. Elements of Seemingly Possible Approaches

This section briefly presents a few aspects present in some papers on boundedly rational models that appear capable of playing key roles in developing, or in providing illuminating first steps to develop, the sort of comprehensive bounded-rationality theory that might meet the requirements just outlined.

The fundamental tool of neoclassical economics is an objective function that maps the space of all relevant decision variables into a real scalar. Thus, profit (that is, present discounted expected profit) is expressed as a function of a vector whose dimensionality covers all quantities, qualities and characteristics of inputs and outputs, all allocative assignments within the firm, all uses and sources of liquidity, all hedging and risk-management options, all beliefs and representations of rivals’ product positioning, pricing and marketing, for now and all future periods.

---

19 Indeed, neoclassical literatures thrive which blithely ignore equilibrium nonuniqueness (e.g., optimal tax models in the literature begun by Diamond and Mirrlees 1971) and limits to aggregatability (e.g., the computable general equilibrium models).

20 Late last century, one of Harstad’s former colleagues was tasked by an unidentified corporation to demonstrate the solvability via supercomputer of a (one-month) subproblem of profit maximization for which more than 100,000 of the variables were integer-valued, and some 73 million constraints had to be satisfied.
A principal characteristic of nearly all bounded-rationality and a few behavioral economic models is a decision process that does not map all of these variables into a scalar. Demonstration of the realism and usefulness of such decision processes—in this aspect, carrying further the theoretical models of Simon (1955) and Selten (1998) or alternatives—can become a key step toward intellectual competitiveness.

Selten, Pittnauer, and Hohnisch (2012) observe in an informationally realistic, dynamic monopoly laboratory setting that subjects do not attempt to formulate a profit objective when it is not set in front of them, textbook-style. Their evidence powerfully suggests that nearly every subject selects a small number of short-run goals to target in the current decision. Each goal selected (e.g., short-run profits, quality, brand visibility) is per se less encompassing than the usual profit function.\(^\text{21}\) Crucially, at any moment in the decision process, the evidence is overwhelming that a subject selects a single goal as the single temporary focus for adjusting the firm’s activities, temporarily ignoring other goals. The computer calculates whether a suggested change in that goal level is feasible, and if so what further goal improvements are possible. The first time the computer reports that the current list of goal levels is feasible and technologically efficient (i.e., no further improvements feasible), we see the subject tell the computer to implement the current list, with no further search for alternative feasible and technologically efficient goal profile levels. The trade-offs mastered by profit maximization are in fact never directly considered.

Only an economist would find this strange. Take any corporation, and examine the calendar of meetings at which subsets of the top hundred executives interact. The topics of the meetings may be: usage of the internal capital market, adjusting procurement procedures, considering entry into new markets, creating adapted variants of current products so as to obtain more shelf space at retailers, controlling employee health insurance costs, and the like. All of these are problems of much lower dimensionality than overall optimization. Is there ever a meeting where the scheduled topic is “maximizing the discounted sum of expected net profit”?

An example of a key objective for bounded-rationality models, one illustrative of objectives in many avenues of research, stems directly from these findings. How general is the characterization of firm management behavior so strongly demonstrated in the monopoly experiments of Selten, Pittnauer, and Hohnisch? As classically defined, a monopoly does not have to anticipate or react to the behavior of rival firms. Does the multiple-goals-singly-attended characterization extend to oligopolies? If so, in what ways is it adjusted, if at all, to face the imperfectly competitive environment?

Similarly expositive are such questions as: How might monopoly pricing be characterized when this model characterizes monopoly decision making? What comparative statics might possibly be predicted from such a model? Should it become clear that a satisfying-aspirations-adjustment model cannot be specified so as to yield clear comparative statics, then presumably the best substitute would be to establish comparable comparative regularities via replicable experiments.

Arad and Rubinstein (2012) use dimensionality-reducing simplification in a theoretical model for a complex game that can be thought of as a simplified symmetric multi-market oligopoly. Their model comports well with the behavior of two quite different subject pools, one online, each numbering a few thousand. In their model, a player edits the strategy space into a few dimensions,
which are the only aspects of a strategy that he will heed. In each dimension, there is a heuristic focus on a “proper response,” which is approximately a best response should rivals also choose among a variety of strategies falling in the same dimension. Then the player [b] solves his problem by first choosing a value in each dimension, perhaps via recursive analysis akin to “level-k” models (Nagel 1995 and Crawford, Costa-Gomes, and Iriberri 2013 cite the literature). Following that, [b] selects a strategy that meets the values in each dimension (perhaps in an ad hoc way, at this final stage). This model makes predictions about a contest with 250 million strategies, with remarkable success. Their hypotheses are: (i) subjects first focus on how many locations in which to compete; (ii) subjects think about what levels of competition to incur where they compete most, and what token competition, if any, to offer elsewhere; then (iii) subjects allocate their forces to locations in line with (i) and (ii). A clever design records the number of seconds spent making decisions, and tracks players with respect to how they played a simpler game where level-k behavior is overwhelming and may allow inference as to a player’s cognitive level; behavior is shown to strongly suggest the sort of reasoning they model, with steps in the order described. The behavior is completely contrary to either eductive or evolutive approaches to Nash equilibrium (Binmore 1987, 1988). The success of the Arad and Rubinstein model may relate to the fact that, while the game has millions of strategies, it is in fact quite easily described.

Segmenting decisions into dimensions as in Arad and Rubinstein is cognitively quite similar to separating behavior into smaller goals in the monopoly model just discussed. In neither case is the behavior consistent with neoclassical theory.

We again see the Arad/Rubinstein model as a welcome mat, encouraging researchers to examine asymmetric games, and multistage games, in order to develop bounds on the applicability of such a model. There may well be other bounded-rationality models that do a good job describing behavior in different classes of games (perhaps with fewer strategies, perhaps with more dynamic structure). It would be valuable to attempt to discern the range of applicability of all such models arising out of particular classes of games. Indeed, when multiple bounded-rationality models have been circulated, each with a canonical game (or, for that matter, with a canonical competitive decision problem) where the model predicts successfully and neoclassical theory predicts poorly; it will then be easier to comprehend a competition between these models that will help determine the application ranges of competing models.

Aggregation is an unavoidably tougher issue when guided not by optimization but by choices of parameters found acceptable in empirical research. This is true whether one is aggregating markets into a macroeconomy, game players’ strategies into an equilibrium, or stage games into a dynamic pattern. Selten’s impulse balance prediction offers a step in the direction of generality, predicting at least as well as four alternatives in games where no pure-strategy Nash equilibrium exists (Selten and Chmura 2008). This success is despite the fact that the three rival behavioral predictions all depend on parameters, and were granted best parameters in the competition. Being parameter-free, impulse balance can be an equilibrium concept, can be applied to consider how stage-game predictions might be altered in a repeated game, and might offer a mechanism to cleanly describe the aggregate behavior of

\[22\] For evaluation of the strength of the conclusions, see Brunner, Camerer, and Goeree (2011) and Selten, Chmura, and Goerg (2011).
players on one side of a market game. How broadly its predictive success ranges is a subject ripe for experimental investigation.

8. Allocative Efficiency

Arguably the crowning glory of neoclassical economics is the capability in so many situations to provide the assumptions necessary to characterize an equilibrium as allocatively efficient. Still, many neoclassical economists regard these theorems as demonstrating the stringency of the assumptions required for efficiency, rather than as suggesting that efficiency is attainable. The most prominent examples may be the two fundamental theorems of welfare economics and the revenue equivalence theorem in auction theory. However, there is a natural tendency to become inured to assumptions a researcher makes continually for decades, and thus to lose perspective on how extreme these assumptions are. This tendency leads economists to treat efficiency characterizations as meaningful and at a minimum plausible.

We suggest allocative efficiency is the most important example where proponents of bounded-rationality approaches need to be prepared for difficulty in meeting neoclassical economics on completely common ground. It is likely to require a combination of a set of bounded-rationality models, less patchwork and more integrated, with boundaries of theory application that are more nearly understood. In addition, presenting an alternative that is also capable of positive and normative characterizations will require widely accepted experimental results and accepted experimental methods of characterization in context.

It is hard to see how particular allocation mechanisms, or less formal incentive-generating arrangements, can be theoretically characterized as attaining allocative efficiency starting from models that deny primacy to optimization of a single objective function. So for efficiency evaluations or analyses, we should expect widely accepted experimental procedures to play as important a role in the alternative economics toolkit as will boundedly rational models.

Laboratory experiments employing the standard technique of inducing values for abstract commodities can use the induced values to benchmark allocative values. An axiomatic method for doing so was developed by Debreu (1951). When marginal valuations are independent of income (thus, in experiments in which marginal valuations are also independent of gains and losses), the standard approach of comparing total subject earnings with maximum feasible subject earnings is Debrevian. More generally, Debreu’s coefficient of resource allocation has to be calculated directly (Harstad and Marrese 1982 provide an example).

However, this methodology measures efficiency as if the induced values were the (complete) basis for the subjects’ optimization of expected utility. When evidence arises that subjects’ decision methodologies differ systematically from such optimization, it becomes unclear whether efficiency measures derived from induced values are valid. Harstad (2012) proposes and demonstrates a general methodology for obtaining efficiency implications solely from observing behavior in controlled settings, laboratory or field. That method assesses the behavioral efficiency

23The underlying presumption of impulse balance equilibrium is that (a) when a higher action could have led to a better payoff, there is an impulse to adjust the chosen action upward, (b) when a lower action could have led to a better payoff, there is an impulse to adjust the chosen action downward, and (c) when an action led to a negative payoff that a lower action could have avoided, there is an impulse to adjust the chosen action downward. The action at which these three impulses are balanced is the impulse balance equilibrium. If instead of being estimated from the data, the three impulses are treated as a priori equally strong, impulse balance equilibrium can be considered a parameter-free model. (To hope for aggregation, it would have to be.)
of an outcome observed in an experimental session, by appending an appropriately designed aftermarket and observing therein the size of mutually beneficial reallocations unreached in the outcome.

9. Article-Ending Remarks

Of course, it is quite at odds with the objective of this paper to call the remarks here “concluding.” It is logically possible that descriptively accurate bounded-rationality models of economic decision making are necessarily fragmented and incapable of representation in a single, coherent model. However, we believe that the appropriate way forward is to treat such a possibility as premature, to suspend disbelief, and to get on with the tough tasks that can potentially create boundedly rational economics that is capable of more directly and thoroughly challenging neoclassical economics. The research welcomed and suggested here will be valuable even if it ultimately falls short of a “unified field” success. To the extent that theoretical models can clearly suggest which empirical regularities might sharpen focus on stationary states, comparative statics, or aggregation, they open up options for other researchers (usually, experimenters) to investigate these regularities. To the extent that experimental studies falsify predictions of neoclassical economics, designing or extending those studies to provide clues to stationary states, comparative statics, aggregation, or efficiency, may sharpen other researchers’ modeling efforts.

References

- Crawford, Vincent P., and Nagore Iriberri. 2007. “Fatal Attraction: Salience, Naïveté, and Sophistication in...


This article has been cited by:

1. Vincent P. Crawford. 2013. Boundedly Rational versus Optimization-Based Models of Strategic Thinking and Learning in Games. *Journal of Economic Literature* **51**:2, 512-527. [Abstract] [View PDF article] [PDF with links]