Bounded Rationality in Industrial Organization[1]

Glenn Ellison
MIT and NBER

August 2005
PRELIMINARY DRAFT

[1]This paper supported by NSF grant SES-0219205. Tom Chang and Moshe Cohen provided valuable research assistance.

1 Introduction

In the last few years there has been a surge of interest in bounded rationality in industrial orgranization. Much of the recent interest is coming from people crossing over into the field from psychology and economics, but there is also a longstanding tradition within the field. Although the field is not yet as coherent and advanced as most fields surveyed at and Econometric Society World Congress, the fact that diverse strands of research and diverse researchers are coming together seemed to me to make it a topic for which a discussion could be very useful.

The terms "boundedly rational" and "behavioral" have been used by different groups of economists over the years to describe different styles of work. In Simon's (1955) early discussion of bounded rationality, the hallmark of the research on the topic was explicit attention to the cognitive contraints of individuals and firms. The "bounded rationality" in my title is intended much more broadly to include work from several different traditions. One tradition is work that proceeds by assuming that players follow exogenously specified rules of thumb. Work in this vein was being developed immediately before the game-theoretic revolution in IO and has continued at a steady albeit measured pace, most notably being popular where IO has intersected with the game theoretic learning theory. A second tradition is work that grows out of recent work in the area often referred to as "behavioral economics" or "psychology and economics". (I'll use the latter term to avoid causing confusion with other "behavioral" literatures.) Work in this vein is usually motivated by psychological or experimental evidence exhibiting that people depart from self-interested rationality in various ways. Analytically, this type of boundedly rational IO is often similar to traditional rational IO, albeit with the important difference that the function agents are maximizing is not their utility function. A third tradition is work that follows Simon's vision more closely and pays explicit attention to cognitive limitations. By now, this is the smallest of the three.

I begin with an discussion of some of the very early literature and also spend some time on papers that are twenty or thirty years old before moving on the current literature.

2 Early History

It is impossible to identify the beginnings of bounded rationality in industrial organization. Throughout the 1920s, 1930s and 1940s there were active debates about whether firms should be modeled as profit-maximizing or if one should instead use behavioral rules like setting prices at cost plus conventional markup. Rothschild (1947) is in many ways quite like modern behavioral IO papers. He assumes that a firm's "desire for secure profits" is a second objective as important as the desire for maximum profits and discusses various implications in oligopoly settings. For example, he argues that this can lead to price rigidity, to varying contractual terms-of-trade, to price wars, to political actions, etc. Analytically, of course, the paper is quite different from modern work. The conclusions are drawn from very loose verbal arguments. Rothschild is, however, prescient in recognizing this limitation. He explains that "A completely novel and highly ingenious general theoretical apparatus for such a solution of the oligopoly problem has recently been created by John von Neumann and Oskar Morganstern ... Unfortunately, at the time of writing this article I had no opportunity of obtaining a copy of this important book."

"Bounded rationality" came to the fore in the 1950s. Simon (1955) wrote that "the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms." Simon discusses the various challenges involved in optimization and proposes "satisficing" as a more reasonable model of behavior: agents search for actions until they find one that achieves a payoff that provides them with at least their aspiration level.

Cyert and March (1956) is an early application to industrial organization. In the modern style, they start by citing papers from the psychology literature to justify their assumptions about deviations from profit-maximization. Among the implications of satisficing they discuss are that firms are most likely to expand sales when profits are low and costs high, and that dominant firms may tend to lose market share over time. The analytic methodology, of course, is still far from modern: other than one set of diagrams using Stigler's dominant firm model, the analysis is completely informal. They do present empirical data related to their hypotheses.

An impression one gets from the early literature is that it was much better at pointing out significant shortcomings of the rational model than in building alternate frameworks and deriving results from them. Another aspect I find striking is that the focus seems to be almost exclusively on firms' being irrational rather than on consumer irrationality.

3 Bounded Rationality at the Time of the Game-Theoretic Revolution

The late 1970's and early 1980's is usually thought of as the beginning of the game-theoretic revolution in Industrial Organization. It was also a time, however, when several distinct "boundedly rational" or "behavioral" approaches were being developed.

3.1 The Rule-of-Thumb Approach

Consider the simple problem a consumer faces on every trip to the supermarket: should he buy any of the dozens of relatively new products he has never tried or should he continue to buy what he habitually buys. To make a rational decision the consumer must start with a prior over the quality of the unknown goods, he must also have a prior over how quality covaries with cost and think through what these imply about the signals inherent in the price of the good and the amount of shelf space the store has devoted to the product. Whether by observing other shoppers or talking to friends, he will also learn something about the popularity of the product. Correctly accounting for this signal requires a correct prior over the space of possible social processes by which popularity has been achieved. How long has the product been on the shelf before I noticed it? Are the previous consumers people who have used the good before or people trying it for the first time? How many previous customers did they talk to before trying it? What did they know about where those customers got their information and how they used it? Recent models by Banerjee (1992), Bikhchandani, Hirschleifer and Welch (1992), Banerjee and Fudenberg (2004), and Bose, Orosel, Ottaviani, and Vesterlund (2005) provide remarkable analyses showing that the rational approach can be made tractable in some such situations.

Despite my admiration for these papers, when I think about consumers' making such decisions dozens of times in the course of an hour-long shopping trip I have the same reaction that Smallwood and Conlisk (1979) had when contemplating this problem twenty-five years ago:

> ... a would-be optimizing consumer who took account of market popularities would be involved in a massive game theory problem with all other consumers. Is it really plausible that he could solve the game?

The approach of the rule-of-thumb literature is to simply posit rules of thumb that consumers are assumed to follow. One can think of this as similar to the game-theoretic approach, but skipping the part of the argument in which one posits utility functions and derives the behaviors as optimizing choices. Proponents of the rule-of-thumb approach see it as having two advantages. First, in some models it seems implausible that consumers would do the "rational" calculations. This does not imply that the behavior the rational model predicts will seem implausible. But when it does, it seems more reasonable to directly assume a behavior consistent with our intuition. If we are going to do this in the end, why not just start with the behavior we like? Second, in practice rule-of-thumb papers tend to provide more analysis of robustness. One reason is inherent: the analysis of rule-of-thumb models is simpler, making it possible to analyze more variants in the same amount of time. Another is probably sociological: rule-of-thumb authors recognize that given the current norms of the profession they must head off critiques about their behavioral rules being ad hoc.

Smallwood and Conlisk's (1979) model of the effect of product quality on market share remains a beautiful example of the rule-of-thumb approach. Consider a market in which consumers choose between K brands at $t = 0, 1, 2, \ldots$. Suppose that a consumer of product k in period t experiences a "breakdown" with probability b_k. This could be a literal breakdown of a durable good or just a disappointing experience with an ordinary consumable.

In this environment, Smallwood and Conlisk examine what happens when consumers follow particular rules of thumb. Specifically, they assume each consumer continues to use the same product until he experiences a breakdown. When he does experience a breakdown he considers switching and ends up choosing product k with probability proportional to $m_k(t)^\sigma$, where $m_k(t)$ is the market share of product k. By varying the parameter σ the assumption can encompass a range of behaviors. For example, $\sigma = 0$ is purely random choice, and $\sigma = 1$ models consumers who ask one randomly selected friend which brand they used and purchased that product.

The first theorems in the paper explore the connection between behavior at the consumer level and the efficiency of product adoption. For example, they show that when $\sigma < 1$, higher quality products eventually become more popular regardless of the initial conditions, but all products have positive market shares in the long run. When $\sigma > 1$ this is not true and an inferior product can come to dominate the market if its initial market share is sufficiently high. A particularly interesting feature of the model is that when $\sigma = 1$ the highest quality product (or products) always dominates in the long run. In this sense, we see that limited rationality at the individual level – recall that $\sigma = 1$ means copying the decision made by one randomly chosen user without using any information about their satisfaction – can make product adoption socially optimal in the long run. When this happens, it makes the rule-of-thumb assumption all the more palatable: there is no strong incentive to deviate to some more complex behavior.

The Smallwood-Conlisk paper has other features that presage future developments in boundedly rational IO. They analyze how firms who were aware of consumers' bounded rationality would distort their quality choices, and they cite psychology papers that provide experimental support for their behavioral assumptions.

When the game-theoretic revolution swept through industrial organization, much of rule-of-thumb literature could not withstand the onslaught. For example, Schmalensee (1978) used a rule-of-thumb approach to analyze a model in which higher quality products were more expensive to produce and firms chose advertising levels. He argues for the rule-of-thumb approach over the rational approach on the grounds that "it seems implausible to assume that households actually compute optimal solutions to a large number of difficult game-theoretic and information-theoretic problems." Part of his argument for this contention is that "buyers' optimal use of the signals transmitted by firms's choices of advertising levels would depend on the strategies being employed by all sellers." Within just a few years, the mechanics of solving such signalling problems had become so routine that I am sure few economists presented with this latter quote would have guessed that it was meant as a critique of signalling theory rather than as an exposition of how to solve such problems.

3.2 Explicit bounded rationality

While Simon's initial motivation of bounded rationality leaned very heavily on limited human capacity for computation, the theory mostly developed along lines that had little to do with this. Agents were assumed to satisfice rather than to maximize, but there was little attempt to formalize why this might be easier than maximizing or to provide criterion on which to assess the feasibility of other behaviors.

One place in which computational limitations were made explicit was in team theory.[1] In the canonical model of this literature, a firm is modeled as a group of agents sharing a common objective. The firm needs to choose a vector-valued action. Which action was optimal depended on an unknown state of nature. Each employee has some information about which action is optimal. The problem of choosing an optimal action is complicated by the potential presence of two additional costs: each agent may incur a cost of gathering information; and there may be costs of communicating information across agents. Given these information costs, it will generally not be optimal to gather all the available information, nor to convey what has been gathered to a single decision maker. Instead, the firm may want to decentralize decision-making and have separate agents or groups of agents choose different parts of the vector-valued action.

Making decisions according to the optimal decentralized procedure is, of course, the fully rational thing to do in this model if one takes all costs into account. It can thus be seen as providing a rational analysis of why firms are organized as they are. Team theory models can also, however, be seen as tool for looking at a range of industrial organization problems. The idea is that team theory provides information-cost-based microfoundations

[1]See Marschak and Radner (1972).

that can guide our choice of rules of thumb. The traditional analysis of monopoly and oligopoly problems ignores information costs. Hence, what we call "rational" models are really rule-of-thumb models that assume firms use a particular rule of thumb: behave as if information costs were not present. If team theory gives us some general insights into how information costs affect firm behavior, then analyzing IO problems by using rules-of-thumb suggested by team theory should be superior to the "rational" approach as it is conventionally applied.

Radner's (1975) model of cost-reduction by satisficing managers connects with both the explicit bounds and the rule-of-thumb literature. He examines the problem of a firm that has only limited managerial attention to devote to minimizing its costs. Analytically, Radner takes a rule-of-thumb approach. He defines procedures for the manager to follow and examines the consequences.

3.3 Empiricism as "behavioral economics"

The 1970s also featured a third "behavioral" approach to industrial organization that would not normally be thought of as behavioral today. Joskow's (1973) "Pricing Decisions of Regulated Firms: A Behavioral Approach," is a good example. What Joskow does in this paper is simply to estimate a probit model. The dependent variable is an indicator for whether a utility applied to the New York Public Service Commission for an increase in its electric rate. He assumes that they apply for a rate hike if and only if $X_i\beta + \epsilon_i > 0$ and estimates the parameter β. The right-hand side variables X_i include things like the firm's earnings growth.

Economists today will wonder why this would have been called "behavioral." Let me explain and argue that it's not completely unreasonable.

In modern structural empirical IO, this would be thought of simply as a method for estimating the firm's profit function. If the firm is rational and profits are of the form $\pi_i = X_i\beta + \epsilon_i$ where ϵ_i is a normally distributed component of profits that is unobserved by the econometrician, then the probit model provides a consistent estimate of β. Note, however, that the probit model also provides a consistent estimate of behavior if the firm follows the irrational rule-of-thumb of applying for a rate hike whenever $X_i\beta + \epsilon_i > 0$. Whether the firm's behavior is actually profit-maximizing is irrelevant in most applications.[2]

It is because empirical work often involves directly estimating behavior (rather than utility or profit) that one can think of empirical economics as a behavioral approach. The empirical literature then becomes a place where "boundedly rational" industrial organization has quietly carried on for decades. Consider, for example, my empirical work with Judy Chevalier (1997) on risk-taking by mutual funds. We start by estimating the relationship between a mutual fund's performance in year t and how much new business it attracts in year $t+1$. The main focus, is then on how fund companies distort their investment decisions in order to attract new business. It may be possible to provide a plausible rational explanation for the precise form of the relationship between investment returns and the inflow of new business, although my prior would be that it is not.[3] For the main question of interest, however, why consumers choose between mutual funds as they do simply doesn't matter. The initial estimates of consumer behavior tell us everything about consumers that we need to know to think about the firm's optimization problem.

A similar argument could in fact be made for much of empirical industrial organization. Consider, for example, a classic "rational" paper, Porter's (1983) study of price wars in a railroad cartel. Porter estimates a structural econometric model in which demand is assumed to be of the form $\log(Q_t) = \alpha_0 - \alpha_1 \log(P_t) + \alpha_2 Lakes_t + u_t$ and firms' supply decisions are of the form predicted by the Green-Porter model of collusion with imperfect monitoring. Again, to answer his main questions about what the firms are doing, it does not matter whether the demand curve comes from a population of consumers who are optimizing some utility function or a population of boundedly rational consumers. By focusing on consumer behavior as an empirically estimable object, the approach is robust.

Discussing the whole field of empirical industrial organization is obviously far beyond what is possible in this format. For the remainder of this paper, I will treat work that is behavioral in the sense that Joskow's paper was behavioral as outside the scope of my talk.

[2] One notable exception is papers doing welfare analyses.
[3] See Lynch and Musto (2003) and Berk and Green (2004) for more on this.

4 Developments in the theory of bounded rationality

In the 1980's the game theoretic revolution in industrial organization seems to have been all-consuming. A number of papers included slight departures form rationality, but these seem mostly to be unintentional or intentional shortcuts.

Psychology and economics was, of course, developing rapidly at this point. Kahneman and Tversky had published their work on prospect theory in *Econometrica* in 1979 and many new areas were opening up. Developments in psychology and economics have had a huge impact on the new boundedly rational IO, but many other surveys of psychology and economics are available (including two in this volume) and I have limited space, so I won't survey developments in psychology and economics that are not directly related to industrial organization.

I will discuss briefly a couple of contemporaneous developments in theory. In the early 1980's rational game theory was also at its zenith in theory. Major advances were being made in bargaining, reputation, repeated games, equilibrium refinements, etc. By the time Tirole's *The Theory of Industrial Organization* appeared, however, theorists were already looking beyond the rational agent equilibrium paradigm. I'll discuss a couple relevant branches of the literature.

4.1 Learning

Fudenberg-Kreps' (1998) draft monograph "A Theory of Learning, Experimentation, and Equilibrium in Games," promoted a rule-of-thumb approach to the question of whether we should expect players to play the equilibrium of a game. Again, a two-fold argument can be made for this approach. First, a fully rational approach to learning is unappealing — it essentially replaces the assumption that agents play the equilibrium of some simple game with the assumption that they don't know enough to do this but somehow know enough to play the equilibrium of some horribly complicated game in which agents with different information are all simultaneously learning from one another. Second, a boundedly rational approach can provide an enlightening discussion of which rules would and wouldn't lead to equilibrium play. One way to do this is to examine a variety of behavioral assumptions.

Fudenberg and Kreps advocated another: they ask the reverse-engineering question of what assumptions about behavior are needed to get to particular conclusions about the outcome of the learning process.

By the early 1990's such boundedly rational learning models were clearly the hot topic in game theory. Most of this literature was written to interest game theorists, but several of the classic papers included little industrial-organization stories. For example, Kandori, Mailath, and Rob (1993) noted that their model could be motivated as a model of students buying computers in a world with network externalities.

The nicest example I know of a paper addressing a real industrial organization problem by such methods is Mobius's (2001) examination of the rise and fall of competition in local telephone service in the U.S. in the early 20th century. The basic facts about telephone competition are striking. AT&T was a monopolist in the U.S. telephone market until key patents expired in 1893/4. Eight years later, three thousand firms had built separate networks not connected to AT&T's and taken approximately half of the rapidly growing market. The entrants continued to do well for another decade or so as the market quadrupled in size. But a period of consolidation followed, and by the mid 1920s competition had essentially ended.

Mobius provides an explanation for this pattern using a rule-of-thumb learning model. In his model, consumers and businesses are connected by social networks. The network consists of many "islands," each containing a several consumers and one business. Everyone interacts disproportionately with others in their island, but consumers also benefit from interactions with businesses on other islands. The behavior rules that describe how consumers and businesses decide between competing noninterconnected telephone networks are similar to those in Kandori, Mailath, and Rob (1993), Young (1993), and Ellison (1993). In each period most consumers take others' actions as given and choose myopically between having no phone or buying service from firm A or firm B to maximize their current period utility. Firms' decision rules are similar, but firms may also want to get two telephones so as to be accessible to consumers on both networks. A small fraction of consumers are assumed to adopt randomly in each period. Essentially, what Mobius has done with these few assumptions is to describe a Markov process governing telephone adoption. What remains is just to analyze the behavior of the process. Using both analytic results and simulations he shows how the model can account for the growth and decline of independents. Roughly speaking, what happens in the early period is that each network grows by adding new subscribers within the islands in which they are dominant. Later, the "duplication" effect becomes important. Businesses in islands dominated by the minority telephone network start to get a second phone to communicate with consumers from other islands. Once this happens, the minority network has only a weak hold on the islands it dominates and can easily lose a whole island in a chain reaction if a few friends randomly decide to switch to the majority system. He presents empirical evidence consistent with this mechanism.

Thinking about the methodology, what makes the boundedly rational approach so useful here is more related to the second advantage mentioned above. The relative tractability of making direct assumptions about behavior allows Mobius to enrich his model in a way (incorporating the social network) that that would have made it too complex to analyze had he insisted on assuming consumer rationality.

4.2 Computational complexity

Explicit concern for computational complexity entered game theory at about the same time. Rubinstein (1986) and Abreu and Rubinstein (1988) discussed complexity in the context of repeated games. In particular, they discussed the concept of strategies' being implementable by finite automata and used the number of states of an automaton as a measure of its complexity. Standard repeated game equilibria can be implemented by automata, but they note that a conundrum appears if one thinks of agents as trying both to maximize the payoffs they receive in the repeated game and to minimize costs inherent in constructing a more complex automata: agents will be tempted to leave off any states that are not reached on the equilibrium path and this may render the kinds of off-path threats necessary to support cooperation infeasible.

Fershtman and Kalai (1993) examine some IO questions using a similar complexity notion. They consider a multimarket monopoly and single- and multimarket duopoly competition. Implicit in the complexity notion is a diseconomy of multimarket operation. If there are two possible states of demand in each market, then there are 2^n possible states for the n market combination, and a fully optimizing monopolist would therefore need a plan with 2^n contingencies. They note that monopolists may want to avoid entering some markets to reduce complexity and discuss the complexity of dynamic collusion.

The models of Rubinstein (1993) and Piccione and Rubinstein (2003) discussed below develop alternate approaches to complexity. Rubinstein (1998) provides a nice survey of old and new work in this vein.

5 The New Behavioral IO: Firm Reactions to Behavioral Biases

One factor that has clearly contributed to the recent surge of interest in bounded rationality in industrial organization is the bigger and slightly less recent surge of interest in psychology and economics. This literature has by now documented a plethora of ways in which real consumers depart from the rational self-interested ideal: they discount in a nonexponential manner, exhibit loss aversion; care about fairness; have self-serving biases; are influenced by how decisions are framed, etc. What is probably more important for the development of behavioral IO is that the psychology and economics literature has developed a number of models of irrational behavior that can be applied in place of utility maximization in various environments.

Once we have a set of alternatives to utility maximization there is a natural set of behavioral IO papers to be written: how will firms price if consumers have behavioral bias X? In fact, we really have a whole matrix of such paper topics. Think of the set of behavioral biases as the column headings, and put all of the standard questions in IO as the row headings: how will a monopolist selling durable goods behave; how will a monopolist price discriminate; how will oligopolists selling differentiated goods set prices; what factors will affect the number of entrants, etc. Most elements of this matrix are worth exploring.

In many cases, of course, researchers will not end up with any conclusions that are sufficiently deep to warrant their being published. For example, Hossain and Morgan

(2005) conduct field experiments on eBay and find that auctioning goods with a high (but not too high) shipping charge raises more revenue than using an equivalent minimum bid and making shipping free. It's a striking fact and they've written a very nice psychology and economics paper around it that discusses how the results can be explained by mental accounts with loss aversion (or other explanations). The companion how-should-firms-price paper, however, seems too obvious to write: if consumers behave this way then firms should use high shipping charges.

In other cases, there will be interesting implications: interesting potential explanations for why firms price as they do and interesting observations about when irrationalities do or don't matter. I'll discuss a couple papers here and speculate about some of the general patterns we'll see.

5.1 One bias: hyperbolic discounting

Della Vigna and Malmendier's (2004, 2005) work on selling goods with delayed benefits (or delayed costs) to time-inconsistent consumers nicely exhibits the potential of behavioral IO. Their motivating example is pricing at health clubs: they think of consumers as incurring a short-run disutility when visiting a club, and enjoying a delayed reward in the form of better health.

The model uses a theoretically appealing (and practically relevant) degree of generality in pricing. Consumers are initially offered a two-part tariff with an upfront payment of L and an additional per visit charge of p. If consumers accept this offer, they learn the disutility d that they will incur if they visit the club, and then decide whether to visit (which costs p and gives a delayed benefit b.) Note that the decision to accept the two-part tariff is made under symmetric information, so in a rational model the health club would extract all consumer surplus via the fixed fee.

Consumers are assumed to have (β, δ) quasi-hyperbolic discounting preferences: from the perspective of period 0, payoffs in period t are discounted by $\beta\delta^t$. They consider both naive hyperbolic consumers who don't realize that they have a commitment problem, and sophisticated hyperbolic consumers who are fully aware of their time inconsistency.

In this model, one can think of two reasons why a health club will want to distort p away from marginal cost. First, sophisticated rational consumers would like to commit themselves to go to the health club more often. The health club can help them to do this by setting p below cost. Second, naive rational consumers will overestimate the number of times that they will go to the club. Reducing p and increasing L widens the gap between the surplus that the consumer expects to receive from accepting the contract and what the consumer actually receives. Hence, distorting the contract in this way makes the contract more appealing to these consumers. The two effects go in the same direction, so regardless of what type of hyperbolic consumers we have, we reach the same conclusion: visits will be priced below marginal cost.

Is this distortion bad? In the case of sophisticated consumers it is not. Essentially, one can view the health club as selling a commitment device that allows the consumer to overcome his or her self-control problem and achieve the first-best. In the case of naive hyperbolic consumers things are less clear cut. The low per-visit price does get consumers to visit the gym more often than they would with marginal cost pricing. Consumers' misperception of the value they will get from the contract, however, leads to their receiving negative total surplus.

Della Vigna and Malmendier's (2005) empirical study lends credence to the view that time-inconsistency is an important factor in health-club pricing. The health clubs they surveyed offer contracts in which the per-visit charge is set below marginal cost, and most consumers take them. In fact, they take these two-part tariffs even though the clubs also offer contracts with a higher per-visit charge that would be cheaper for most consumers. Clearly, people either must value the commitment or must overestimate their future usage. Survey data suggest that overestimation is at least part of the story.

5.2 Another bias: loss aversion

Heidues and Koszegi (2004) similarly explore many aspects of pricing to consumers who experience loss aversion. The formal specification of consumer behavior follows Koszegi-Rabin (2005). Consumers have a reference-dependent utility function that depends not only on the number of units x of the good they consume and on the amount of money m they have left over, but also on reference levels r_x and r_m which are their ex ante expectations: $u_t(x, m) = xv_t + m + \mu(x - r_x) + \mu(m - r_m)$. Loss aversion comes in via the function μ, which is assumed to be have a kink at the origin. Note that consumers are averse both to being unable to buy a product they expected to buy and to paying more than they expected to pay.

Heidues and Koszegi consider a world in which demand and cost are both random and time-varying. In the base model, firms are assumed to be able to commit to a distribution of prices. Some interesting effects arise because this commitment allows the firm to manage consumer expectations. One observation is that prices will be sticky and firms may set a constant price (or choose prices from a discrete set) even when cost shocks are continuously distributed. The disutility that loss-averse consumers feel when they pay more than they were expecting is greater than the utility they derive from paying symmetrically less than they were expecting. This provides an incentive to keep prices constant. A related observation is that markups are countercyclical. Another fundamental conclusion is that there is substantial scope for indeterminacy: as in Koszegi-Rabin (2005) the purchase decision contingent on the price is not necessarily unique because the consumer has a higher valuation when he expects to buy, and even apart from this there can be multiple pricing equilibria when the firm lacks commitment power.

5.3 Some thoughts

To think about the general principles behind how firms "distort" pricing to exploit boundedly rational consumers, it helps to go back to the basic quality-selection framework of Spence (1975). A good way to review the basic ideas of this literature is to consider a population of consumers with unit demands. Suppose that a consumer of type θ gets utility $v(s; \theta) - p$ if he or she buys one unit of a good of quality s at price p. Consider a monopolist that must choose a single quality level s for its product (putting aside for now the possiblity of using multiple quality levels to practice price discrimination.) Assume that the good is produced under constant returns to scale, with a quality s good having a unit production cost $c(s)$. Assume θ is uniformly distributed on $[0, 1]$ and that $\partial v/\partial\theta < 0$ so that a monopolist that sells q units will serve consumers with $\theta \in [0, q]$.

The monopolist's problem is

$$\begin{aligned} \max_{q,s,p} \quad & qp - qc(s) \\ s.t. \quad & v(s;\theta) - p \geq 0 \quad \text{for all } \theta \in [0, q] \end{aligned}$$

Conditional on q and s the optimal value for p is $v(s; q)$, so this reduces to

$$\max_{q,s} \; qv(s; q) - qc(s)$$

The first-order condition for the monopolist's quality choice s^m is

$$\frac{\partial c}{\partial s}(s^m) = \frac{\partial v}{\partial s}(s^m, q^m).$$

The marginal cost of providing higher quality is equal to the marginal benefit for the marignal consumer.

The social planner's problem would be

$$\max_{q,s} \int_{\theta=0}^{q} v(s;\theta)d\theta - qc(s).$$

The first order condition for the first-best quality choice s^{FB} is

$$\frac{\partial c}{\partial s}(s^{FB}) = \frac{1}{q^{FB}} \int_{\theta=0}^{q^{FB}} \frac{\partial v}{\partial s}(s^{FB}, \theta)d\theta.$$

The marginal cost of providing higher quality is equal to the marginal benefit for the average consumer.

Spence emphasized that monopolists typically will not provide optimal quality: the marginal and average consumer can have different valuations for quality; and the pool of customers served by the monopolist differs form the pool the social planner would want to serve. One case in which the monopolist's quality choice is optimal, however, is when the population is homogeneous. In this case, both the monopolist and the social planner would serve all consumers and there is no difference between the marginal and average consumer. Hence, the monopolist provides optimal quality.

The quality-selection model is relevant to the question of how firms respond to behavioral biases because one can regard the things a firm commits to in the contract it offers as quality choices. For example, the per-visit charge a health club imposes can be thought of as a product characteristic that affects the quality of a membership in the club. A contract with a lower per-visit fee is a higher quality contract.

The "optimal quality" result of Spence's model implies that a monopolist with a homogenous unit-demand customer population will choose to sell the product s that maximizes $v(s) - c(s)$. In a rational model, the $v(s)$ in this equation is the equilibrium utility the consumer will receive if her or she buys a product of quality s. To apply the same result to an irrational model, one just has to keep in mind that the appropriate $v(s)$ is more generally the willingness to pay for a good s at the moment at which the consumer makes the purchase decision.

In Della Vigna and Malmendier's sophisticated hyperbolic model, just like in the rational model, the willingness to pay of the time zero agent is the equilibrium utility that the time zero consumer will receive if he or she signs up for the health club. Hence, the outcome is that the health club chooses the contract that is optimal from this perspective. In a naive hyperbolic model, willingness to pay is the time zero agent's (incorrect) forecast of his equilibrium utility. Hence, the contract is designed to maximize the difference between this and cost.

The same basic intuition will apply to other behavioral biases. Hence, a general priciple is that monopolists will distort product characteristics along whatever dimensions increase irrational consumers' willingness-to-pay. What these distortions are under alternate behavioral specifications, of course, still needs to be worked out, and as in the papers mentioned above doing so has the potential to provide interesting insights into why firms price as they do.

Another question that will come up frequently in discussions of firms' exploiting behavioral biases is whether competition eliminates the exploitation of behavioral biases.

In the homogeneous case, the answer is clearly no. Competitive firms will behave just like monopolists and make product- and contract-design choices s so as to maximize

$v(s) - c(s)$. The only difference is that the surplus will be returned to consumers in the form of lower fixed fees.

In heterogeneous populations, the precise choices of monopolists and competitive firms will differ. How this will work will differ depending on whether firms are or are not assumed to be able to offer multiple quality levels.

For some applications one would want to restrict firms to offering a single product quality. For example, suppose that developing a version of the product requires fixed costs sufficiently large so that only one quality will ever be produced in equilibrium. (For the competitive version of this model suppose firms offer contracts (p, s) and compete for consumers before incurring the production cost so one firm will serve the whole market but earn zero profits.) In this case, the monopolist will design its product to appeal to its marginal consumer, whereas the competitive firms will design their products to appeal their marginal consumer (who has a lower valuation). Although the monopoly and competitive product designs will be different, one imagines that qualitative conclusions about the direction in which product designs are distorted from the design that would maximize a rational consumer's utility will be similar in the two cases. This should be the case as long as the behavioral biases of the two marginal consumers are similar.

For other applications it will be better to assume that there are no fixed costs of product introduction so firms can introduce a continuum of products at different quality levels and use these to price discriminate. A monopolist, will typically introduce a continuum of goods. The type $\overline{\theta}$ with the highest willingness to pay is typically sold a product of the "optimal" quality, which here means maximizing $v(s;\overline{\theta}) - c(s)$. Other types will be sold products of quality lower than that which maximizes $v(s;\theta) - c(s)$. Whether this adds to or offsets the distortion that comes from exploiting the difference between v and utility will depend on the application. With perfect competition there are no quality distortions. What this means in the irrational context is that each type θ is sold the product that maximizes $v(s;\theta) - c(s)$. This is higher than the quality offered by the monopolist. Whether "higher" is better or worse socially than the monopoly choice will depend on how t he monopolist's choice differs from the utility-maximizing one given the form of the consumer bias.) Models of oligopolistic price discrimination will produce other results.

6 The New Behavioral IO: Bounded Rationality as Part of the IO Toolbox

My impression from reading papers from the 1980's was it was extremely important that IO theory papers be fully rational. Today, being less that fully rational is much more acceptable. Indeed, if I had to guess I'd say that at the top general interest journals an IO paper actually has a better chance of being accepted if it sells itself as being "behavioral" or "boundedly rational."[4] Regardless of whether this guess is right, I think the change is a good one. The legitimization of boundedly rational explanations provides an opportunity to explore new explanations for old phenomena and opens up whole new topics. In this section, I'll discuss some recent research in several parts of IO. Some of the papers take a rule-of-thumb approach, some cite work in psychology and economics, some examine explicit computational constraints, some are hard to classify. I organize my discussion by the IO topic being addressed.

6.1 Technology adoption

In the field of technology adoption, the old critiques of rationality still ring true. It would be awkward to assume that consumers have correct priors over the set of new products that might be invented and even more awkward to assume that consumers fully understand the social processes by which some products achieve popularity. Ellison and Fudenberg (1993) builds on the Smallwood-Conlisk approach to explore whether dispersed information about product quality will be aggregated at the population level and come to be reflected in adoption decisions.

We identify two mechanisms by which such "social learning" can occur. In one information is aggregated via agents paying attention to popularity. In the other it is aggregated by geography. In both models, there are two technologies f and g. The payoffs are random

[4] One anecdote to this effect is that the *Journal of Political Economy* (which I've always thought of as the ultimate proponent of rational explanations) recently explained that it was rejecting a paper not because of anything the referees said but because the editor's "gut feeling is that most of the time, forgetfulness, confusion, or tother types of bounded rationality play a much more crucial role."

variables related by $u_t^g - u_t^f = \theta + \epsilon_t$ with ϵ_t uniformly distributed on $[-\sigma, \sigma]$. Boundedly rational players observe $u_t^g - u_t^f$, but don't know the full history of payoff realizations. The popularity model looks at a homogeneous population in which there is inertia and in which players who do switch are more likely to use more popular actions: players choose g if $u_t^g - u_t^f > k(m_t(f) - m_t(g))$, where k is a positive constant and $m_t(f)$ is the market share of technology f at time t. In this model, we show that with no popularity weights ($k = 0$) both technologies get positive shares in the long run. With $k > \sigma$ there can be herding on an inferior technology. With $k = \sigma$ full learning occurs in the long run and all consumers eventually adopt the superior technology.

In the geographic model, players are arrayed on a line (which could be a physical location or a position in some social network or taste space). There is no longer an uniquely best technology – which technology is superior depends on the location. Players observe the technology choices of those in a neighborhood around them, and also observe the average payoff of any technology used in their neighborhood. In this model, we show that social learning occurs via a geographic mechanism: the dividing point between the f-adopters and the g-adopters shifts over time and a law-of-large-numbers mechanism keeps it close to the optimum when the observation neighborhood is small. There is a tradeoff between the speed of adoption and its long-term efficiency.

Ellison and Fudenberg (1995) consider a model closely related to the nonspatial model described above. There are two primary differences: there is a random idiosyncratic component to the payoff each agent receives in each period; and we consider rules-of-thumb in which an agent asks k randomly selected members of the population about the payoffs they received in the current period and chooses the technology that provided the highest payoff in this random sample (with the proviso that an agent never adopts a technology that was used by no one in his sample.) We focus on the question of how the structure of information flows affects the learning process.

Considering first the case where the two technologies are equally good, we show that the value of k affects whether we see "herding" on one technology or "diversity" with both technologies retaining a positive market share in the long run. Intuitively, when agents rely on small samples, they are unlikely to hear about unpopular technologies, which makes it more likely that these technologies will die out. We then introduce differences in average payoffs across the technologies, and show that long-run efficient social learning will occur for a range of values of the sample size k. When k is smaller herding on an inefficient technology may be possible, and when k is larger both technologies may survive in the long run. Intuitively, there is a degree of popularity-weighting implicit in the assumption that agents don't adopt technologies not being used in their random sample, and this can support social learning as in the earlier model.

Spiegler (2004) introduces price-setting into a related model. There are N firms. Each firm has a distinct (albeit equally good) technology that provides random payoffs that are independent and identically distributed both over time and across individuals: each technology gives a payoff of one with probability α and zero with probability $1-\alpha$. There is also an $N+1$st technology that provides this same payoff at no charge. (Spiegler discusses unscientific medical practices as a potential application and refers to the firms as quacks.)

Consumers are assumed to use a rule-of-thumb that involves finding one person using each technology, treating the payoff, v_i, that person received as if it were the expected payoff of technology i, and choosing the technology for which $v_i - p_i$ is maximized. This is similar to the decision rule of Ellison and Fudenberg (1995), with the main difference being that each technology is sampled once rather than a popularity-influenced random number of times. Spiegler notes that the rule-of-thumb can also be motivated as an extreme form of Rabin's (2002) model of "believers in the law of small numbers" or as an application of a solution concept in Osborne and Rubinstein (1998).

The pricing game between the firms turns out to have an unique mixed strategy Nash equilibrium (not unlike the equilibria one sees in search-based models of price dispersion). In this equilibrium, expected prices are inversely related to the common product quality. Industry profits are nonmonotone in the number of firms. Industry profits initially increase in N because there is more likelihood that consumers' samples will contain at least one firm success, but decline as N becomes large because the price competition becomes more intense. Adding one higher quality firm would not change the other firms' strategies.

Spiegler's paper also illustrates an advantage of the rule-of-thumb approach. In a rational model with asymmetric information the firms' prices would be potential signals of quality, and the only thing one could say is that a wide variety of outcomes are possible depending on the beliefs consumers hold about out-of-equilibrium signals.

6.2 Models of sales

The classic models of sales by Varian (1980) and Sobel (1984) are described by their authors as rational models. At a different time, the papers could also have been sold as boundedly rational or behavioral: Varian's model features some 'high search cost' consumers who are ignorant of the prices each store offers and end up going to one store chosen at random; and Sobel's features some infinitely impatient consumers who buy immediately regardless of the potential gain from waiting for a good to go on sale. In each case, a very rough intuition for why there are sales is that firms want to price discriminate and give discounts to the more sophisticated shoppers.

Rubinstein (1993) and Piccione and Rubinstein (2003) develop new formal approaches for modeling differences in agents' cognitive abilities and use these frameworks to provide models of sales in which the discounts-for-sophisticated-consumers intuition is more firmly grounded. In Rubinstein (1993), cognitive complexity is captured by the order of the "perceptron" needed to implement a strategy. Some agents can only implement very simple strategies (buying if price is above a threshold), whereas others can implement nonmonotone strategies involving two or more cutoffs. He writes down a model in which a monopolist wants to charge high-cognitive-ability agents a lower price in some states, and in which the monopolist can achieve this by randomly choosing prices in a manner that makes it unprofitable for low cognitive-ability customers to also buy in this state. Piccione and Rubinstein (2003) introduce an alternate form of differential cognitive ability: they assume that agents differ in the length m of the price history they can recall. They again consider an environment in which the firm would like to charge high-cognitive-ability agents a lower price, and show how this can be done by alternating between regular and sale prices in a manner that high-ability agents can recognize (letting them buy only when the item is on sale) and low ability agents cannot (which forces them to pay the time-average price).

Kahnemann, Knetsch, and Thaler (1986) had much earlier proposed that sales might have a different irrational origin. They conduct a survey, and find that many subjects say that it is "unfair" for firms to raise prices when demand goes up and speculate that this may give firms an incentive to hold sales rather than reducing "regular" prices: if firms lower regular prices when demand is low, they will be branded as unfair if they raise prices back to normal when demand returns to normal. Rotemberg (2005) proposes a more complex fairness-based model to account both for firms' occasional use of sales and for the stickiness of prices within and across non-sale periods. In his model, consumers have reciprocal-altruism preferences and punish firms discontinously if their estimate of the firm's altruism crosses a threshold. The model also relies on the firms' objective function being a concave function of profits and on consumers feeling regret. He argues that similar assumptions can also help explain how firms adjust prices in response to demand shocks and inflation.

6.3 Price dispersion

The IO literature developed a number of search-cost based explanations for price discrimination in the early 1980s.[5] These models are compelling for many real-world applications, but Baye and Morgan (2004) note that dispersion seems to occur even in enviroments where it is less plausible that substantial search costs exist: it is common for products listed for sale on Internet price search engines; and it even occurs in laboratory experiments in which a computer plays the buyer's role to perfectly recreate a Bertrand duopoly![6]

Baye and Morgan propose that one reason why this occurs may be that the equilibrium of the Bertrand competition game is somewhat nonrobust to departures from rationality. Suppose that firms are ϵ-optimizers rather than being fully rational, i.e. they may pick any action that earns profits that are within ϵ of the maximum possible profit. One might think at first that the Bertrand model is fairly robust to such a change because in any *pure strategy* ϵ-equilibrium industry profits would be at most ϵ. Baye and Morgan note, however, the Bertrand game also has mixed strategy ϵ-equilibria with much higher profits.

[5] Baye and Morgan (2005) provide an excellent survey.
[6] See Dufwenberg and Gneezy (2000).

For example, for any $x \in (0, \pi^m)$, suppose firms use the mixed strategy with CDF

$$F^x(p) = \begin{cases} 0 & \text{if } p < \underline{p} \\ 1 - x/pD(p) & \text{if } p \in [\underline{p}, p^m) \\ 1 & \text{if } p \geq p^m \end{cases}$$

where $\underline{p}$ is such that $\underline{p}D(\underline{p}) = x$. In the Bertrand duopoly game with zero costs, the maximum gain from deviating from this profile is $x^2/2\pi^m$. If we set this to ϵ, we find that a firm can earn at least $\sqrt{2\epsilon\pi^m} - \epsilon$ in an ϵ-equilibrium. This is a strikingly large number for reasonable values of ϵ. For example, if ϵ is one percent of the monopoly profit, the aggregate industry-profits in and ϵ-equilibrium can be approximately 26% of the monopoly profit. For $\epsilon = 0.04\pi^m$, we get about half of the monopoly profit.

6.4 Obfuscation

One of the most basic results on the economics of information disclosure, is that firms will disclose all relevant information to consumers if this is possible and costless (Grossman 1981; Milgrom 1981). The simple intuition is that the firm with the highest possible quality will aways want to disclose to increase consumers' willingness to pay, so any attempt at constructing a pooling equilibrium will fail because things unravel from the top.

Ellison and Ellison (2005) note that there are startling differences between this theoretical prediction and what we see in many real-world environments. For example, mattress manufacturers put different model names on products sold through different stores and provide sufficiently few technical specifications so as to make it very difficult to compare prices across stores. Credit cards are another good example: it is also hard to imagine that the complex fee schedules in small print on the back of credit card offers could not be made simpler. Our empirical work examines a group of small retailers selling computer parts over the Internet via the PriceWatch search engine. Many of these firms have clearly adopted pricing practices that make it time-consuming and difficult for consumers to understand what exactly a firm is offering and at what price. For example, products are described incompletely and consumers may have to go through many pages to learn all of nonstandard aspects of the offer: the restocking fee that will apply if the product is returned, how much extra the consumer will need to pay for a warranty, etc.

An interesting empirical result is that retailers appear to obtain substantial markups over marginal cost even though there is easy entry, minimal product differentiation, and minimal inherent search costs. One imagines that this is not unrelated to the obfuscation. One doesn't find any discussion of obfuscation in the fully-rational IO literature. Models of bounded rationality seem like a natural starting point.

The most straightforward way to think about obfuscation using standard IO tools would be to regard it as increasing search costs in a model with costly search like Stahl (1989). One would want, however, to extend the standard models to allow the search costs to vary by firm (instead of just across consumers). Even then, we would have a fairly black-box model and it would be nicer to derive the search costs as arising from some model with costly computation.

Gabaix and Laibson (2004) offer one very simple proposal. They suggest that one could regard obfuscation as increasing the variance of the random evaluation error ϵ_i in a model in which consumers have noisy estimates of the utility they will receive from consuming a product: they think they will get utility $\delta_i + \epsilon_i$ from consuming product i when they actually get utility δ_i. Such a model is formally equivalent (from the firm's perspective) to a model in which firms can invest in product differentiation, and results about firms having an incentive to invest in differention to raise markups will carry over. They emphasize that for some error distributions adding firms to the market will drive down prices very slowly.

Spiegler (2005) offers another rule-of-thumb model. In his model, products inherently have a large number of dimensions. Consumers evaluate products on one randomly chosen dimension and buy the product that scores most highly on this dimension. In this model, consumers would evaluate the products correctly if products were designed to be equally good on all dimensions. Spiegler shows that this will not happen, however. Essentially, firms randomize across dimensions making the product very good on some dimensions and not so good on others. He thinks of this cross-dimensional variation as intentional obfuscation. The comparative statics of the model may help us understand why there is more obfuscation in some markets than in others: obfuscation increases when there are more firms in the market, and when the outside option is more attractive.

6.5 Add-on pricing

In many industries, it is common to sell high-priced add-ons. Hotels charge high prices for dry cleaning, telephone calls, minibar items, and restaurant meals. Credit cards have high late-payment fees. Upgrading to a larger hard drive and adding more memory adds substantially to the price Dell advertises for their computers. Ellison (2005) shows how the joint adoption of add-on prices can raise industry profits in a competitive price discrimination model. The conclusion derives from an assumption that differentiates the model from most of those discussed in Armstrong (2005) and Stole (2005). Whereas most competitive price discrimination models assume that brand preferences and quality preferences are independent, I assume that price-sensitive consumers are less likely to buy high-priced add-ons. In such an environment, firms that rely on sales of add-ons for a large part of their profits face a severe adverse selection problem: price cuts disproportionately attract cheapskates who don't buy any add-ons. This discourages pricecutting and raises equlibrium profits.

An interesting question raised by the profitability of the joint adoption of add-on pricing is when is it individually rational for firms to adopt add-on pricing. The classic arguments of Diamond (1971), Klemperer (1987), and Lal and Matutes (1994) are based on informational limitations. If it is costly for consumers to observe a firm's add-on price, then at any price less than the monopoly price firms would have an incentive to deviate and make add-ons slightly more expensive than consumers are expecting. This is a good story for many applications, but not for others. Typically, consumers would be able to buy a more efficient bundle if add-ons were priced closer to cost. If firms could commit to setting such prices and inform consumers via advertising (without incurring extra advertising costs), then firms in a rational-consumer model would typically want to do this.

Gabaix and Laibson (2005) formally develop a boundedly rational explanation along lines suggested in Ellison (2005). In this model, there are two types of consumers: rational consumers and irrational consumers who will buy the expensive add-on only if it has not been advertised. In such an environment, the profits earned serving rational consumers can be increased by advertising a higher base-good price and a lower add-on price, but profits obtained by serving irrational consumers are reduced. This is a consequence of the assumption that the irrational consumers will avoid buying the add-on if they are made aware of it in advance. I agree with Gabaix and Laibson's belief that assuming some customers are irrational is realistic. In an earlier era where such assumptions were not allowed, their paper could probably have also been sold as a rational model. One could have assumed that there was a state the world in which add-ons exist and states of the world in which they do not (with irrational consumers not knowing the state).[7]

Spiegler (2005) can also be thought of as providing an explanation for add-on prices. One can think of the differences in product quality on different dimensions as representing both pure quality differences and differences in add-on fees charged in various circumstances.

6.6 Performance standards

How well an individual or firm has performed (or will perform) typically has a multidimensional answer. Performance assessments, however, often ultimately result in a zero or one decision. Should a student be admitted to this college? Should a student be given an A or a B? Should a job applicant be hired? Should a firm's contract be renewed? Should a paper be published? Should a professor be granted tenure? Should she be elected a fellow of the Econometric Society? These decisions establish performance standards. Two questions about standards are of primary interest: how will multiple dimensions be weighted and how high will the overall hurdle be?

In some applications, one could try to develop fully rational models of performance standards, e.g. a firm should hire a worker if and only if the firm's profits are higher with the worker than without. Often, however, no clear objective function is available, e.g. are colleges supposed to admit the smartest students, those who will earn the highest grades, those who will add most to campus life, those who will benefit most from the education, those who will go on to accomplish great things, or those who will eventually donate the most money to the college? Even when there is a clear objective, judges will typically lack the understanding necessary to maximize the objective, e.g. how does a student's GPA in math classes or his having had a particular position at the school newspaper predict

[7] I thank Jean Tirole for this observation.

subsequent job performance? As a result, the way that many such decisions are made is to compare the current candidate with past candidates, and to use relative performance to judge whether to admit the candidate.

Sobel (2000, 2001) uses a rule-of-thumb approach to explore how various factors affect whether standards tend to rise or fall. In these papers, candidates are assumed to choose a level of costly effort and to direct this effort to achieve a multidimensional vector of accomplishments. In the two-dimensional case, for example, the candidate's achievements are described by a pair (q, r). Candidates are then judged relative to the pool of recent successful candidates.

In Sobel (2000) the way the judging is done is that a single judge aggregates the multiple dimensions using a function $v(q, r)$, and deems the t^{th} candidate successful if $v(q_t, r_t) \geq z_t$. It is assumed that the standard z_t is set so that a fraction τ of the previously successful candidates would meet the standard. When the function v is held fixed and there is no randomness in the achievement process, this model predicts a bunching of performance. Starting from any initial condition candidates who are able to do so exert exactly the effort necessary to achieve the $1 - \tau$th percentile level of performance, and eventually the entire pool of "recent" sucessful candidates has achieved exactly this performance level. The most interesting results concern what happens when weighting functions change over time, e.g. shifting from v to v' and back to v again. Such shifts cause standards to decline. Intuitively, when standards shift from v to v', the performance level of the earlier candidates (measured in terms of v' can be achieved at lower cost by tailoring effort to v'. Hence, equilibrium effort declines. When the standards shift back to v, candidates can again achieve a performance that grades as highly as the performance of the candidates who had worked toward v' with less effort. Comparing the first and third cohorts of agents, performance will have unambiguously declined.

Sobel (2001) considers institutions involving voting by a panel of judges. Each judge has a different v function. Similarly to the previous paper, judges vote to acceept a candidate if the candidate's performance is in the top τ of recent successful candidates. What happens in this model depends both on the voting rules and the dimensionality of the performance space. When votes from only a small fraction of the judges are sufficient for admission, standards decline. When near unanimity is required, standards increase over time. In intermediate cases the results are more nuanced.

Ellison (2002) focuses on the weighting of different dimensions rather than the overall level of the standard. It is motivated by changes in the standards of academic journals over the last thirty years. There are several clear trends in economics (and other fields) over that period: papers have been getting longer; they have longer introductions, more references and discuss more extensions; authors are required to carry out more extensive revisions before papers are published. The paper's view is that all of these changes can be thought of as reflecting a shift in quality standards. Think of an academic paper as having a two-dimensional quality (q, r), where q reflects the importance of a paper's main contributions and r reflects other dimensions that can be improved with incremental effort, e.g. improving the exposition, generalizing and extending results, and performing robustness tests. It is argued that all of the observed trends can be thought of as caused by or reflecting an increased weight being placed on r-quality. Mathematically, suppose articles are judged acceptable if $v(q, r) = \alpha q + (1 - \alpha)r \geq z$. What we may be seeing is a decrease in α.

One way to account for such a trend would be to use the comparative statics of an rational model in which α is chosen to optimize some social welfare function. One can compute how the distribution of paper qualities is affected by α in equilibrium, and one could argue that the distribution produced by a lower α is now preferable because of some exogenous change in the profession. Such changes are hard to identify, however, and it is also hard to find any evidence that changes in standards were intentional. This motivates the search for explanations in which the shift is an unconcious byproduct of rule-of-thumb behavior.

My model focuses on the behavior of referees. Referees are assumed to try to faithfully apply the profession's standards (rather than imposing their own preferences over quality dimensions). They do not, however, inherently know what values of (α, z) to apply. They try to learn these by seeing what parameters best fit the data they get in the form of seeing decisions made on papers they refereed, seeing papers in journals, and seeing decisions on their own papers. In each period authors work to maximize their chance of acceptance given their current understanding of the norm, referees apply the standard they currently believe in, editors accept the fraction τ of papers that score most highly, and beliefs are updated given the acceptance decisions.

When referees are unbiased, this model has a continuum of equlibria. Given any weight α, there is a corresponding achievement level $z(\alpha)$ that yields an equilibrium. I then introduce a small behavioral perturbation, in a sense combining the rule-of-thumb and psychology and economics approaches. Players are assumed to be overconfident about the quality of their own work and think that it is ϵ higher in quality than it truly is. This destabilizes the former equilibria: players will be puzzled about why it is that their papers are being rejected. The first thing players would conclude if they started at an equilibrium of the behavioral-bias-free model is that z must be higher than they had thought. This model, however, also cannot fully explain the data. The result of players' continuing struggles to reconcile inherently contradictory data is a slow, gradual evolution of beliefs about the social norm in the direction of decreasing α. At the most abstract level, the idea of the paper is that one way to explain a long trend is to perturb a model with a continuum of equilibria. A slight perturbation of such a model can have a unique equilibrium and the disequilibrium dynamics can feature a slow evolution along the near equilibrium set.

7 Concluding Remarks

In this essay I've tried to give some perspective on the burgeoning field of boundedly rational industrial organization. It's a hard literature to summarize because one can depart from rationality in so many ways. The irrational actors can be the firms or the consumers. Several approaches to modeling irrationality can be taken: one can specify rules of thumb, behavior can be derived as the maximizer of something other than utility/profits; or one can explicitly introduce cognitive bounds. I've stressed two advantages of boundedly rational approaches: boundedly rational behavior seems more realistic in some applications; and the tractability of boundedly rational models can sometimes allow researchers to incorporate additional features into a model. For the next few years at least there is probably also an additional important advantage: after twenty-five years of focusing on rational models, the questions rational models are best suited to address have been much more thoroughly explored than other questions.

Which of the approaches to boundedly rational IO is likely to be most successful in the coming years? One key difference that seems relevant is that the psychology and economics motivated "behavioral" approach to is less behavioral than the earlier rule-of-thumb approach. The psychology and economics literature has been focused on writing down models involving objective functions, e.g. Laibson (1997), Fehr and Schmidt (1999), Eyster and Rabin (2005), in which behavior is endogenously determined by equilibrium conditions. The recent behavioral IO papers growing out of this literature share this approach. A big potential advantage of using objective functions as the primitive rather than behavior (as do rule-of thumb models) is that one may be able to calibrate the objective function experimentally or empirically, and then apply them to a variety of environments. Rule-of-thumb specifications are less portable. There is debate, however, about how portable behavioral objective functions really are, and the tractability benefits of the rule-of-thumb are absent because of the difficulty of doing equilibrium calculations. I think all of the bounded rationality approaches should and will continue to grow, but suspect that the psychology-and-economics motivated work will be most successful for practical reasons if no other: there is a clear matrix of papers waiting to be written; and the equilibrium derivations are an easy way to fill out an idea into a 35 page paper containing the kinds of calculations referees are used to seeing.

To conclude, I'd like to make one last set of remarks on the shift that has taken place from focusing on irrational firms to focusing on irrational consumers. For many reasons this is very appealing. There is a great deal of experimental evidence on consumer behavioral biases. Firms can hire consultants to advise them on how to maximize profits and market competition may tend to eliminate firms that don't maximize profits. I agree with all these point, but think there may still be good reasons to keep the boundedly-rational firm literature alive. I noted above that Baye and Morgan (2005) have shown that outcome of the Bertrand competition game is quite sensitive to whether firms are perfect- or ϵ-optimizers.

Although the mixed equilibrium they exhibit immediately strikes one as being a very special feature of this particular game, it is not clear that the sensitivity they are showing is at all unusual. As Akerlof and Yellen (1985) have noted, in smooth models, the derivative of a firm's profits with respect to its action is zero at the Nash equilibrium. If a firm changes its action by ϵ, profits will be within ϵ^2 of the best-response profits. A change in one firm's action does have a first-order effect on consumers and on the other firms' profits. Hence, in an ϵ-equilibrium, profits can be of order $\sqrt{\epsilon}$. For example, the simplest Cournot duopoly with $D(p) = 1 - p$ and zero costs any (q_1, q_2) with $q_i \in [1/3 - 2\sqrt{\epsilon}/3, 1/3 + 2\sqrt{\epsilon}/3]$ is an ϵ-equilibrium. This includes the monopoly outcome if $\epsilon > 1/64$, which is approximately 6% of the monopoly profit level.

Slight departures from traditional assumptions about consumer rationality can also have very large impact on economic models. Diamond's (1971) search model is a striking example in which prices go all the way from the competitive level to the monopoly level if consumers have an ϵ search cost. The evolving standards model of Ellison (2002) is an example of a situation in which an ϵ behavioral bias can have a big impact. In other models, however, this is not true: adding an ϵ mass of behavioral consumers to a Bertrand or Cournot model will probably only shift equilibrium play by an $O(\epsilon)$ amount. In such examples, small departures from rationality on the part of firms can be as important as much larger departures from rationality on the part of consumers.

I look forward to seeing boundedly rational IO develop in the years to come.

References

Armstrong, Mark (2005). "Recent Developments in the Economics of Price Discrimination," (eds.) *Advances in Economics and Econometrics.* Cambridge: Cambridge University Press.

Banerjee, Abhijit V. (1992). "A Simple Model of Herd Behavior," *Quarterly Journal of Economics* 107, 797-817.

Banerjee, Abhijit and Drew Fudenberg (2004). "Word-of-Mouth Learning," *Games and Economic Behavior*, 46, 1-22.

Baye, Michael and John Morgan (2004). "Price Dispersion in the Lab and on the Internet," *RAND Journal of Economics*, 35 (3), 449-466.

Baye, Michael, John Morgan, and Patrick Scholten (2005). "Information, Search, and Price Dispersion," *Handbook on Economics and Information Systems.* Elsevier.

Berk, Jonathan B. and Richard C. Green (2004). "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112 (6), 1269-1295.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992-1026.

Bose, Subir, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund (2005). "Dynamic Monopoly Pricing and Herding," mimeo.

Burdett, Kenneth and Kenneth L. Judd (1983). "Equilibrium Price Dispersion," *Econometrica*, 51, 955-969.

Chevalier Judith, and Glenn Ellison (1997). "Risk-Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105 (6), 1167-1200.

Cyert, R. M. and James G. March (1956). "Organizational Factors in the Theory of Oligopoly," *Quarterly Journal of Economics*, 70(1), 44-64.

Della Vigna, Stefano and Ulrike Malmendier (2004). "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 119 (2), 353-402.

Della Vigna, Stefano and Ulrike Malmendier (2005). "Paying Not to Go to the Gym," *American Economic Review*, forthcoming.

Diamond, Peter (1971). "A Model of Price Adjustment," *Journal of Economic Theory*, 3, 156-168.

Dufwenberg, Martin and Uri Gneezy (2000). "Price Competition and Market Concentration: An Experimental Study," *International Journal of Industrial Organization*, 18, 7-22.

Ellison, Glenn (1993). "Learning, Local Interaction and Coordination," *Econometrica*, 61 (5), 1047-1071.

Ellison, Glenn (2002). "Evolving Standards for Academic Publishing: A q-r Theory," *Journal of Political Economy*, 110 (5), 994-1034.

Ellison, Glenn (2005). "A Model of Add-on Pricing." *Quarterly Journal of Economics*, 120 (2), 585-637.

Ellison, Glenn and Drew Fudenberg (1993). "Rules of Thumb for Social Learning," *Journal of Political Economy*, 612-643.

Ellison, Glenn and Drew Fudenberg (1995). "Word-of-Mouth Communication and Social Learning," *Quarterly Journal of Economics*, 110 (1), 93-125.

Ellison, Glenn and Sara Fisher Ellison (2005). "Search, Obfuscation, and Price Elasticities on the Internet," mimeo.

Eyster, Erik and Matthew Rabin (2005). "Cursed Equilibrium," *Econometrica*, forthcoming.

Fehr, Ernst and Klaus M. Schmidt (1999). "A Theory Of Fairness, Competition, And Cooperation," *Quarterly Journal of Economics*, 114 (3), 817-868.

Fershtman, Chaim and Ehud Kalai (1993). "Complexity Considerations and Market Behavior," *RAND Journal of Economics*, 24(2), 224-235.

Gabaix, Xavier and David Laibson (2004). "Competition and Consumer Confusion," mimeo.

Gabaix, Xavier and David Laibson (2005). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," mimeo.

Grossman, Sanford J. (1981). "The Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24, 461-483.

Heidhues, Paul and Botond Koszegi (2004). "The Impact of Consumer Loss Aversion on Pricing," mimeo.

Hossain, Tanjim and John Morgan (2005). "...Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay," *Advances in Economic Analysis & Policy*, forthcoming.

Joskow, Paul (1973). "Pricing Decisions of Regulated Firms: A Behavioral Approach," *The Bell Journal of Economics and Management Strategy*, 4(1), 118-140.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986). "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 76(4), 728-741.

Kandori, Michihiro, George Mailath, and Rafael Rob (1993). "Learning, Mutation and Long Run Equilibria in Games," *Econometrica*, 61(1), 29-56.

Klemperer, Paul (1987). "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, 102, 375-394.

Koszegi, Botond and Matthew Rabin (2005). "A Model of Reference-Dependent Preferences," mimeo.

Laibson, David (1997). "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443-477.

Lal, Rajiv and Carment Matutes (1994). "Retail Pricing and Advertising Strategies," *Journal of Business*, 67, 345-370.

Lynch, Anthony W. and David Musto (2003). "How Investors Interpret Past Fund Returns," *Journal of Finance*, 58, 2033-2058.

Milgrom, Paul (1981). "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics* 12, 380-391.

Mobius, Markus (2001). "Death Through Success: The Rise and Fall of Local Service Competition at the Turn of the Century," mimeo.

Osborne, Martin and Ariel Rubinstein (1998). "Games with Procedurally Rational Players," *American Economic Review*, 88, 834-849.

Piccione, Michele and Ariel Rubinstein (2003). "Modeling the Economic Interaction of Agents with Diverse Abilities to Recognize Equilibrium Patterns," *Journal of the European Economic Association*, 1(1), 212-223.

Rabin, Matthew (2002). "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 117, 775-816.

Radner, Roy (1975). "A Behavioral Model of Cost Reduction," *Bell Journal of Economics*, 6(1), 196-215.

Rotemberg, Julio (2005). "Fair Pricing," mimeo.

Rothschild, K. W. (1947). "Price Theory and Oligopoly," *Economic Journal*, 57, 299-320.

Rubinstein, Ariel (1993). "On Price Recognition and Computational Complexity in a Monopolistic Model," *Journal of Political Economy*, 101 (3). 473-484.

Rubinstein, Ariel (1998). *Modeling Bounded Rationality.* Cambridge: MIT Press.

Schlag, Karl (2004). "Competing for Boundedly Rational Consumers," mimeo.

Schmalensee, Richard (1978). "A Model of Advertising and Product Quality," *Journal of Political Economy*, 86(3), 485-503.

Smallwood, Dennis E. and John Conlisk (1979). "Product Quality in Markets where Consumers are Imperfectly Informed," *Quarterly Journal of Economics*, 93(1), 1-23.

Sobel, Joel (1984). "The Timing of Sales," *Review of Economic Studies*, 353-368.

Sobel, Joel (2000). "A Model of Declining Standards," *International Economic Review*, 41 (2), 295-303.

Sobel, Joel (2001). "On the Dynamics of Standards," *RAND Journal of Economics*, 32 (4), 606-623.

Spence, A. Michael (1975). "Monopoly, Quality, and Regulation," *Bell Journal of Economics*, 6(2), 417-429.

Spiegler, Ran (2004). "The Market for Quacks," mimeo.

Spiegler, Ran (2005). "Competition over Agents with Boundedly Rational Expectations," mimeo.

Stahl, Dale O. (1989). "Oligopolistic Pricing with Sequential Consumer Search." *American Economic Review.* 79, 700-712.

Stole, Lars (2005). "Price Discrimination and Imperfect Competition," in (eds.) *Handbook of Industrial Organization.* North Holland.

Varian, Hal (1980). "A Model of Sales," *American Economic Review*, 70, 651-659.

Young, H. Peyton (1993). "The Evolution of Conventions," *Econometrica*, 61 (2), 57-84.