Bounded Rationality

Glenn Ellison

Massachusetts Insitute of Technology

Introduction

There is a long history of incoporating bounded rationality in IO

- In the 1940's the profit-maximizing firm was actively debated
- In the 1950's Simon highlighted bounded information processing
- In the 1970 "team theory" developed models of firms organized to efficiently process information
- In the 1970s Smallwood and Conlisk advocated for models with rule-of-thumb consumers over then-emerging rational models.

The more recent literature has focused on sophisticated firms exploiting boundedly rational consumers. Approaches include:

- **Rule of thumb**: Posit behaviors that seem reasonable and make models tractable
- **Explicit bounded rationality**: Maximize utility/profit net of computation costs. Search can be an example. See Spiegler's book.
- **Psychology and economics**

The psychology and economic branch has been the most active in recent years.

Behavioral Economics

Over the past 25 years behavioral economists have developed a large number of models to capture ways in which people depart from the rational ideal.

Many are portable models that can be plug-and-play replacements for utility maximization.

- Loss aversion
- Fairness preferences
- Quasi-hyperbolic discounting
- Imperfect Bayesian updating
- Self-serving
- Overconfidence

Here are some quick overviews of a few models:

Loss aversion

- In financial problems consumers maximize $E\,f(w(a) - w_0)$ rather than $E\,u(w(a))$.
- f overweights losses relative to gains, e.g $f(x) = \begin{cases} x & \text{if } x > 0 \\ 2x & \text{if } x < 0 \end{cases}$
- f sometimes also assumed to be concave for $x > 0$ and convex for $x < 0$.
- With multiple goods consumers sometimes assumed to have separately loss-averse preferences in each good, $\max_a \sum_j E\,f(x_j(a) - x_{j0})$.
- Reference point x_0 can be current allocation or equilibrium expected allocation or a random variable. See Koszegi-Rabin (*QJE* 2006).

Fairness preferences

- Fehr-Schmidt consumers $\max_a \pi_i(a) - \alpha \max(\pi_j(a) - \pi_i(a), 0) - \beta \max(\pi_i(a) - \pi_j(a), 0)$ with $\alpha > \beta > 0$.
- Bolton-Ockenfels consumers $\max_a v(\pi_i(a), \frac{\pi_i(a)}{\sum_j \pi_j(a)})$.

Quasi-hyperbolic discounting

- At time t consumer $\max_a u(c_t(a)) + \beta \sum_{s=1}^{\infty} \delta^s u(c^*_{t+s}(a))$.
- Consumers can have sophisticated, naive, or partially naive expectations about future behavior.
  - Sophisticated realize future selves are (β, δ) hyperbolics.
  - Naive think future selves are rational with discount factor δ, so future selves will carry out desired plan.
  - Partially naive think future selves will be $(\hat{\beta}, \delta)$ hyperbolics. The $(\beta, \hat{\beta}, \delta)$ model nests the previous two cases.

Imperfect Bayesian updating

- Eyster-Rabin posteriors are $E^{IB}(\theta|x) = \chi E(\theta) + (1-\chi)E(\theta|x)$.

Overconfidence bias

- Overestimate ability, e.g. underestimate prediction error.

Self-serving bias

- Like to think/act to make previous actions optimal.

Research Opportunities

1. One can think about topics/applications where some bias seems potentially relevant.
2. One can view the entirely of non-behavioral IO as just one column of a matrix of papers waiting to be written.

	Rational	Quasi-hyp.	Loss Averse	Fair	···
Monopoly	✓				
Price disc.	✓				
Durable goods	✓				
Horizontal diff.	✓				
Vertical diff.	✓				
Search costs	✓				
Dynamic comp.	✓				
Strategic inv.	✓				
Auctions	✓				
···	✓				

Della Vigna-Malmendier "Contract Design and Self Control ..."

Asks basic behavioral IO questions:

- How would a monopolist design products/sales mechanisms to exploit irrational consumers?
- How does competition affect exploitation?

Model is a multistage game motivated by health clubs:

1. Firm offers two-part tariff $L + px$ for x visits. Consumer with (β, δ) hyperbolic preferences says yes/no.
2. Consumer learns disutility $d \sim F$ of attending gym.
3. Consumer chooses $x \in \{0, 1\}$ and incurs disutility $(d + p)x$. Firm incurs cost cx.
4. Consumer gets delayed health benefit xb

Main monopoly observation is that p is distorted down for two reasons:

- A lower p increases the probability of going to the gym. A sophisticated hyperbolic consumer values this.
- A naive hyperbolic consumer overestimates gym usage. Charging upfront for a discount conditional on visiting exploits this mistake.
- In the sophisticated hyperbolic case p is socially efficient.

Della Vigna-Malmendier (*QJE* 2004)

Welfare statements are tricky in models with hyperbolic consumers. Here, efficiency means from the perspective of the firm and the time 0 consumer.

The efficiency of the contract for sophisticated hyperbolic consumer is a consequence of result we saw earlier that a monopolist chooses the optimal quality for the marginal consumer:

- Think of the price p as a dimension of quality for the contract.
- Consumers are homogenous so the marginal consumer and average consumer coincide. The monopolist chooses the contract that maximizes joint surplus and then extracts all of the surplus with the fixed fee L.
- More generally a monopolist serving a homogenous population will design the product to maximize the difference between consumer WTP and cost. Whether the WTP is utility or not doesn't matter.

Della Vigna-Malmendier note that we get the same distorted contract with perfect competition as with monopoly.

Again, it is a general result that competition does not eliminate exploitation. Perfect competition forces firms to maximize the gap between consumer WTP and costs.

Della Vigna-Malmendier "Paying Not to Go to the Gym" (*AER* 2006)

The paper is a reduced-form empirical exercise arguing that the form of health club contracts and consumer attendance is best explained by a hyperbolic discounting model.

The dataset includes 7,752 members at three Boston-area health clubs.

The data are inconsistent with several predictions of a fully rational model.

- Consumers who sign up for a monthly contract pay too much per visit. Those on $70 plans average 4.3 visits per month. There was a $100 ten-visit pass available.
- Consumers who sign up for the (auto-renewing) monthly plan are more likely to still belong in month 13 than customers who chose the (manually renewing) annual plan.
- Customers on average spend $187 after their last visit to the gym.

Some other observation suggest a partially naive hyperbolic model may be appropriate.

- In surveys members overestimate future monthly usage (9.5 vs. 4.2).
- Lower usage is correlated with being slower to cancel after the last visit.

Explore distortions in credit markets when consumers are quasi-hyperbolic and firms can offer teaser rates.

- An agent with $(\beta, \hat{\beta}, 1)$ preferences borrows c at $t = 0$ to buy a durable good that provides utility at $t = 1$.
- If the agent repays q at $t = 1$ and r at $t = 2$ the selves utilities are:

Time 0 self	c	$-k(q)$	$-k(r)$
Time 1 self	c	$-k(q)$	$-\beta k(r)$
Time 2 self			$-k(r)$

  where k is a convex function with $k(0) = 0$, $0 < k'(0) < \beta$, and $\lim_{x \to \infty} k'(x) = \infty$.
- The credit market is perfectly competitive with zero interest rate. Firms can offer contracts $c, \{(q_s, r_s)\}_{s \in S}$ consisting of an up front loan amount and a set of repayment options that can be chosen at $t = 1$.

If the consumer is rational with $\beta = 1$ then the optimal repayments q^{FB} and r^{FB} satisfy $k'(q^{FB}) = k'(r^{FB}) = 1 \Rightarrow q^{FB} = r^{FB}$. Competition implies $c^{FB} = q^{FB} + r^{FB}$.

The loan contract can be a simple contract with just one option.

Heidhues-Koszegi (*AER* 2010)

Proposition 1: When $\hat{\beta} = \beta$ the equilibrium contract is $c^{FB}, (q^{FB}, r^{FB})$. This is the optimal for the Time 0 self.

This is as in Della Vigna-Malmendier. Competition leads firms to offer the contract that maximizes the surplus of the Time 0 consumer. This contract lets the Time 0 self commit future selves to do what he wants.

Proposition 2: When $\hat{\beta} > \beta$ the competitive equilibrium contract is $c, \{(\hat{q}, 0), (q^{FB}, r')\}$ where $r' > r^{FB}$ satisfies $k'(r') = 1/\beta$ and $c = q^{FB} + r' > c^{FB}$. Consumers think they will choose the $(\hat{q}, 0)$ option, but actually choose (q^{FB}, r').

Note:

- The model predicts two distortions: Consumers borrow more than is optimal and repayments are shifted toward the second period.
- Firms offer seemingly attractive repayment plans that consumers don't take advantage of.
- Competition ensures zero profits, but exploitation of biases still hurts consumers.
- A slight departure from perfect sophistication, e.g. $\hat{\beta} = \beta + \epsilon$, can lead to large changes in outcomes and welfare.

Proposition 2: When $\hat{\beta} > \beta$ the competitive equilibrium contract is $c, \{(\hat{q}, 0), (q^{FB}, r')\}$ where $r' > r^{FB}$ satisfies $k'(r') = 1/\beta$ and $c = q^{RB} + r' > c^{FB}$. Consumers think they will choose the $(\hat{q}, 0)$ option, but actually choose (q^{FB}, r').

Proof:

We can restrict attention to offers $c, \{(\hat{q}, \hat{r}), (q, r)\}$ where time 0 consumers think their future selves will choose $(\hat{q}, \hat{r})$ but they actually choose (q, r). Firm competition to make the most attractive offer implies the contract will solve:

$$\begin{aligned} &\max\ c - k(\hat{q}) - k(\hat{r}) \\ &s.t. \\ IR:&\ \ c \leq q + r \\ IC_0:&\ \ k(\hat{q}) + \hat{\beta}k(\hat{r}) \leq k(q) + \hat{\beta}k(r) \\ IC_1:&\ \ k(\hat{q}) + \beta k(\hat{r}) \geq k(q) + \beta k(r) \end{aligned}$$

The *IR* constraint obviously holds with equality. The IC_1 constraint must also hold with equality. Otherwise we increase q, r. When IC_1 holds with equality, IC_0 is equivalent to

$$(\hat{\beta} - \beta)k(\hat{r}) \leq (\hat{\beta} - \beta)k(r) \Leftrightarrow \hat{r} \leq r$$

Proof: (cont'd)
The substitutions reduce the problem from the one on the left to the one on the right.

$$\begin{aligned} &\max\ c - k(\hat{q}) - k(\hat{r}) \\ &s.t. \\ &IR:\ \ c \leq q + r \\ &IC_0:\ \ k(\hat{q}) + \hat{\beta}k(\hat{r}) \leq k(q) + \hat{\beta}k(r) \\ &IC_1:\ \ k(\hat{q}) + \beta k(\hat{r}) \geq k(q) + \beta k(r) \end{aligned} \qquad \begin{aligned} &\max\ q + r - k(\hat{q}) - k(\hat{r}) \\ &s.t. \\ &IC_0:\ \hat{r} \leq r \\ &IC_1:\ k(\hat{q}) + \beta k(\hat{r}) = k(q) + \beta k(r) \end{aligned}$$

Substituting the expression for $k(\hat{q})$ in IC_1 into the objective function the problem becomes

$$\max q + r - k(q) - \beta k(r) - (1-\beta)k(\hat{r}) \quad s.t. \quad \hat{r} \leq r$$

Clearly we want to set $\hat{r} = 0$.

The FOCs for q and r then give $k'(q) = 1$ and $k'(r) = 1/\beta$.

Grubb "Selling to Overconfident Consumers" (*AER* 2009)

Grubb notes that many sellers seem to use three part tariffs with a fixed fee, a number of units offered at no charge, and then a high fee for excess unit, e.g.

Cell phone	\$39.99 per month	2 GB data free	Extra \$0.05 per MB
Car lease	\$12,000 fixed fee	36K miles free	Extra \$.29 per mile
Credit card	\$1.9% transfer fee	6 months free	19.8% APR afterwards

He suggests that this may be a result of firms exploiting an "overconfidence bias": consumers think that they can forecast usage better than they can.

Here is a simple illustrative example:

- Firms have fixed costs of \$50 and marginal costs of \$0.05 per unit. Can offer nonlinear tariff $T(x)$.
- Consumers get benefits of \$0.45 per unit for θ units where $\theta \in \{100, 400, 700\}$. All three are equally likely at the time of purchase. Consumers then learn θ and chooose how many units x to consume.

With rational consumers the monopoly plan is $T^m(x) = 160 + 0.05x$. Perfectly competitive firms choose $T^c(x) = 50 + 0.05x$.

Both are socially efficient.

Grubb (*AER* 2009)

Consider now a version with overconfidence:

- Firms have fixed costs of \$50 and marginal costs of \$0.05 per unit. Can offer nonlinear tariff $T(x)$.
- Consumers get benefits of \$0.45 per unit for θ units where $\theta \in \{100, 400, 700\}$. All three are equally likely at the time of purchase. Consumers then learn θ and chooose how many units x to consume.
- Overconfident consumers believe $\theta = 400$ with probability one.

The monopoly tariff is now $T^m(x) = \begin{cases} 180 & x \leq 400 \\ 180 + 0.45x & x > 400 \end{cases}$

Firms charge 0 for the first 400 units because consumers overestimate the probability of consumption. If you charge p you need to cut your fixed fee by $400p$. You only get $300p$ in extra revenue.

Firms charge 0.45 for units beyond 400. Lower prices lead to a revenue loss that cannot be recouped with a higher fixed fee. Higher prices will cause consumers not to purchase these units when $\theta = 700$.

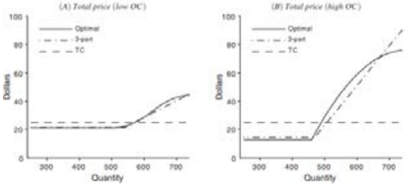
Perfectly competitive firms would use a similar tariff with a fixed fee of \$25.

Grubb (*AER* 2009)

Grubb does a number of things to turn this basic idea into an *AER* paper.

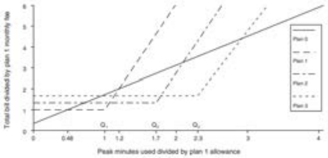
- Generalizes model to consider version where $\theta \sim F$, value is $v(q,\theta)$, and consumers believe $\theta \sim F^*$. Defines appropriate sense of overconfidence.

  With continuous F the solution is not exactly a three-part tariff, but has some similarities.



- Discusses extension with price discrimination where consumers get signal s, have posteriors $F^*(\theta|s)$ and the firm can offer a menu of nonlinear tariffs.
- Discusses alternate explanations with rational consumers and show price discrimination can result in similar tariffs but only with unusual preferences.
- Presents empirical estimates from a dataset containing 40 months of data on 2332 cell phone users.

The provider Grubb studies offers one two-part tariff and three three-part tariffs.



Observations:

- In a given month about half of the consumers on a plan that was not ex post optimal. Consumers on the two most popular three-part tariffs would on average have been better off on the two-part tariff.

	Plan 1 customers (60)		Plan 2 customers (61)	
	Bills (523)	Potential saving†	Bills (578)	Potential saving†
Underusage	56 percent	20 percent	59 percent	30 percent
Intermediate	28 percent	(8 percent)	33 percent	(16 percent)
Overusage	16 percent	30 percent	9 percent	23 percent
Total	100 percent	42 percent	100 percent	37 percent

- Overages account for about 23% of seller revenues.
- Usage is FOSD increasing in the plan chosen, but does not have the properties needed for the price discrimination explanation.

Heidhues-Koszegi-Murooka "Procrastination Markets"

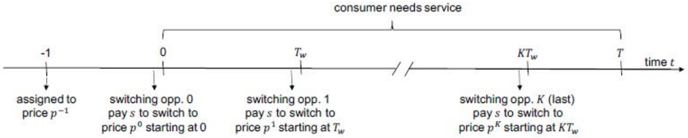
An applied motivation is to understand why minimal product differentiaion does not seem to lead to low prices in some markets with switching costs: retail electricity, health insurance, auto insurance, credit cards, cable TV.

Relative to the behavioral IO literature it explores the intuition that competition will return rents to consumers and the effects of flexibility in pricing.

In my diagram it fills the quasi-hyperbolic consumers - switching costs box.

- Consumers purchase a good at $t = 0, \Delta, \ldots, K\Delta$ with $T = (K+1)\Delta$.
- N firms just move once. They simultaneously choose offers $(p_I^n\ ,\ p_S^n)$ that apply throughout to new and switching consumers. Must be at most v.
- At $t = -1$ consumers are exogenously randomly assigned to one of the new-consumer offers.
- At $t = 0, \Delta, \ldots, K\Delta$ consumers receive one switching offer selected at random. They can pay a switching cost of s to accept this price in place of their previous price. $\Delta = \frac{1}{m(N-1)}$, where m is the rate at which firms send switching offers.
- Consumers have (β, δ) hyperbolic preferences with $\delta = 1$ for simplicity. Switching costs are treated as incurred in the "present". All prices are paid in the "future".

Heidhues-Koszegi-Murooka "Procrastination Markets"



From "Procrastination Markets", Feb. 20, 2021. Courtesy of Paul Heidhues, Botond Koszegi, and Takeshi Murooka. Used with permission.

- Consumers purchase at $t = 0, \Delta, \ldots, K\Delta$ with $T = (K+1)\Delta$.
- N firms once. Choose offers $(p_I^n\ ,\ p_S^n)$ that apply throughout to new and switching consumers. Must be at most v.
- At $t = -1$ consumers are exogenously randomly assigned to one of the new-consumer offers.
- At $t = 0, \Delta, \ldots, K\Delta$ consumers receive one switching offer selected at random. They can pay s to accept this price. $\Delta = \frac{1}{m(N-1)}$
- Consumers have $(\beta, 1)$ hyperbolic preferences. Switching costs are treated as incurred in the "present". All prices are paid in the "future".

Consumers may fail to switch to better offers for two reasons: (1) They overweight the switching cost relative to the later price savings; and (2) They may wrongly think that they'll switch next period anyway.

Let $p^{NIS} = \frac{1}{\beta}\frac{s}{T}$. If $p_I < p^{NIS}$, then $\beta T p_I < s$ so a consumer would not pay s to switch to a firm offering a zero price.

Let $p^{IP} = \frac{1-\beta}{\beta} m(N-1)s$. If $p_I < p^{IP}$, then $\beta p_I \Delta + \beta s < s$ so a consumer would prefer to wait and switch at the next opportunity.

Proposition

The model has an equilibrium in which all firms set $p_S^n = 0$ and $p_I^n = Min\{v, Max\{p^{NIS}, p^{IP}\}\}$ and no consumers ever switch. All equilibria have this initial price and no switching.

Observations:

Prices can be very high in this equilibrium.

Prices increase in number of firms (with advertising intensity held constant). The perceived benefit of putting off switching is assumed to remain constant, and the perceived cost from paying high prices until the next offer arrives gets smaller.

Heidhues-Koszegi-Murooka "Procrastination Markets"

One extension adds competition to be the initial supplier:

- Consumers receive initial offers from N_I firms. Signing up with one provides ϵ utility at $t = -1$.
- N_S firms then send offers to switch at $t = 0, \Delta, \dots, K\Delta$.

Proposition

Suppose $v > p^{IP} > p^{NIS}$ and $\frac{N_I}{N_S} < \frac{(1-\beta)m}{\beta}$. Then, for any $\epsilon > 0$ the model has an equilibrium in which all firms offer $(p_I^n, p_S^n) = (p^{IP}, 0)$ and consumers never switch.

Observations:

Competition to be the initial provider fails to reduce prices and transfer surplus to the consumers.

The small deviations that lead to $p = 0$ in the Bertrand model don't have the same effect here: consumers think they'll switch away from any offer with a high initial price at $t = 0$, so they only pay attention to the sign-up benefit.

Heidhues-Koszegi-Murooka "Procrastination Markets"

Some other extensions:

1. Endogenous advertising levels

   Assume firms send new-consumer and switching offers at the same rate m.

   Outcomes are discontinuous in the cost of sending offers.

   With high ad costs, firms send few ads and the equilibrium introductory price is p^{NIS}. When costs are below a threshold. firms send many more ads and prices jump up to p^{IP}.

2. Flexible pricing

   Suppose $N = 2$ and firms can include sign up bonuses in offers.

   Prices can end up even higher than in the base model.

   Consumers take the offer with the worst price first, assuming they'll collect the bonus and then switch to the better price. Overcutting incentives can push p above p^{IP}.

   Bonuses are just s and don't dissipate profits.

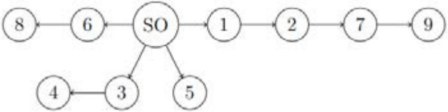
   The flexibility leads to exploitation message differs from the rational model, but is consistent with other behavioral papers.

Antler “Multilevel Marketing: Pyramed-Shapes Schemes or Exploitative Scams?”

Develops a model of multilevel marketing firms like Amway, Herbalife, and Mary Kay in which people can become distributors and and earn money from (1) selling products, (2) recruiting new distributors, and (3) sales and recruitments by downstream distributors.
The model is a random network model:

- The organizer commits to price/compensation scheme including fees for buying the product, becoming a distributor, and commissons for sales made directly, by your direct offspring, by your level-2 offspring, etc.
- One new agent arrives at each $t = 0, 1, 2, \ldots, n$. Each is randomly interested/not interested in purchasing and quiet/talkative. Each connects to one random agent who arrived previously.
- They are offered the chance to purchase if the person they connect to is a distributor or talkative and connected to a distributor through a chain of talkative people.
- They are offered the chance to become a distributor if the same is true and the distributor they reach wants to sell a distributorship (which increases future offspring sales but reduces future direct sales).

Antler “Multilevel Marketing: Pyramed-Shapes Schemes or Exploitative Scams?”



Antler discusses both a model with rational consumers and a model with consumers with beliefs as in Jehiel’s “anology based equilibrium”. These consumers believe others will purchase/buy distributorships/make offers with the correct average probability, but do not realize that these vary with t.

The bounded rationality matters because consumers are less likely to purchase a distributorship later in the game.

Some observations are:

- With no purchases, no distributorships are sold in a rational model.
- With no purchases and analogy-based consumers, the organizer can earn positive profits. All profitable schemes involve offspring commissions.
- With purchases and rational consumers there is a profit-maximizing scheme that does not involve selling distributorships.
- With purchases and analogy-based consumers, the organizer only sells distirbutorships when n is small.

On Wednesday I'll discuss some work on bounded rationality with rule of thumb consumers and cover some basic advertising theory.

If I have time I'll come back to Heidhues, Koszegi, Murooka.

See you then!