Bounded Rationality

Glenn Ellison

Bounded Rationality: Other approaches

Using theoretical models from the psychology and economics literature is now most common, but other approaches are also seen:

1. Rule of thumb consumers
2. Computational costs

Potential advantages of these approaches include that we can capture the behavior we think is relevant in a more tractable manner and think similarly about boundedly rational consumers and firms.

Rule of Thumb Consumers

Smallwood and Conlisk (*QJE* 1979) argued that rule-of-thumb models were more plausible and fruitful than the then emerging models of information acquisition by "rational" consumers.

Consider a consumer buying cereal at a supermarket. The consumer will not know how much they like the taste of most products. Suppose also that they observe the consumer in front of them pick a box off the shelf. What will they do?

- The prior consumer's decision could signal how much they like cereal.
- It could also reflect what they saw a previous consumer choose, whether they saw a TV commercial for the product, and the convenience of the shelf the box was on.
- The previous consumers action could in turn reflect what they saw, the choice to purchase a TV commercial could signal something about the seller's confidence that consumers will like the product, the shelf placement could reflect aggregate popularity or payments for shelf space, which are themselves signals.
- Do we really want to work with models in which consumers have priors over what every consumer has seen, purchased, and experienced since the dawn of the cereal industry?

Smallwood and Conlisk, "Imperfectly Informed Consumers"

Suppose a continuum of consumers choose among K brands at $t = 0, 1, 2, \ldots$

- Products differ in quality. A consumer of product k has a bad experience "breakdown" with probability b_k.
- Consumers buy the same product they bought previously until they experience a breakdown. When this occurs they buy k with probability $m_k(t)^\sigma$, where $m_k(t)$ is product k's market share at t.

All of these models involve naïve consumers. $\sigma = 0$ is purely random choice. Copying the consumer in front of you would be $\sigma = 1$. Larger σ could reflect recognition of the information value of popularity.

S&C's main result is that social learning can work well despite the naivete:

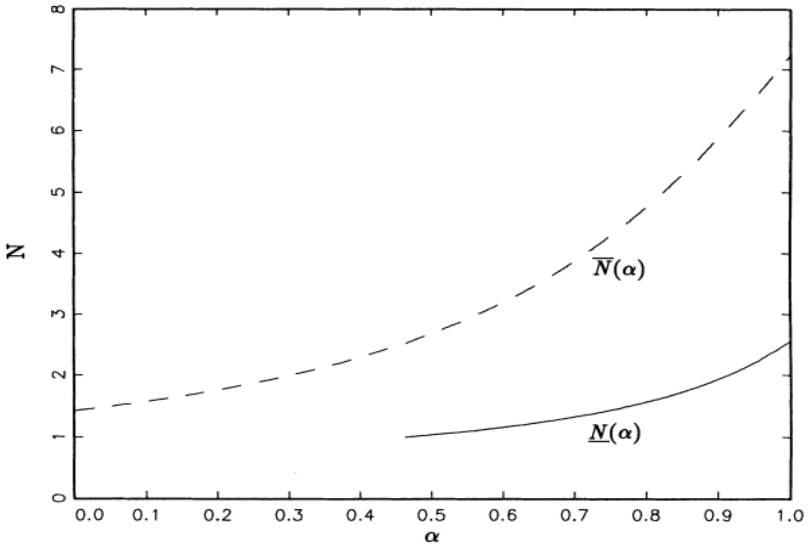
Propositon: When $\sigma = 1$ the most popular product dominates in the $t \to \infty$ limit. When $\sigma < 1$ we converge to state with all K products active. When $\sigma > 1$ an inferior product can come to dominate the market.

Ellison and Fudenberg , "Word of Mouth Communication and Social Learning," *QJE* 1995

Consider a two-product model like S&C's but with utilities:

- Consumer i's utility if they consume product *k* at t is $u_{ikt} = \pi_{kt} + \varepsilon_{ikt}$. Suppose $\pi_{1t} - \pi_{2t} \in \{-\theta, \theta\}$ is iid equal to θ with probability p and ε_{ikt} is normally distributed.
- Consumers buy the same product they bought previously with probability $1 - \alpha$. With probability α they ask N friends about their most recent purchase and buy the product that gave the higher average utility. (If all friends bought the same product they buy that.)

Observation: We sometimes get optimal social learning in this model. Typically, this occurs when N is fairly small. The all-friends-bought-the-same possibility creates some popularity weighting.



Spiegler, "The Market for Quacks," *REStud* 2006

Spiegler introduces pricing into a K product word-of-mouth model:

- Consumer i's utility from product k at t is $u_{ikt} = \begin{cases} 1 - p_{kt} & \text{w prob } \alpha \\ -p_{kt} & \text{w prob } 1 - \alpha \end{cases}$.
- Consumers ask one user of each product about their 0/1 experience and buy the product would give the highest utility assuming the same experience if it gives positive utility (otherwise they choose the outside option.)

Observations:

1. The model has a symmetric mixed strategy equilibrium.
2. Prices are decreasing in quality α.

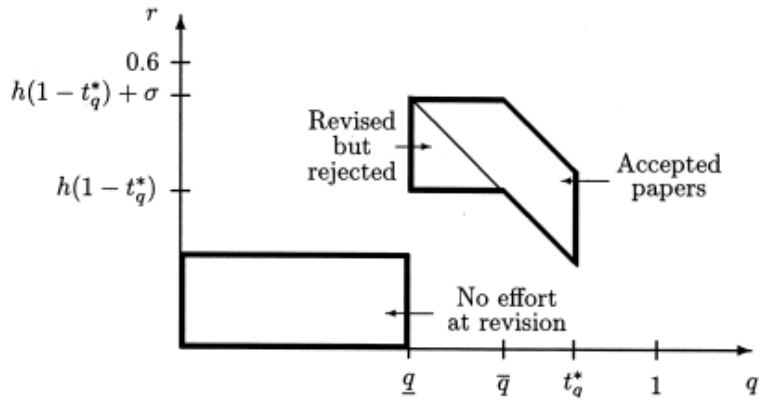
Ellison, "Evolving Standards for Academic Publishing: A q-r Theory," *JPE* 2002

Consider the problem of journal referees deciding whether to recommend that a journal publish paper.

- Papers have two-dimensional quality (q, r). q captures importance of main idea. r can be improved via revision.
- Referees know they should recommend acceptance if $\alpha q + (1-\alpha)r \geq z$, but only learn α and z by experience. At time t use α_t and z_t.
- Authors rationally divide their time between producing the two types of quality. They start by spending $t_q \in [0,1]$ to get a random draw $q \sim F(q|t_q)$. Authors choose $t_r \in [0,1-t_q]$ in response to reports, generating $r = h(t_r) + \eta$ with $\eta \sim U[0,\sigma]$.
- Editors accept the highest-quality papers that are resubmitted. They fill the journal by accepting a fraction τ of papers.

Ellison, "Evolving Standards for Academic Publishing"

The model has a continuum of consistent social norms. For any quality weight $\alpha \in [0,1]$ there is an overall threshold z that just fills the journal.



Suppose that referees are trying to learn (α, z) from two types of observations:

- They observe editors yes/no decisions on papers they refereed.
- They observe the requested r(q) on papers they submit.

They estimate parameters with a loss function combining a sum-of-squared residuals on r(q) and a sum of absolute deviations on acceptances. They update $\begin{pmatrix}\alpha_{t+1}-\alpha_t\\ z_{t+1}-z_t\end{pmatrix} = k\begin{pmatrix}\hat{\alpha}_t-\alpha_t\\ \hat{z}_t-z_t\end{pmatrix}$

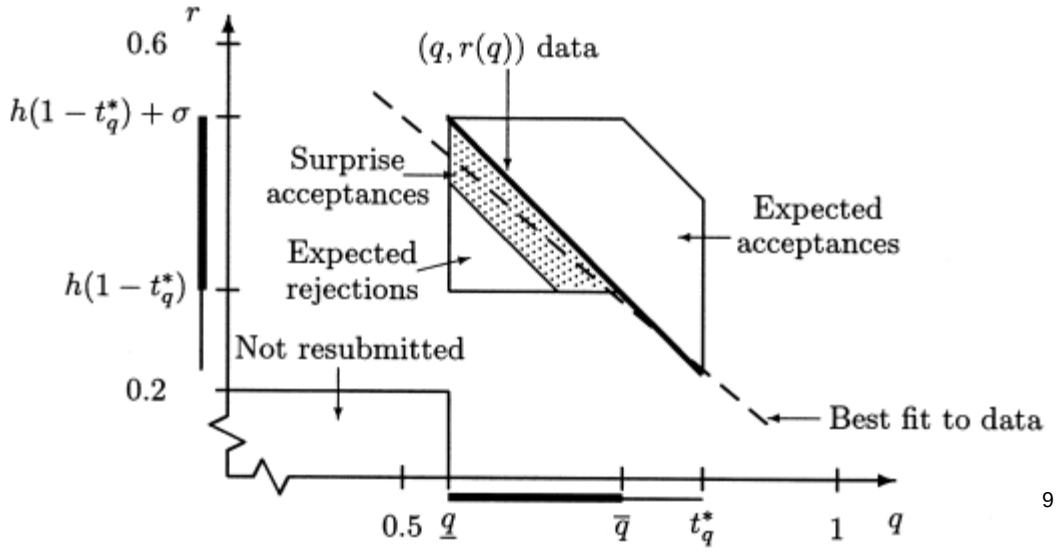
Referees have an overconfidence bias and think own papers are ε better than they are.

Ellison, “Evolving Standards for Academic Publishing”

With no bias the model could converge to any point on the continuum of consistent social norms.

With the bias, referees always hold authors to a slightly too-high standard and reconcile the unexpected acceptances of some papers they refereed by inferring that the weight α on q quality is slightly lower than they thought.

The result is a long, slow evolution toward lower and lower α.



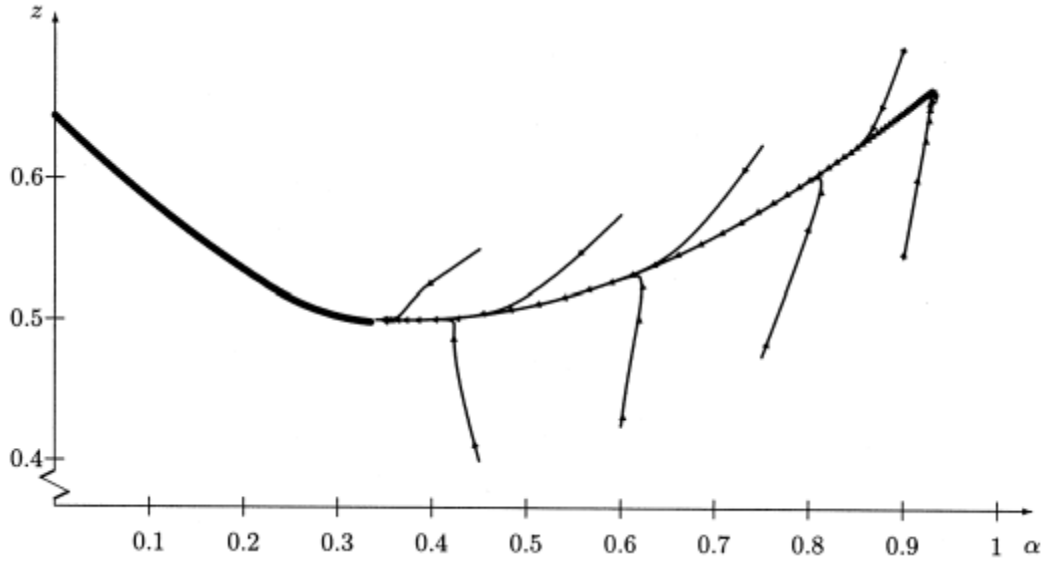
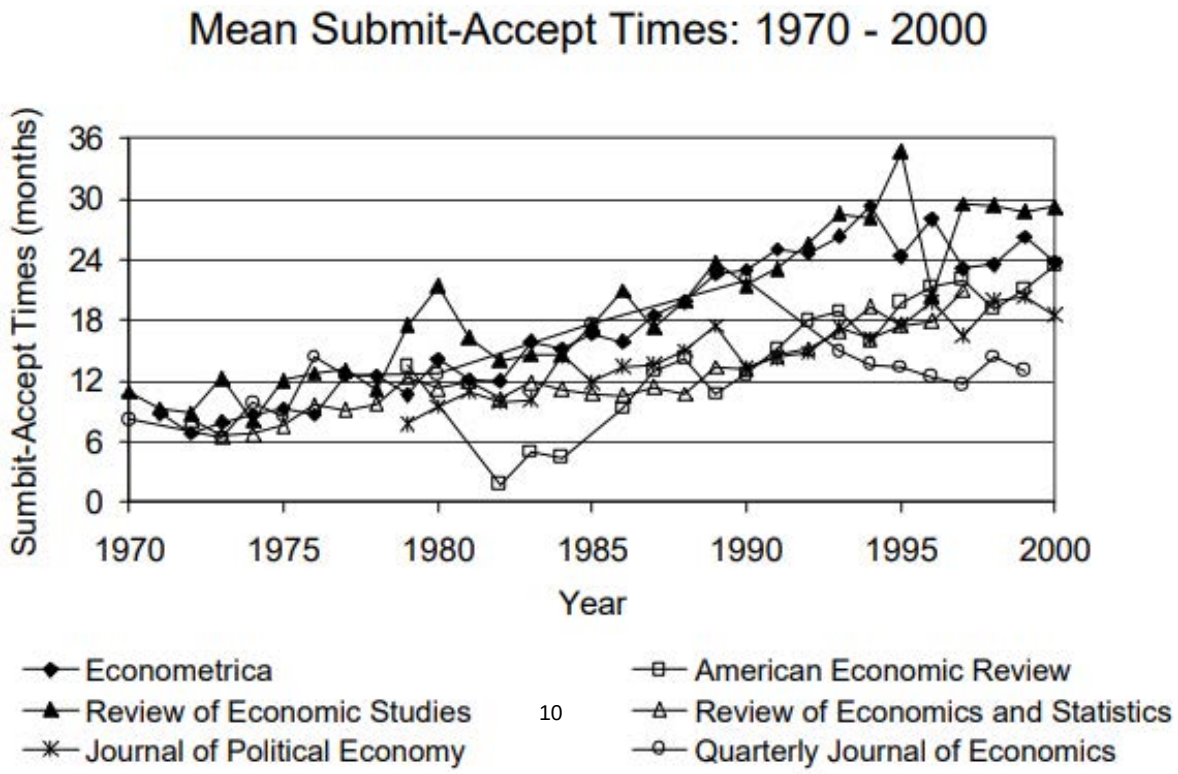


FIG. 9.—The disequilibrium evolution of social norms when $\epsilon > 0$

Ellison, "Evolving Standards for Academic Publishing"

In a related paper (*JPE* 2002) I noted that submit-accept times for economics papers rose dramatically from 1970-2000.

A depressing update is that the median time that star IO students have needed to publish their job market papers is now 7 years.

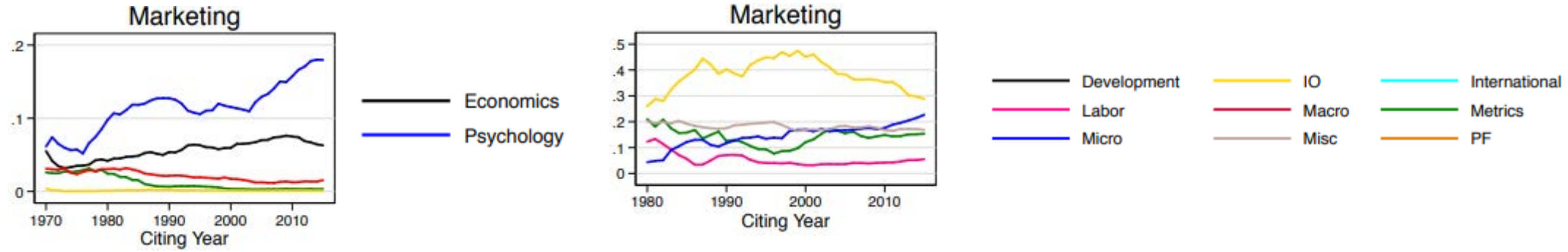


Advertising

Glenn Ellison

Advertising and Marketing

Marketing is an enormous field. Parts are more influenced by psychology, but there is a lot of work that looks like theoretical and empirical IO. A number of our IO PhD students have taken marketing jobs.



I can't cover much of the field in 1.5 lectures. So what I'll do is to go over some classic and not-so-classic models of how economists think about advertising and cover a few empirical papers.

Online advertising has been a particularly active overlap area recently.

Models of Advertising

The theoretical literature on advertising has noted many distinct mechanisms through which it could affect purchasing.

- Advertising can have informational content, e.g. it can inform consumers about a product's existence, its price, or its attributes.
- Advertising can signal something about a product like quality.
- Advertising can change consumer preferences or directly affect the purchase decisions of nonrational consumers.

Butters, "Equilibrium Distribution of Prices in Advertising," *REStud* 1977

Butters is a classic model of informative advertising:

- N firms sell identical goods
- Unit mass of identical consumers get with utility $v - p$ if buy
- Cost c of production
- Advertising costs $A(x)$ to reach random fraction x of population

Considers a game where

- Firms simultaneously choose x_i and p_i.
- Consumers who receive at least one ad buy from the cheapest firm j they see if $v - p_j \geq 0$.

Butters, "Equilibrium Distribution of Prices in Advertising"

Observations:

- The model has an equilibrium with price dispersion as in Stahl, Spiegler, etc. Firms mix over p because they want to set $p = v$ if consumers see only their ad, but benefit from being lower than others when consumers see multiple ads.
- Interior symmetric pure strategy equilibrium exists for x^* if $A'(0)$ small and $A'(1)$ very large.
- Advertising levels x^* are socially efficient.

The argument for efficiency uses that firms are indifferent over all prices. Hence the FOC for x^* gives

$$A'(x^*) = (v - c)Prob\{\text{Consumer reached gets no other ad}\}$$

cost of reaching one more consumer

profit from reaching to one more consumer

The RHS is the social value of an advertisement, so ad levels are socially optimal.

Milgrom and Roberts, "Price and Advertising Signals of Product Quality," *JPE* 1986

Milgrom-Roberts is a classic signaling model:

- At t=0 Nature chooses quality $q \in \{L, H\}$ of a firm's product.
- At t=1 the firm observes q chooses (p, A) paying cost A.
- At t=2 consumers observe A, decide whether to buy, learn quality if they do buy, and then can buy again. The combination of these various decisions gives the firm profits $\pi(p; q, \hat{q})$ that depend both on the true quality q and the quality $\hat{q}$ consumers expect to receive when making the first purchase decisions.

Observations:

1. The model can have a separating equilibrium in which only firms with high quality advertise. Little word-of-mouth spread and more repeat purchases help support this.
2. Though technologically "money burning" the advertising can increase social welfare if it improves consumer-product match quality.
3. In some models like this firms can also signal quality purely through pricing.

Shapiro, "A Memory Jamming Theory of Advertising," 2006

Shapiro has an interesting take on advertising changing preferences: he considers the possibility that it can do this by changing peoples memories. The paper is well motivated by psychological evidence.

- A continuum of consumers have types $\theta_i \sim F$ on [0, 1]. Type θ consumers get benefit v with probability θ when they consume the good.
- Consumers have consumed on N prior occasions and gotten these draws.
- Consumers don't know θ_i. They estimate it by recalling one or more prior times when they consumed and using Bayes rule to form $E(\theta_i|r_i)$.
- Models advertising as potentially altering memory/recall in one of two ways:
  - An ad could replace a true memory in a consumer's mind with planted positive memory.
  - Ads could make positive memories more likely to be recalled than negative ones.
- Mostly focuses on "rational" case where consumers are aware that their memories are being manipulated and try to offset the manipulation.

Shapiro makes several observations about how manipulation depends on technologies/parameters and welfare effects. Issues are like those in info design.

On Monday I'll discuss auctions. It will mostly be theory, but I'll also cover

- Hendricks and Porter

See you then!

14.271 Industrial Organization I
Fall 2022