14.271: Industrial Organization I

Summary of the Readings
Fall, 2006

Contents

Part I
Before Midterm

1 Monopoly Pricing and Durable Goods

1.1 *Are Durable Goods Consumers Forward Looking?*, Chevalier, Goolsbee

Objective

Do publishers revise textbooks to kill off the secondary market for used books? This behavior would not be profitable if forward-looking consumers take the resale price into account when making purchasing decisions - in that case, purchases of used books would translate into higher prices for new books. If consumers are myopic or have extremely high discount rates, revisions may increase profits.

Data Source - Textbooks

Why are textbooks ideal?

(1) Little scope for delaying purchases in response to expected changes in textbook characteristics - demand is based on the book being assigned this semester.

(2) Quality differences are readily observable to students, so no adverse selection.

(3) College bookstores do not sell older editions, so there is no complicated decision as to whether to buy the new or old version.

(4) Student has no cross-price elasticity between different books.

(5) Most textbooks sold at college bookstores.

(6) Textbooks are an important category of spending for students so they should take these decisions seriously.

Actual data from a consulting company that collects data from college bookstores includes textbooks in economics, biology, and psychology from 1698 schools from 1997-2001.

Also, data on textbooks assigned and the number of students enrolled in each course, and student characteristics at these schools.

Problems controlled for: missing enrollment data, bundled goods

Results and Analyses

"Prices remain fairly constant over the life of the edition while the probability of revision varies dramatically; thus, the purchasing behavior of a forward looking consumer should change over the life of an edition, but that of a myopic or high discount rate consumer should not."

Evidence: regression of prices on linear time trend or different dummy for each time period - prices rising with inflation over time (not at all for books over $40); used book price is 75% of new book price, buyback price is 50% of new price

"Consumers are forward looking - the magnitude of responses is what we would expect if they have ordinary discount rates and take into account the resale price when making purchasing decisions"

First need to estimate $\Pr[\text{Revision}]$ to compute expected resale price $(1 - \Pr[\text{Revision}]) \cdot u \cdot P$, where u is fraction of original price obtained at resale.

$\Pr[\text{Revision}]$ is estimated using Hazard model which controls for age, intro and category (no interaction terms?)

Next, use logit demand framework:

$$U_{ijt} = \begin{cases} x_{jt}b - a\left(P_{jt} - d\left(1 - P_{DIE}\right)up_{jt}\right) + K_{it} + e_{ijt} & \text{if buy new book} \\ f\left(x_{jt}\right) + e'_{ijt} & \text{if buy used book} \\ e''_{ijt} & \text{if buy no book} \end{cases}$$

With e, e', and e'' independent, this allows you to use quantity purchased as the dependent variable.

Note that du not separately identified. But if we assume that $u \in [0.5, 0.75]$, we can infer values for the discount rate.

Instrument for price using the share of books published by a non-profit publisher (these charge systematically lower prices)

Estimated discount rates between 0.5 and 0.75 depending on precise specification.

"Simulations based on the estimated demand system suggest that publisher revenues would remain constant or fall if the revision cycle were accelerated."

2 Price Discrimination

2.1 *Price Discrimination and Retail Configuration*, Shepard

Objective:

Shepard tests whether price discrimination can persist in extremely competitive markets. She compares pricing of full-service or self-service gas at gas stations that offer both types of service with those that offer only one type of service. Standard models of price discrimination predict that with free entry prices will be equal to cost and thus the same within service type across stations, whereas if firms can price discriminate then multi-service stations should charge more for full-service gas if, by doing so, they can target a group of customers with higher willingness to pay.

Data:

Data comes from 1,527 gas stations in four counties of eastern Massachusetts over a 12 week period in 1987. It includes information about the brand of gas station, physical characteristics of the station, and (specifically) the prices and service types (full/self) of gas at all the stations.

Model:

Shepard uses a fairly standard model of consumer tastes, where utility of consumers is

$$U = \begin{cases} V(g)(t_i - p_g) & \text{if she consumes one unit of service level } g \\ V(o)\, t_i & \text{if she does not purchas} \end{cases}$$

where $g \in f$ (full service) $, s$ (self service) and $t_i \sim U[0,1]$. She then solves for the demands of each consumer in the standard Vertical Competition on a line method (find t such that consumer is indifferent, etc) and calculates the optimal price for full and self serve gas in multi-service (MP) and single service stations (SP). She then shows that, if multi-service stations are price discriminating, they should charge weakly more for full-service gas and weakly less for self service gas than the single service gas stations.

Results:

The primary result comes from a regression of price in station i, of service g, in market j,in a station of type (multi/single) k on dummies for full service (D_g) and multi-service gas station:(D_k):

$$p_{ikgj} = \beta_0 + \beta_1 D_g + \beta_3 D_k D_g + \phi \mathbf{X}_{ikg} + \varepsilon_{ijkg}$$

where $\mathbf{X}$ sometimes contains market fixed effects. Shepard finds that "*on average, the price differential [between full and self-serve] at multiproduct stations is 9-11 cents higher than the differential across single-product stations*". Controlling for market fixed effects only makes the effect stronger.

Shepard is careful to note that any cost differences between gas stations would bias the results in the opposite direction, towards less of a difference between full and self-serve at multi-service stations than single service type stations. It is interesting to note that she does not attempt to measure these costs directly, an example of the distrust of cost data that is often found in I.O. papers.

3 Static Competition and Models of Differentiation

3.1 *Games Economists Play: A Noncooperative View*, Fisher

Executive Summary

Since the 1970s everybody in IO has jumped on the game theory bandwagon. Fisher does not think that this has gotten us any closer to a "generalizing theory" of oligopoly behavior than we were in 1950. The reasons include that the game theoretic framework, especially normal form, is not the most illuminating way to learn about how context affects oligopoly interactions. Also, theorists have been too focused on analyzing Nash equilibria in cute "exemplifying" models. While much of this work has been useful (one way to get to a "generalizing theory" is by distilling general principles from the stripped down models), too much has been simply an exercise in elegant analysis, rather than an attempt to explain the real world in terms of a unified theory linking structure to conduct to performance.

Slightly more detailed summary

Recently (i.e. from 1970s) oligopoly theory has been dominated by game theoretic approach–this is not warranted, and has not gotten us very far. To see this, let's go back in time:

Brief History of Oligopoly Theory:

- Cournot started it all, but no good, because why would firms only set quantity?
- So along comes: Stackelberg, Bertrand, Edgeworth->messy collection of cases.
- In response, Chamberlin (1933): firms will reach cooperative "joint maximization" solution. Of course this has problems when incentives to deviate are strong. Nevertheless, the impact was oligopoly theory has since tended to focus on join maximization is achieved or not and why.

In short, by 1950, not much satisfactory theory, only messy collection of cases.

Now after having applied game theory:

Economists are still focusing on Cournot as a Nash equilibrium (in fact, academics are testifying in antitrust cases that we should look at markets in terms of Cournot). Ie., we have not made much progress–Cournot himself knew all this.

Problems with Nash with more than two oligopolists: it's not the only appealing concept: Bernheim (1984) and Pearce (1984): Nash only a subset of rationalizable equilibria.

Another problem: focus has been on one-shot games, not realistic. So move to repeated games.Nash suffers from folk theorem, so, refinements, but still too many equilibria. So focus on pareto efficient equilibria: this leads us right back to Chamberlin, and the state of the world in 1950!

So, what do we want a theory to do?

There are "generalizing theories" (e.g., welfare theorems, slutsky) that tell us, given wide assumptions, what must happen. Also, "exemplifying theories" (typical stripped down models -"MIT" style) that tell us what can happen in special cases. "Exemplifying" theories is where IO stands–we need a generalizing theory about how the context an oligopoly is in influences the outcome.

Example of exemplifying theory at it's best (worst): IBM anti-trust case. Fisher testified and wrote that "premature announcements" can't be anti-competitive (unless made in bad faith). Farrell and Saloner (1986), however, produced a counter example, where early truthful announcements can be anti-competitive. But the problem with this and other exemplifying stories is that they will stay simply stories unless they point to a generalizing principle.

So what should a generalizing theory do for us? Fisher's example: an old teacher of his claimed you can't understand the rubber tire industry without knowing that Harvey Firestone was an aggressive guy who believed in cutting prices. However, a generalizing theory would tell us why it was that this aggressive strategy was successful. But we don't have such a theory yet. Instead we have:

Example of bad state of affairs: the Merger Guidelines used by the Department of Justice: oppose mergers which increase a Herfindahl-Hirschman index already over 1,800 by at least 100 points. (impossible to know if this is sensible).

Two reasons why game theory is not doing what we need:

1. Normal form games are awkward for intuition and general principles

2. Existing literature is mostly on one-shot games, not realistic ones. Until economists stop doing cute, elegant models for the models' sake, and instead focus on developing a general theory for the real world, economists will just be playing games.

Funny remarks:

- While it may be true that teaching Latin is useful for systematizing thought, the same can't be said for game theory.
- When discussing what he suspects is a generalizing principle about oligopolies with multi-market contact, he refers to Shakespeare's comment in Macbeth about how drink provokes desire, but hampers performance, but multimarket contact provokes desire and assists the performance (of collusion).

3.2 *Competition and Collusion in the American Automobile Industry*, Bresnahan

Objective

Auto sales in 1955 were 45% higher than in the two surrounding years, an increase that could not be attributed to macroeconomic conditions alone. (Samuelson threatened to flunk anyone who tried to provide an explanation for this.) In one of the first ever structural estimation I/O papers, Bresnahan credits this increase in sales to a fundamental difference in the nature of competition in 1955 as compared to the years 1954 and 1956. It should be noted that Bresnahan does not speculate as to why the nature of competition changed in 1955.

Data Source and Issues

Data were collected for the years 1954, 1955, and 1956. The data come from contemporary trade publications. There are difficulties determining what constitutes a separate product, so some assumptions were made that yielded 80-85 different products for the years estimated. The prices he uses are "list prices as of mid-April in the model year." Note that a model year begins the fourth quarter of the previous year. (i.e. the 2007 model cars are already out.)

In a footnote, Bresnahan writes, "The period 1954-56 is not entirely arbitrary. The Korean war price and quantity controls were lifted in February/March 1953. Foreign competition of any consequence begins with Volkswagen's entry in 1957. No labor-based work stoppages lasted more than ten days within the period."

The Model

This paper uses the standard vertical differentiation framework. Suppose there are N cars being offered each year. Each car i is assigned a quality, x_i based on a real valued function of the observable characteristics (weight, length, horsepower, cylinders, and whether or not the car had a hardtop). The demand side of the model is given by (assuming individual tastes are distributed according to a $U\left[0, v_{MAX}\right]$ distribution.)

$$\begin{aligned} q_1 &= \delta\left[\frac{P_2 - P_1}{x_2 - x_1} - \frac{P_1 - E}{x_1 - \gamma}\right] \\ q_i &= \delta\left[\frac{P_{i+1} - P_i}{x_{i+1} - x_i} - \frac{P_i - P_{i-1}}{x_i - x_{i-1}}\right] \\ q_N &= \delta\left[v_{MAX} - \frac{P_N - P_{N-1}}{x_N - x_{N-1}}\right]. \end{aligned}$$

The variable cost side of the model is given by the simple expression $q \cdot u \cdot e^x$. (And it is also assumed that $E = \mu e^{\gamma}$.) The first order conditions for the firms are then given by

$$0 = q_i + \left(P_i - \mu e^{x_i}\right)\frac{\partial q_i}{\partial P_i} + H_{i,i+1}\left(P_{i+1} - \mu e^{x_{i+1}}\right)\frac{\partial q_{i+1}}{\partial P_i} + H_{i,i-1}\left(P_{i-1} - \mu e^{x_{i-1}}\right)\frac{\partial q_{i-1}}{\partial p_i},$$

where $H_{i,j} = \begin{cases} 1 & \text{if } i,j \text{ are produced by the same company or there is collusion} \\ 0 & \text{otherwise} \end{cases}$.

Bresnahan then estimates δ, v_{MAX}, μ, and $\gamma$by maximum likelihood under two regimes:

Regime 1 (Collusive): $H_{i,j} = 1$ for all i, j.

Regime 2 (Nash-Competition): $H_{i,j} = 1$ if and only if i and j are produced by the same company.

Note that if $H_{i,i+1} = 0$, we would expect tougher price competition between i and $i+1$. (Especially if the difference between the qualities, $x_{i+1} - x_i$ is close to zero.)

Results

Bresnahan begins by testing the overall fit of the collusive model to the Nash-competition model in each of the three years using what amounts to a likelihood ratio test. (He also compares them both to two other models he terms the "products" model and the hedonic model, so as to make the results more robust.) He finds that the 1954 and 1956 data are best explained by the collusive model and that the 1955 data are best explained by the Nash-competition model. He then estimates the parameters δ, v_{MAX}, μ, and γ (as well as the parameters that he uses as proxies for car quality) and finds the following.

Specification	Collusive			Nash-competition		
Year	1954	1955	1956	1954	1955	1956
μ	0.1753	1.344	0.1880	0.2518	0.1747	0.2902
γ	4.593	1.604	4.441	6.352	3.911	1.204
v_{MAX}	$1.92E+7$	$1.46E+8$	$2.83E+7$	$9.81E+5$	$2.41E+7$	$1.03E+7$
δ	0.4108	$5.75E-2$	0.4075	5.04	0.4024	7.334

where

μ is a measure of marginal cost

γ measures the quality of the outside option (e.g. used cars)

v_{MAX} measures the maximum quality taste parameter of consumers

δ is a measure of the mass of consumers.

Note that these parameters should be relatively constant over time. This lends evidence to support that the collusive specification is most fitting for 1954 and 1956 and the Nash-competition specification is best for 1955.

3.3 *A Model of Add-on Pricing*, Ellison

See notes from October 1, 2007. This paper concerns all of the models we looked at including Diamond (1971) about search costs (from other lecture), Deneckere and McAfee (1996) about "damaged goods" (from other lecture), Lal and Matutes "loss leaders", and Ellison (this paper) "add-on pricing."

Vertical and horizontal differentiation between two firms (like Hotelling, but on two lines). High and low types (related to high and low marginal utilities of income or sophistication, or other characteristics allowing for differentiation in preferences). High and low qualities of good. Constant cost, c, to produce a good (of either quality). Consumer utility is given by

$$u\left(q_{1L}, q_{1H}, q_{2L}, q_{2H}; \alpha, \theta\right) = \begin{cases} v - \theta - \alpha p_{1H} & \text{if } q_{1H} = 1 \\ v - (1-\theta) - \alpha p_{2H} & \text{if } q_{2H} = 1 \\ v - w - \theta - \alpha p_{1L} & \text{if } q_{1L} = 1 \\ v - w - (1-\theta) - \alpha p_{2L} & \text{if } q_{2L} = 1 \end{cases}$$

Standard pricing model where firms choose prices, consumers observe all prices, then decide where to shop–and if they don't like the prices that they see, they can choose to go to another firm at some cost, c.

Add-on pricing model has the firm advertising low good prices, choosing high prices (which consumers don't observe), and then the consumers decide

Toward the end of the paper, Glen explains that practicing add-on pricing in the simplest incarnation of the endogenous advertising game is not individually rational (slight deviations provide higher profits). He then provides four modifications to correct this problem:

1. Per-product advertising costs (make them high), potentially incremental costs per price. But then have to argue about profitable deviations of advertising one price v. another.

2. Advertising costs are determined by consumer search patterns (too many goods for too many customers to advertise to everybody–i.e. "nationwide media campaign to tell a few potential customers that its rate for a three-day rental of a Pontiac Grand Am at the Detroit Airport on August 2, 2002 was \$74.97." Consumer search determines advertising costs.

3. Exploitation of Boundedly Rational Consumers–behavioral explanation about "tricking" unsophisticated consumers into paying more than they would if the firm advertised all of the prices of the add-ons (like rental car insurance \$97/wk for car vs. \$244/wk for insurance). Only need a moderate fraction of irrational type consumers.

4. Tacit collusion – add-on pricing must be easier than colluding directly on price to make it individually rational. Glen gives several reasons for not colluding on price (coordination in response to shocks, price preferences, monitoring deviations, etc.). Firms just make sure no one advertises price of high quality good.

One other main point of the paper is the effect of the "cheapskates" or high marginal utility consumers. They drive the result. Nobody wants the cheapskates, but the cheapskates still bring in profits. So, there is regulation between deviations to undercut to bring in profits and deviations to overcut to dump cheapskates.

One final main result is that welfare would be higher if the low quality good didn't exist–high type consumers would get the high good at a lower price and the low type consumers would get a better product, at no extra cost to the firm. The firm just extracts fewer rents, but consumer surplus is higher.

4 Dynamic Competition

4.1 *A Study of Cartel Stability*, Porter

Main Question

Did observed prices reflect switches from collusive to non-cooperative behavior? Finding: price wars are consistent with *unpredictable* shocks to demand (i.e. not the Great Lakes opening/closing).

Background and Data

First, the JEC is a cartel that controlled *eastbound* freight shipments from Chicago to the Atlantic seaboard. The Sherman act had not yet been signed, so this was a legal agreement. There was a separate westbound agreement, but the goods transported did not include grain (Porter does not mention what was shipped westbound and the agreement is not mentioned in the paper)

The Data come from price and quantity data kept by the JEC. There are several entries and exits over the span of time the paper covers. Collusion measures come from newspaper reports and a trade magazine, *Railway Review.*

Model and Estimation

A model in which firms know the market price, but not the quantity produced by the other firm. Output is homogeneous. Market demand has a stochastic component. The model predicts a patter in which price wars will occur for a certain period after low demand is observed, even when firms are in a collusive equilibrium.

Porter assumes that the probability of collusion is the same in every period (i.e. no autocorrelation). He estimates that probability using an iterative process that starts out as a sum of the collusion dummies divided by the number of time periods and uses Bayes' rule to re-estimate this probability and come up with a fitted sequence of collusion dummies. He uses both the empirical sequence and this simulated sequence in the estimation.

Demand

$\log Q(t)$ is a linear function of $\log P$ and Lakes with normal and independent errors with mean 0.

Supply

$\log P$ is a linear function of $\log Q$, structural dummies (reflecting entry and acquisition), a collusion dummy and independent (but correlated to errors in demand), non-identically distributed errors with mean 0. This is important, because we expect a switch in firm behavior to show up in the constant of the supply equation.

Costs

Small fixed costs plus aq^δ, where $\delta < 1$ is firm-specific

Conclusions

Prices during collusion are higher than Cournot but lower than implied by static joint profit maximizing, implying that imperfect monitoring is keeping prices low. The opening/closing of lakes affects supply significantly. After entry of more firms, the number of price wars increased, but the length decreased. The question of what triggered the price wars remains an open one. The collusion dummies are not systematically related to the closing/opening of lakes.

Issues

Collusion dummy is highly correlated across weeks, but this is not reflected in Porter's estimations. Note: There were no predatory reactions to entry (prices went down, but Porter maintains that this is due to the increased number of firms, not a new entrant per se). There was a "no-exit" constraint - a bankrupt railroad was forced to continue operating at lower prices but relieved of most of its fixed costs.

4.2 *Theories of Cartel Stability and the Joint Executive Committee*, Ellison

Objective and hypothesis of the paper

(1) Did the price wars in the JEC result from the strategies proposed by Green and Porter? Many strategies could lead to price wars. He focuses on transition probability between collusive and non-collusive periods. (Some evidence says "yes.")

(2) Was there countercyclical pricing, or Rotemberg-Saloner, "price wars during booms"? In this case "booms" can be seasonal changes in demand. Ellison notes that this model is illy applied to the JEC because of imperfect information and the possibility of secret price cuts. He finds little support for this.

(3) A prediction common to both models: firms don't cheat in equilibrium. Finds support for this.

Data sources and nuances

Porter's data. Notes:

(1) Only six price wars are relevant

(2) Because of secret price cuts, price is not viewed in the data. What we have in the data is the prices the firms were supposed to charge, and we use these instead of the real prices in hopes that this was the case.

(3) Assume I_t evolves according to a logit model: $P\left(I_{t+1}=1|\, I_t, Z_t\right)=\frac{e^{\gamma w_t}}{1+e^{\gamma w_t}}$, where Z_t is a set of exogenous variables, and W_t can depend on I_t (if it doesn't, we are back to Porter).

(4) Dummies DM_1 to DM_4 mark significant changes in the structure of the cartel.

(5) Variables Ellison designs: $BIGSHARE1$ - when one firm had a big share, based on deviations in $\log q$, $BIGSHARE2$ - same based on deviations in q, $BIGSHARE3$ - same based on quotas.

Model

Green and Porter

Different from Porter in two ways:

(1) Serial correlation in demand (which seems to have an effect and positive coefficients) - this has an important result which is that elasticity is greater than 1, whereas in Porter, it was less than 1. Another result is that it increases the degree of collusion from 0.4 to 0.85.

(2) Markov structure for transition probabilities (suggests more collusion).

Rotemberg and Saloner

"Booms" are defined in one of two ways:

(1) High demand is usually followed by high demand, so in a period after high demand, there is more incentive to cheat.

(2) In certain seasons, demand is higher, or at least future demand is lower (for example, right before the lakes melt).

Ellison adds a ratio of current demand to near-future demand to the supply regression (and to the transition logit).

Price cuts

Secret price cuts are like unexplained shifts upwards in demand.

Conclusions

Green and Porter

Demand shocks did cause price wars. However, small deviations do not raise the probability of a price war by much. The effect of deviating is too small to support collusion (but this might be because of unobserved things in the data).

Rotemberg and Saloner

Things seem to go in the right direction, but it is all highly insignificant

Price cuts

Amazing results, highly significant. It is shown that small price cuts happened. They show that this "unobservable" shift in demand is correlated with newspaper reports of price cuts.

4.3 *A Theory of Dynamic Oligopoly II*, Maskin, Tirole

Coming before generals.

4.4 *Price Cycles in Retail Gasoline Markets*, Noel

Objective

To find evidence of retail price cycles similar to theoretical Edgeworth Cycles (strongly asymmetric, tall, rapid, and highly synchronous prices across stations). Using a unique high-frequency data set, Noel tries to test structural and behavioral predictions consistent with the theory.

Hypothesis

Based on discussions with industry insiders, Noel believes that the price cycles are rapid in Toronto, so they should come up in a 12-hourly data series. Also, the data should allow tests to reject other explanations (e.g. demand cycles, collusion, etc.)

Data

Unique data set collected specifically for this paper. The author's assistant "hand-collected" data every 12 hours for more than 130 days in Toronto (girlfriend collected along "commute" to work). Previous work did not have data at this high frequency, which is why many cycles went "undetected." Price data collected for each type of gasoline; attributes of gas stations collected later. Data appears "representative" even though only stations along "commute" were collected.

Both national (large) firms and independent (small) stations represented.

Even in time-series graph, cycles clearly appear and they are rapid and asymmetric.

Model

Maskin-Tirole markov equilibrium which predicts theoretical Edgeworth Cycles.

Noel uses this model to test structural and behavioral predictions.

Structural Prediction: single price increase followed by consecutive sequences of small price decreases.

Behavioral Predictions: (1) firm reactions fast but not simultaneous, (2) small firms lead prices down, and (3) large firms lead prices up.

Conclusion

Data robustly confirms all predictions.

Author can test other explanations. Demand cycles seem unlikely, because of asymmetric cycles (the mean length is about one week, but the standard deviation of cycle length is high). Inventory does not seem likely either. Because of the folk theorem, collusion using this strategy is still possible, but it seems unlikely that this would be the "focal" equilibrium supporting collusion.

5 Search

5.1 *Oligopolistic Pricing with Heterogeneous Consumer Search*, Stahl

Coming before generals.

5.2 *Competition in a Market for Informed Experts' Services,* Wolinsky

At $t = 1$, firms choose (p_L, p_H, x), where x is the probability that they pretend a minor diagnosis is a major one; costs L and H to fix minor and major problems respectively.

Consumers visit experts and either buy the fix or continue searching. Costs k to get a second opinion.

(1) There is a symmetric mixed equilibrium in which: consumers always accept minor diagnosis, experts accept major with probability y and charge $(L + c, H, x)$. (We assume prior probability $w = \frac{1}{2}$.)

(1) x not equal to 0 or 1 implies indifference between lying and not lying when expert observes L. If they don't lie, they get c. If they lie, they lose out on $(1 - y)$ of the first time consumers, but otherwise get $H - L$.

(2) Indifference between accept/reject: If they accept, they pay H and their problem is fixed. Let p denote the probability that they have H given a diagnosis of H. If they don't accept, then with probability $p + (1 - p)\, x$, they pay $H + k$ (extra search cost) and with probability $(1 - p)\,(1 - x)$, they pay $L + c + e$.

(3) Check price cuts.

For sufficiently low search costs, there is an equilibrium in which:

(1) Some firms specialize in L, choosing $(L, \infty, 0)$

(2) Others do H, choosing $(H, H, 1)$

(3) Consumers go to an L firm first, then to an H firm.

Both equilibria are inefficient, because (with some probability) they involve extra search costs on the part of the consumers.

Other stuff in Wolinsky:

(1) With a continuum, there is vertical specialization and customers proceed sequentially up the category ladder.

(2) Reputation: Expert fixes problem and then presents a bill. Experts do not know whether customers will be repeat customers. Added feature - experts can reject consumers if they don't want to serve them once they observe their type.

Equilibria:

(a) Experts charge the same price for everything, serve everyone truthfully, and consumers return to the same expert unless he rejected them - reputation offsets the temptation to reject consumers with the H problem.

(b) Charge different prices, always truthful, customers return only if expert charged L on the first visit.

Diagnostic errors: Lower search costs $\Rightarrow$ lower consumer surplus because of negative search externality of consumers needing the H treatment. H customers continue searching until they meet an expert who errs and offers the low price. Low price experts treat more H customers and take a loss, forcing them to raise their equilibrium low price.

5.3 *A Theory of Fraud and Overtreatment in Experts Markets*, Alger, Salanie

Coming before generals.

5.4 *Equilibrium Price Dispersion in Retail Markets for Prescription Drugs*, Sorensen

Objective

In this paper, Sorensen tries to justify the price dispersion in the pharmaceutical industry as being a result of rational consumer search. In particular, he claims that for drugs which are purchased frequently, there is a higher return to searching on the basis of price for the consumers, and that we would therefore tend to see less price dispersion. He also controls for the idea that the price dispersion results are being driven by pharmacy heterogeneity, and he asserts that the wholesale cost difference for pharmacies is also not important.

Data Source and Issues

State law in New York requires the posting of the prices of the 152 top-selling prescriptions on a prominent sign in each pharmacy. Alan Sorensen went to twenty one pharmacies in two cities in upstate New York (Middletown and Newburgh) and copied down the prices off these signs. He had problems collecting data from two of these pharmacies, as one pharmacist had not updated it in months, and another pharmacist forced him to leave before he could write down all the prices. In addition, Sorensen used data on drug characteristics derived from the pharmaceutical reference book *Mosby's GenRx.* Sorensen's preliminary findings suggest that Rite-Aid has, in general, the highest prices, and Walmart and Price Chopper have the lowest, but nevertheless, there is still sufficient dispersion in the prices so as to warrant consumer search.

The Model

It is unclear what search model is underlying Sorensen's primary hypothesis. The Stahl model we covered in class describes price dispersion as an equilibrium outcome of an economy in which there is a point mass of consumers who have zero shopping costs. (In contrast to the Diamond model in which, if no one enjoys shopping, then all prices charged are exactly equal to the monopoly prices.) Here, we are led to believe the (plausible) claim that heterogeneous benefits of search can lead to differences in price dispersion among commodities, but the mechanism through which such effects occur is not made explicit.

Nevertheless, the empirical model Sorensen uses is as follows

$$RANGE_{ij} = \beta_0 + \beta_1 PFREQ_i + \beta_2 AWP_i + \beta_3 BR1_i + \beta_4 BR2_i + \beta_5 NEWB_i + \sum_{k=6}^{25} \beta_k D_{ik} + \varepsilon_{ij},$$

where

$RANGE_{ij}$ is the difference between the highest and lowest prices for drug i in city j

$PFREQ_i$ is the frequency with which drug i must be purchased

AWP_i is the average wholesale price for drug i

$BR1_i$ is equal to 1 if drug i faces competition from generic drugs

$BR2_i$ is equal to 1 if drug i does not face competition from generic drugs

$NEWB_i$ is equal to 1 for all observations in Newburgh

D_{ik} are indicators for 20 categories of drug therapy

Note that this is actually a system of two equations and, to make use of the fact that $E[\varepsilon_{i1}\varepsilon_{i2}] \neq 0$, he estimates this model using generalized least squares in the seemingly unrelated regression framework. Sorensen also estimates a model in which he "partials out" the fixed effects of the pharmacies.

Results

The results of the primary specification are given in the first column, and the results of the specification in which he partials out the pharmacy fixed effects and measures the range of the residuals are given in the second column.

	Range	Residual Range
$PFREQ$	-0.336^{**}	-0.266^{**}
AWP	0.280^{**}	0.215^{**}
$BR1$	-0.803	-1.842^{**}
$BR2$	-1.505	-1.967^{*}
$NEWB$	-2.686^{**}	-1.493^{*}
$CONST$	20.070^{**}	14.570^{**}
R^2	0.371	0.258
$\hat{\rho}$	0.338	0.149

As Sorensen's hypothesis predicts, the coefficient on $PFREQ$, which is a proxy for the benefits to consumer search, is negative. That is, for drugs which consumers purchase more often, there is less price dispersion. The importance of the second specification is that, even after controlling for heterogeneity across pharmacies, the coefficient on $PFREQ$ is still negative.

In addition, Sorensen computes the average markups for drugs and regresses them on the same right-hand side variables as in the previous model, finding that the coefficient for $PFREQ$ is -0.262^{**}.

Finally, Sorensen does a more detailed test as to whether or not pharmacy heterogeneity is responsible for the observed price dispersion. He concludes that pharmacy heterogeneity can account for approximately 33 percent of the dispersion, but he acknowledges that this estimate might be high.

5.5 *Search, Obfuscation, and Price Elasticities on the Internet*, Ellison, Ellison

Objective

To describe the market effects of "obfuscation" on the internet. Using proprietary quantity data matched to an hourly search-engine price series.

Hypothesis

To recover fixed costs in the face of search engines, retailers will practice "add-on" pricing and obfuscate results.

Data

Hourly data from "pricewatch" and quantity data from Ellison relative (who sells chips and memory). The hourly data allows identification of price changes (which the authors take to be exogenous for most of the paper, although they have some IV specifications which attempt to instrument for price). Quantity data allows econometrician to recover elasticities given own price changes and other firms' price changes.

Model

Glenn's "add-on pricing" model and also a simple model of search engines. No formal model is strictly taken to the data, but the models conceptualize the empirical results.

Conclusions

Price elasticity extremely high for "low-quality" products, but much more inelastic for medium and high quality products. The results are hard to pin down, however. There is some evidence of "loss leading" and some evidence of "add-on pricing."

Most unexpected result is that within-firm elasticities are different from between-firm elasticities across qualities (i.e. raising low-quality price seems to induce consumers to switch to your own firm's medium-quality price). There is also clear evidence of search obfuscation. Given that one would expect these small retailers to be a very "competitive" industry (i.e. there is likely no collusion among many many small independent CPU and memory retailers), it appears that the "obfuscation," which encourages consumers to purchase more expensive CPUs and memory units, is an important part of their firm's pricing strategy. This might be crucial in allowing them to recover fixed costs.

6 Entry

6.1 *Free Entry and Social Inefficiency in Radio Broadcasting*, Berry, Waldfogel

Objective

The "quantify the inefficiency" of free entry in the radio broadcasting industry. In other words, the authors assume (by using a particular model) that there is inefficiency in the market and attempt to estimate its size. Thus, they do not really have a hypothesis - rather, they estimate the size of inefficiency using a set of assumptions. They also consider under what set of assumptions there *wouldn't* be inefficiency in the market (i.e. take the number of firms as optimal and compute the estimated consumer surplus from listening to the radio.)

They also compare the efficient number of radio stations to the number of stations under monopoly control and find that (surprise) the monopolist underprovides the number of radio stations.

Data

Cross sectional (Metropolitan Statistical Areas - MSA's)

Stations broadcasting from *within* the area

Stations broadcasting from *outside* the area (they later use this as an instrument for the number of stations broadcasting from within the area. Population is also used to instrument for this).

Number of listeners for each station (Arbitron data gathered from sample population paid to fill out radio listening surveys) - "average quarter hour rating" - the number of people listening to a station for at least 5 out of each 15 minutes.

Aggregate annual revenue for *some* radio stations in MSA (sum only - the authors calculate advertising revenue based on total revenue and which radio stations reported)

Commercial radio stations only

Population, other demographics included

Radio station characteristics (FM, mega-watt, tower height, music category, etc.)

Model

Standard entry model with fixed costs (which are distributed log-normally). Excessive entry is possible when average costs are falling and entrant's products are substitutes. Consider social welfare to be the sum of radio and advertising profits, without taking listeners' surplus into account.

Empirics

New programming (more stations) does not greatly increase total listening.

Advertising price per listener declining in total listening share of a station.

Positive relationship between population and fixed costs.

Distribution of costs is estimated from entry decisions. Costs are jointly estimated, but the authors state that equation-by-equation estimation does not affect the results much. However, the joint estimation produces the "right" standard errors.

Error terms on utility from listening to a station in a market are i.i.d. extreme value(depend on both individual and stations).

Assumptions

Identical distribution of fixed costs for all radio stations. Equal shares of listeners for each station in the market. Advertisers have no preferences for a specific radio station.

Extensions

Allow for station heterogeneity (but this still does not reduce the estimate of business stealing, which drives the result of the paper). It is not possible to model entry with station heterogeneity (they say it is "beyond the current empirical literature.")

Endogenous fixed costs. This is something I can see Glenn asking about - it is not too complicated, but he did not talk about it in class. The argument is that stations are choosing the quality of their programming (and thus their costs) and providing it to listeners at a given marginal cost so there is no inefficiency. But the retort is that increases in product quality still produces business-stealing, because it enters the demand for other products and has a negative effect on a rival's profit that the entering firm doesn't consider.

Government restrictions on broadcasting (doesn't change estimates)

Reducing the number of outside stations as well as inside stations in optimality simulation causes the estimation of deadweight loss from free entry to increase even more.

Endogeneity of outside-the-metro stations - include "isolated" markets only - results unchanged.

Conclusions

When considering only advertising and radio station surplus, there is substantial excess entry. The surplus to listeners would need to be about 15 cents/hour to balance out the business stealing effect. The authors consider this number to be large. The business stealing result is robust under all the specifications the authors try. R&D-intensive industries, print media, computer software, and TV broadcasting are expected to share the characteristics of high fixed costs and nearly zero marginal costs and thus be susceptible to excess entry.

6.2 *Patterns of Firm Entry and Exit in U.S. Manufacturing,* Dunne, Roberts, Samuelson

Objective and Hypothesis of the Paper

This paper wants to test entry and exit patterns of firms.

(1) Identify 3 types of entrants and "assess their importance." The three types of entrants are: new firms, old firms opening new plants, old firms embarking on new production in old plants.

(2) Persistence of entry and exit patterns and correlation between industry entry and exit rates.

(3) Assess the post-entry performance: size, market share, and failure rate.

Data Sources and Nuances

Plant level data from US census of manufactures (63-82: which is five, because it comes out every five years). What is so great about this data set:

(1) Because it is important to make observations that are industry-specific, they treat the 4-digit Standard Industrial Classification as separate markets.

(2) Panel data - allows aggregation to the firm level.

(3) Data on the value of production that can be used to compute market shares

(4) Allows identification of the "three types" of entrants.

(5) Mergers and changes in ownership won't be mistaken for entry and exit.

There was a major change in CIS in 1972, so they had to reclassify 7-digit industries (1963,1967) into 4-digit industries. Also, sometimes there were changes in a plant's identification number for legal reasons which leads to measurement error. To reduce measurement error, they delete the bottom 1% of output. Single plant firms account for 93 percent of the firms, but only 17 percent of production.

Weaknesses of this data set:

(1) Five year intervals - underestimates year-to-year entry and exit.

(2) Interval 1963-1967 is 4 years, while the other intervals are 5 years (they don't correct for this)

(3) Census changed its treatment of small firms, which can result in a slight increase in estimated exit rates.

The Model Used

They examine the following:

(1) Entry and exit rates

(2) Entering and exiting market shares

(3) Average size of exiting firms relative to non-exiting firms, and the average size of entering firms relative to existing firms.

Conclusions

General entry and exit results

(1) About 40% of firms are entrants (compared to previous census)

(2) Market share of entrants is smaller (16%)

(3) Entrants produce about 35% of average production

(4) Numbers on exiting firms are similar

New firms versus diversifying firms

(1) New - 55%, diversifying new plant - 8.5%, diversifying old plant - -the rest

(2) Diversifying accounts for more production than their share of entry (50% versus 44%)

Industry specific results

(1) Entering and exiting results are positively correlated: both in industry and in time

(2) Persistent differences in entry and exit rates across time (instruments, lumber and printing have high entry and exit rates)

Longitudinal aspects of entry and exit

(1) The market share of a group of entrants ("cohort") declines across time. This is a result of two conflicting forces: firms exit and surviving firms grow larger.

(2) Diversifying new-plant cohorts' market shares decline less.

6.3 *Pioneer Advantage,* Golder, Tellis

Coming before generals.

Part II
After Midterm

7 Networks and Two-Sided Markets

7.1 *Network Effects and Microsoft,* Bresnahan

- Motivation: Network effects important in modern high tech industries
  - Theory tells us:
    - Positive feedback in decision rules of individuals
    - Indeterminacy of equilibrium
    - First-mover advantages or barriers to entry
    - High inertia for established standards, high volatility for nascent ones
    - Strategic competition intense while establishing standards, little competition after lock-in to a particular standard
  - These theories are difficult to test
    - Hard to distinguish behavior of different actors econometrically, because of positive feedback and coordination.
    - Theory implies indeterminacy, but scholars cannot observe what did not happen
  - Solution: look at business documents
- Data: Microsoft business documents
  - These are documents from a Microsoft antitrust case, covering 1994-1998
  - Documents discuss strategy and marketing of browsers, operating systems, and applications - dividing technology Java
  - Three major themes of these documents:
    - "Browser war" between Internet Explorer and Netscape Navigator, 1995-1998
    - Competitive position of Microsoft's PC operating system (the de facto standard): did Microsoft have a "permanent" monopoly or was it susceptible to entry?
    - The relationship between their operating system (popular among users) and Java (popular among developers)
- Results: Theory's implications in the documents
  - Network effects and positive feedback cycles
    - Microsoft argues that to win a platform battle, they need to have a large installed base of users and drive adoption of the platform among developers of complementary applications: need both sides of the 2-sided market!
  - Indeterminacy
    - Markets can "tip" to any of several standards
    - This is an implicit assumption in the browser war - market can tip to IE or Navigator
  - Leader Advantages/First Mover Advantages/Tipping
    - Microsoft doesn't worry too much about competing with the "network computer," because network effects make a barrier to entry
  - Browsers and tipping

        * Netscape had the leading position at first in the browser war
            · For users: Microsoft's strategy was to make sure all the good content was available for IE only to incentivize shifting
            · For developers: Microsoft worries that developers are more and more tied to Netscape since Netscape has a higher share of browser usage

- Results: Underpinnings and Refinements of Theory
    - Sunk costs
        * The theory says that it is easier to influence the network effects and direction of the positive feedback loop early because of standard-specific sunk costs of nonstrategic actors (users and developers)
        * In practice, there seems to be some cases in which people act outside of their own self-interest to coordinate on strategies, e.g. developers try to thwart Microsoft from learning Netscape's revenue sources, even though their best interest may be to encourage openness
        * This departure from rationality doesn't really threaten theoretical predictions
    - Expectations
        * Expectations matter in determining the dynamic results when the static long-run equilibrium is indeterminate
        * Firms try to influence expectations
    - Entry via complements
        * Theory doesn't cover this, but Microsoft believes that lock-in may end when a complement enters and disrupts the market equilibrium
    - Vertical integration
        * Theory doesn't describe this, but Microsoft sees vertical integration as a way to remove divided technical leadership
    - Compatibility decisions
        * Consumers and developers like choice, but Microsoft wants to lock people in. For example, it doesn't want a cross platform Java environment to encourage developers to write programs that will run on Windows or Mac OS.
        * Solution: Make the Microsoft product work on different platforms, so there is no advantage to using the other product. e.g. IE runs on Macs

- Bresnahan's conclusion:
    - There is congruence between theory and practice
    - Practice is richer than theory, and the inductive logic used by businessmen has some value in thinking about theory

- Glenn's take:
    - This may be an interesting source of new theory ideas, but he's a little skeptical

7.2 *Two Sided Markets: A Progress Report,* Rochet, Tirole

What is a Two Sided Market?

This paper is very much related to the first lecture that Tirole gave in November, yet it covers a wider range of topics. A cursory summary will be given here, because (1) the details of the various extensions are probably beyond the scope of this class and (2) [the summarizer] understood very little beyond what was covered in class.

Examples of two-sided markets: video game platforms (gamers and developers), credit cards (consumers and merchants), operating systems (end-users and developers), newspapers (eyeballs and advertisers). Why isn't a standard market (buyers and sellers) considered to be a two-sided market? This paper provides two equivalent definitions of two-sided markets. Loosely speaking, a market is considered to be two-sided if the allocation of per-usage charges affects the volume of trade.

To see why the standard market setting with one buyer and one seller is not two-sided, note that the overall level of per-usage charges is usually zero. However, if a tax t is imposed, then the overall level of per-usage charges becomes t. It is well known result of tax incidence analysis that it does not matter who is required to pay the tax. The level of trade will be the same regardless of who bears the official tax burden.

Typically, it is the inability of one side of the market to transact with the other side of the market that leads to two-sidedness. For example, if merchants could charge different prices to consumers who pay with credit cards and to consumers who pay with cash, then the credit card market would no longer be two-sided. For two-sidedness to occur, it must be the case that one side cannot "pass the cost" onto the other side. This is why it is often said that the failure of the Coase theorem is a necessary, but not sufficient, condition for markets to be two-sided. The reason why the failure of the Coase theorem is not sufficient for two-sidedness is that asymmetric information (which is another failure of the Coase theorem) need not cause a market to be two-sided.

The Basic Model

Let B and S be two sides of a platform. Let $i \in \{B, S\}$ and define the following variables

b^i - usage benefit to an individual of side i

a^i - usage charge assessed to side i by the platform

B^i - membership benefit to an individual of side i

A^i - membership charge assessed to side i by the platform

N^i - number of members of side i on the platform

C^i - fixed cost of membership

c - usage cost per transaction

Utility of an individual of type i is given by

$$U^i = \left(b^i - a^i\right) N^j + B^i - A^i,$$

where $N^i = \Pr\left[U^i \geq 0\right] \equiv D^i\left(p^i, N^j\right)$. It can be shown that, under some regularity conditions,

$$\begin{aligned} N^B &= n^B\left(p^B, p^S\right) \\ N^S &= n^S\left(p^B, p^S\right), \end{aligned}$$

where $p^i = a^i + \frac{A^i - C^i}{N^j}$. Next, we have that for a given level p, the optimal price structure is given by

$$V(p) = \max\left\{ n^B\left(p^B, p^S\right) n^S\left(p^B, p^S\right) \middle| \, p^B + p^S = p \right\}.$$

If there exists a finite number of solutions to this problem, then the market in which the platform is operating is considered to be two-sided. Otherwise, if this problem is solved for any $\left(p^B, p^S\right)$ such that $p^B + p^S = p$, then the market is not two-sided. (i.e. the allocation of prices does not affect the volume of trade.) Note that, since p^i is a function of both a^i and A^i, this says that if either the allocation of (1) the usage fees or (2) the membership fees matters, then the market is two-sided.

Extensions

Multi-Homing

If it turns out that, say, buyers join two platforms, whereas sellers join only one of two platforms, then the buyers are said to be multi-homing. Since each platform receives revenues only if the buyers and sellers transact on that given platform, it will be the case that the platform will cater to the buyers by offering them low membership fees (possibly negative) and low usage fees (also possibly negative) and charge the sellers high membership fees and high usage fees. Examples here include Acrobat (reader is free, whereas the writer is very expensive - unless you pirate it), video game consoles (the consoles are usually priced below cost, whereas the development kit is very expensive)

Payment between end-users

The canonical preferences $U^i = \left(b^i - a^i\right) N^j + B^i - A^i$ can be altered to be consistent with payments between end-users. If the trade between end-users is the outcome of bargaining (with take-it-or-leave-it offers being an extreme form of bargaining), then the platform's optimal decision is to charge a transaction charge a to maximize

$$v\left(a\right) = E\left[\left(b^B + b^S - c\right) x\left(b, a\right)\right],$$

where $x\left(b, a\right)$ is the probability of trade resulting from the bargaining process, and to choose its membership payments, as in the canonical model, so as to

$$\max\left[p^B + p^S + v\left(a^*\right)\right] n^B n^S.$$

8 Strategic Investment

8.1 *Capital Structure and Product Market Competition,* Chevalier

Objective and Hypothesis

- The objective of this paper is to see the effects of leveraged buyouts (LBOs) on rivals' decisions in the supermarket industry. Particularly, whether they "soften" or "toughen" competition.
- Given that they were triggered by the threat of takeovers, Chevalier takes them as exogenous (she finds this assumption consistent with her results later).
- Analyzes impact on expected future profits (first part), expansion and entry/deterrence behavior (second part).
- First part: analysis in the stock-return response to LBO announcement
  - If investment is soft, $\frac{dE\pi_2}{dK} > 0$. Rivals' stock valuation should have positive abnormal returns when an LBO is announced.
- Second part: analysis of the entrance and expansion behavior in the presence of LBOs
  - If investment is soft, less competition should promote more entry and expansion.

Data Source

- First part: 1985-1988. Stock-market return response for 13 rival supermarkets to LBOs of Safeway. Compares it to general stock-market trends to see if it is significantly different. Uses exit behavior of Safeway to separate competing and non-competing markets (according to Safeway CEO, they sold out those stores that were less profitable - i.e. those that were subject to stronger competition).
- Second part: 1985-1991. Supermarket composition in 85 cities. Dependent variable: change in number of stores. Explanatory variables: number of households, Herfindahl index, deviation from the expected number of supermarkets in 1985, share of supermarkets with high LBO, and some other controls. Wants to see if the coefficient of LBO supermarkets is significantly positive after correcting for some other variables that could explain a change in the number of firms.

Model Used

- This is a paper on strategic investment and is based on Tirole's benchmark "soft/tough". However, it does not address the model deeply, but only mentions that one expects LBOs to "soften" market competition.
- Econometric models vary on the tests: SUR (first part), OLS, ordered probit (second part).

Conclusions of the Paper

- Results
  - First part: coefficient seems to be positive in competing markets, although some caveats might invalidate it (as we saw in problem set 9).
  - Second part: coefficient is significant for rival incumbents, but not for supermarket stores in general. Argue that the effect of LBO on entrance behavior might take longer than expansion reaction. Compares early and late LBOs to control for it and finds consistent evidence. Early LBOs have stronger effects for the analyzed period.
- Conclusion: LBOs seem to "soften" product market competition. Thus, it is important to take into account product market effects of capital market decisions when choosing optimal capital structure.

8.2 *Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration*, Ellison, Ellison

Summary

This paper is an example of one out of three ways to test for strategic investment - find a qualitative prediction of the model and test it. The qualitative prediction is that investment levels will be non-monotonic in market size if there is strategic investment. The authors study the behavior of pharmaceutical companies prior to patent expiration - when the patents expire, generic entry is possible.

Basic prediction: investing to deter entry is worth it in medium sized markets. In small markets, entry is more likely to be unprofitable, so deterrence is more likely to be superfluous. In large markets, entry is more likely to be profitable no matter what the incumbent tries to do, so deterrence is more likely to be ineffective.

The authors conclude that there is evidence of non-monotonicity in detail advertising - less detail advertising in medium sized markets, since detail advertising has a positive externality for entrants. They also find evidence of "presentation proliferation" in medium-sized markets.

Model

The one we covered in class - incumbent makes an investment decision, which is observed by the potential entrant before the entry decision is made. The entrant then makes its decision based on profitability prospects in the market.

Direct effect: the relationship between a market characteristic z (here, size) and the optimal level of investment, A, in the absence of competition.

Competition effect: the difference between the effect of investment on monopoly profits and its effect on duopoly profits, $\frac{\partial \pi_1^{d^*}}{\partial A} - \frac{\partial \pi_1^{m^*}}{\partial A}$. If this is positive, for example, firm 1 will invest more when the market is larger.

The assumption the authors need in order to have monotonicity be the prediction in the absence of entry deterrence motives: the competition effect and the direct effect have the same sign.

Data

63 drugs that lost patent protection between 1986 and 1992.

Data source: Historical IMS audits. What is IMS? (Good question.)

Advertising - detail (sending representatives to doctors, most advertising spending by drug companies is in this category) and journal (print ads). The problem with advertising is that the direct effects and competition effects go in the opposite directions, which makes the authors' methodology potentially not applicable.

Product proliferation (number of forms in which the drug is available). Another case where effects go in opposite directions.

Prices - paid by retail/hospital sector (wholesale)

Revenues - annual presentation-level (i.e. different for the pill and liquid version of the same drug). Measures market size. The authors find that revenue is a strong predictor of generic entry.

Data for each drug spans three years prior to patent expiration and the year after.

To test for monotonicity, the authors use an approach by Hall and Heckman. They modify this approach and suggest their own test statistic. They use simulations rather than proofs to show the validity of their test statistic. They address potential problems of measurement error and endogeneity and suggest that it is not a problem.

Control variables: shares of sales to hospitals; generic entry dummy (measured by whether any other firm had an approval to produce the drug generically); chronic, psychological, topical drug dummies; whether the drug tends to be prescribed by a specialist.

Results

Detail advertising: evidence of entry-deterrence - detail advertising reduced in medium-sized markets prior to patent expiration.

Journal advertising: looks non-monotone, but significance depends on the test used.

Presentation proliferation: departure from monotonicity at about 11% significance level.
Caveat (for this and other strategic investment studies): we cannot see the long-run effects of investment.

8.3 *Entry, its Deterrence, and its Accommodation*, Kadiyali

Objective

The objective of this paper is to answer the following questions.
(1) Why does entry take place or not take place?
(2) Why and how is entry accommodated?
(3) Are pre- and post-entry market structure and strategies consistent with theoretical predictions?

Context

Kodak was a monopolist in the 1970s. Fuji had a very minor presence. In 1979, Fuji announced its intent to grow its market share from a previously very minor presence. Beginning in 1980, Fuji and Kodak compete.

Data

Quarterly firm-level time-series on quantities, prices, advertising dollars, costs of factors and materials. The data are from many sources. Reported in an appendix (so probably won't be tested). Here are a few of the sources: *Popular Photography* magazine, *Wolfram Report*, *Kodak v. Berkey Phone* case, *Monthly Labor Review.*

Structural Modeling Approach

Demand

$$q_i = a_{i1}P_1 + a_{i2}P_2 + \gamma_{i1}\sqrt{A_1} + \gamma_{i2}\sqrt{A_2} + \alpha_{i0} + \sum_{j=1}^{3}\alpha_j Q_j + \alpha_{i4}Y$$

Quantity as a function of P price, A advertising, Q_j quarterly dummies, and Y, real income. These are estimated separately for Kodak in the pre- and post-periods, and for Fuji only in the post-period. By assumption, there are diminishing marginal returns to advertising.

Cost

$$MC_i = ni1 \cdot \text{Price\_capital1} + ni2 \cdot \text{Price\_labor1} + ni3 \cdot \text{Price\_silver} + ni4 \cdot time,$$

and they assume constant marginal cost.

Collinearity problem $\Rightarrow$ use estimates from pre-entry period to reduce parameters estimated in the post-entry period.

Optimization

Assume that firms choose price and advertising to maximize profit

$$(p - c) \cdot D(p, A)$$

They consider many forms of interaction: 9 forms of noncooperative interaction (simultaneous and sequential move games), and 3 forms of collusion (in pricing, advertising, and both). Given all these estimates, they find the best fitting game.

Results:

- Pre-entry period: evidence of limit pricing and entry deterrence
  - Price elasticity of demand is only -0.64. That is, Kodak is pricing below the monopoly level.
  - Seems that there were low price and advertising elasticities because of low prices and big advertising budget.
  - However, there is no counterfactual, so we don't know if this was a pooling equilibrium where all incumbents would use limit pricing, or a separating equilibrium where only low-cost incumbents would use limit pricing.

- Post-entry period: evidence of collusion
  - Run a race between the 12 games proposed in the previous section
    - Best fitting game: Kodak and Fuji collude on price and advertising
  - Low own-price and own-advertising elasticities: are Kodak and Fuji colluding to actively deter further entry?
    - Advertising by one firm increases *both* firms' demand
      - Collusion ⇒ more advertising
    - Colluding on ads and prices gives higher profits than just colluding on ads
      - Collusion keeps prices low enough to deter entry, but higher than a more competitive equilibrium.
- Maybe they collude on price and ads, but not on entry
  - Fuji and Kodak also compete in foreign markets.
  - Find Kodak's profits would have been lower if it did not accommodate entry
  - Both firms chose entry and accommodation optimally given the estimated parameters.
- Tirole's animal framework: Kodak appeases, Fuji aggressive
  - Kodak does not cut its prices and lets Fuji have the lower price: appeasing stance - must be puppy dog or fat cat
  - Kodak's advertising budget increased, benefiting Fuji proportionally more than it benefited Kodak - Kodak acts soft, the action is a strategic complement - Kodak is a fat cat
  - Fuji was aggressive: lower advertising and lower price - Fuji is the lean and hungry guy.

Glenn's comments on this paper:

- Maybe demand is not estimated correctly: (1) bad instruments can cause bias, (2) need dynamic model.
- Very strong functional form assumptions allow her to do this without very much data.

9 Predation and Limit Pricing

9.1 *A 'Signal-Jamming' Theory of Predation*, Fudenberg, Tirole

Objective

This paper suggests a model of rational predatory behavior

Model

This is a two period model. Firm one is the "older" firm in the sense that its (fixed) cost is known to it. The cost of firm two (the "entrant") is unknown to either firm. In both periods, firms engage in some sort of competition. The funny thing about this competition is that there is no market clearing price (firms can charge different prices and still get positive sales). The firms do not observe each others' prices.

The signal jamming story is that firm one wants to lower its price in order to interfere with firm two's inference problem about its own cost. Firm one wants to cause firm two to have lower profits in order to "trick" firm two into thinking its cost is high, which will cause it to exit the market. In a separating Bayesian equilibrium (which exists under all sorts of concavity assumptions), firm two learns its cost and exits only when it is sequentially optimal. However, predatory behavior still occurs, because otherwise firm one would have an incentive to cut prices to "fool" firm two (contract people: this is just like Holmstrom's "career concerns" story, which they even elaborate on in this paper).

Note that predatory behavior causes expected profits of entrants to be lower and hence deters entry. Welfare implications (compared to complete information on costs) are ambiguous, because:

(1) Prices are lower in the first stage (could be welfare improving or not)

(2) Entry deterrence (also can be welfare improving or not)

Differences between signal-jamming and reputation models (as in Milgrom and Roberts)

(1) Entrant needs to enter for predation to occur in signal-jamming

(2) Typically finite set of equilibria as opposed to continuum of equilibria in reputation models.

Another theory of predation is the "long purse story," which is that firms are financially constrained and low profits in the first stage would make it impossible for them to finance themselves in the second stage. Though this theory is widely used, it is theoretically unsound, because with perfect information, there is no reason why the firm would not be able to take a loan in the second stage. The authors integrate this story with their model: the bank is unaware of the manager's ability and tries to infer it from the firm's first-stage profit. Differences with the pure long-purse story:

(1) There is no need for uncertainty in long purse story: the effect is on wealth, not on knowledge.

(2) Predatory behavior in the long-purse story will not occur unless successful.

9.2 *Predation and Its Internal Rate of Return*, Genesove, Mullin

Price wars following the entries of Spreckels, Sr. and Arbuckle were predatory:

- Comparisons to marginal cost, and constructing the predicted competitive price cost margins.
- Entry into the American sugar refining industry following the formation of the Sugar Trust
- Organized in 1887. Reorganized as the American Sugar Refining Company (ASRC) in 1891, the ninth largest corporation in the US by 1909.
- In December 1887, 18 firms controlled 80% of the capacity
- After reorganization, refined prices rose 16%.
- Claus Spreckels, Sr. completed a plant in December 1889: two year price war
- ASRC acquired the plant by April 1892, capacity share 95%
- A series of small scale firms entered (five firms, each 1340 barrels/day, while ASRC 49,500 barrels)
- In August 1898, Arbuckle Brothers finished a refining plant: a large wholesale grocer, dominant roaster of coffee, patent on a packaging machine. They applied the packaging to sugar from 1892-1896. Another entrant, Doscher refinery began production in November 1898. (each plant had a capacity of 3000 barrels). ASRC dropped to 87%.
- Price war until May 1900. Doscher merged with two of the major independents to form the NSRC, under the control of ASRC.
- The federal government filed an antitrust suit against ASRC in 1910 (ASRC alone 66%, ASRC+NSRC 80%)

The history of the sugar industry: Eichner (1969) and Zerbe (1969)
Weekly statistical sugar trade journal (Willett and Gray)
Testimony of the ASRC president before the Industrial Commission, 1899
The Hardwick Committee hearings of 1911
Audits conducted by the US Tariff Commission between 1914 and 1919

- Common technology, marginal cost up to plant capacity
- $c = c_0 + 1.075 p_{raw}$ (marginal cost to produce hundred pounds of refined sugar)
- 1.075 fixed coefficient production technology between raw sugar and refined sugar.
- Estimates for c_0 between 16 cents and 26 cents (per hundred pounds, in 1898 dollars)
- Proper margin $= P - 1.075 p_{raw}$
- Arbuckle: proper margin sometimes fell below the estimated range for c_0, indicating pricing below marginal cost.
- Spreckels: proper margin below marginal cost for a number of weeks and well above marginal cost only during the high season (summer).
- On average, the proper margin was 53 cents in high season, 46 cents in low season.

Possible explanations for the price war

- To drive the added capacity out of the industry: no - it was kept after the war

- A lower purchase price for the plant: no - the plausible range of an entrant's beliefs about ASRC's costs too small
- A greater market share in a collusive equilibrium: no - during the price war with Arbuckle, an upper bound for an internal rate of return only 0.67percent.
- Deterrence of entry: gross loss of 21 million dollars from the price wars. Given the internal rate of return, predation was profitable if one or two 3000 barrels per day plants were to enter, but not profitable for a plant with a capacity of 1840 bpd.
- Long purse hypothesis: no evidence, the internal rate of return was 5.8%.

10 Auctions

10.1 *Auction Theory: A Guide to the Literature,* Klemperer

- Four basic types of auctions
  - Ascending bid (open, oral, or English) - price raised until only one bidder remains
  - Descending bid (Dutch) - price lowered until one bidder calls out
  - First-price sealed bid auction
    - Strategically equivalent to descending bid
  - Second-price sealed bid auction
    - Strategically equivalent to the ascending bid auction (true for private values or with two bidders; otherwise, drop out price for lower valuation guys tells you something about your own valuation.)
- Revenue equivalence theorem - risk-neutral buyers, privately-known signal from a common, strictly atomless distribution; any auction mechanism in which (1) the object always goes to the buyer with the highest signal, and (2) any bidder with the lowest feasible signal expects zero surplus, yields the same expected revenue (and results in each bidder making the same payment as a function of her signal)
  - Note: applies to both private value models, and common value models, provided signals are independent
- Marginal revenues: imagine a firm whose demand curve is constructed from bidders whose values are independently drawn from the same distribution (i.e. demand at price $p = v$ is $1 - F(v)$). A bidder's marginal revenue is defined as the marginal revenue of this firm at the price that equals the bidder's actual value $\left(v - \frac{1-F(v)}{v}\right)$.
  - The expected revenue from an auction is the expected marginal revenue of the winning bidder (e.g. in an ascending auction, the winning bidder pays the second highest valuation: v^*)
- Risk aversion
  - Bid more in first-price auction (higher probability of winning, lower value of winning - better with risk aversion). Thus, first-price auction produces more revenue.
  - Risk-averse auctioneer prefers first price auction: conditional on knowing the highest valuation, the price is fixed in the first-price auction, random with the same mean in the second-price auction; risk averse $\Rightarrow$ prefer certainty to randomness.
- Affiliated values
  - In terms of revenue: ascending auction is better than second-price sealed bid auction which is greater than first-price auction.
    - The more the price depends on other bidders' information, the lower is the informational rent to the winner (the price is closer to their valuation)
    - In general, the principle that revenue is raised by linking the winner's payment to information affiliated with that payment is known as the Linkage Principle
    - With affiliated private values, lower reserve price (intuition: more certainty about any one bidder's valuation conditional on other bidder's information.)
- Asymmetries

  - First-price auction discriminates in favor of selling to bidders from the weaker distribution, since they bid more aggressively.
  - First-price auction is allocatively inefficient, but may be more profitable in expectation.
  - With common values, one player having slightly more information can translate into winning a large percentage of the time.

- Entry is optimal provided the reserve price is the seller's valuation.
- Collusion - easier to sustain in second price auction - all bid 0 except the winner.
- Multiunit auctions - revenue equivalence sometimes holds (homogeneous units, each buyer interested in just one unit), but efficiency very hard to achieve.
- Royalties: if the winner's value can be observed ex post, the seller can do better by making the winner's payment depend on that value.
- Double auctions - multiple buyers and sellers - typically efficient as the number of buyers and sellers becomes large.

10.2 *Internet Advertising and the Generalized Second-Price Auction*, Edelman, Ostrovsky, Schwarz

Executive Summary

Generalized Second Price (GSP) auctions have become the mechanism of choice for search engines to sell advertising. Some key features of the GSP are:

- It is not the Vickrey-Clarke-Groves mechanism (which I guess is a famous mechanism for multi-unit auctions)
  - Has no equilibrium in dominant strategies
  - Truth-telling is not an equilibrium
- An equilibrium of GSP has same payoffs as VCG
- The English Auction corresponding to GSP has a unique "continuous strategy" equilibrium, which involves same payoffs as VCG

Evolution of Internet Advertising Market

1. This market started out as a "pay per impression" market, where advertising banners were sold on a per-appearance bases, typically in large batches, very expensive, deals negotiated on a case-by-case basis. Very slow, high entry costs.
2. Next came Overture's (now part of Yahoo!) "pay-per-click" first-price auction. Step forward because it lowered costs, but because it was a first price auction, it was susceptible to "gaming" and had very unstable revenues– bidders could use robots to game the system.
3. Finally Google introduced GSP, which stabilized revenues, and the market has converged to this mechanism.

Supposedly VCG would further reduce incentives to game the system, but search engines haven't switched, probably because VCG's more complicated to explain, and switching costs are high.

So how does GSP work?

For each keyword, K bidders bid on N ordered advertising spots on the search page. Each spot i has a click rate of α_i, and each advertiser k has a value per click of s_k. Thus the payoff to bidder k winning spot i is $\alpha_i s_k - p_{ik}$, where p_{ik} is the payment if he wins. Bidders bid their willingness to pay-per-click. The spots are allocated in order of bids, and the winner of spot i pays $\alpha_i b^{(i+1)}$, or in other words, the per-click payment is the next-highest bid (hence second price).

By contrast, the VCG mechanism has the winner paying the externality imposed on the other bidders by taking that spot, so the payment in the VCG mechanism is $p^{V,(i)} = (\alpha_i - \alpha_{i+1})\, b^{(i+1)} + p^{V,(i+1)}$.

Analysis of GSP

Although the setting of these auctions is dynamic, the authors avoid the folk theorem and cut down on the number of equilibria by restricting attention to locally "envy-free" equilibria of the static game. The make an argument that the dynamic game would converge to this. They show that any envy free equilibrium has at least the revenue of the VCG mechanism.

Next the authors study the generalized English auction that corresponds to GSP. They find that if they restrict attention to continuous strategies, there is a unique equilibrium that has same payoffs as VCG.

10.3 *An Empirical Study of an Auction with Asymmetric Information*, Hendricks, Porter

Note: Glenn discussed this paper extensively in class, so [the summarizer] added a few insights to his discussion.

Objective and Hypothesis

- Objective: to see to which extent the auctioning model with asymmetric information in which a single agent has an informational advantage over the other N bidders is consistent with the bidding attitudes in auctions for drainage leases. Check if evidence seems to follow such a Bayesian Nash equilibrium.
- Two different tracts: wildcats (where no firm has an informational advantage) and drainage (where neighbor firm has private information).
- Bidding equilibrium should be different in wildcat and drainage tracts: average profits and tract value should be lower for wildcats, but yet participation should be larger, because there are no informational asymmetries. Furthermore, in drainage tracts, bidding strategy should be different for neighbor (informed) and non-neighbor firms. After displaying some summary statistics that suggest the former is true, this paper concentrates on the latter aspect.
- Assume that there is just one neighbor (informed agent). Indeed, evidence shows that most of the times just one neighboring firm bids (they seem to coordinate on bidding).
- Hypothesis: the paper states some stylized facts true of formal auction models with one informed agent and contrasts them with empirical evidence (see table below).

Data Source

- US (Louisiana and Texas) offshore drilling rights auctions between 1959 and 1969.
- 114 tracts are selected among drainage tracts.
- Date and size (acreage) of the tract auctions, bid for each participating firm, reservation price announced by the government.
- Estimate *ex post* gross profits: annual production - costs of drilling
- The *ex post* gross profits of the adjacent tract represent the estimate of the true tract value of a tract and are considered public value. However, the true *ex post* gross profits of that tract are only known by the informed agent.

Model Used

- First-price sealed auction with reservation price R with one agent with asymmetric information

$$ValueDrainageTract = \begin{cases} v & \text{if neighbor} \\ v - c & \text{if not neighbor} \end{cases}$$

  - Non-neighbors have expectations over $v - c$ based on a public signal.

$$E\left[v - c \middle| z\right] \geq r \text{ with } v \sim F\left[\underline{v}, \bar{v}\right].$$

  - One informed bidder (neighbor) sees v and z.
  - Prop 1: There is no BNE in which uninformed bidders never bid (intuition: given this deviation, informed bidder bids no more than R when it is worthwhile do so. The uninformed bidders could always bid $R + \varepsilon$ and be better off, however.)

- Prop 2: There is no pure strategy BNE with $b_U^*(z) \geq R$ (intuition: if uninformed have a fixed strategy based on z, because z is publicly observed, the informed bidder can infer from it b_U^* and bid $b_U^* + \varepsilon$)
- In equilibrium, uninformed mix over 0 and $(R, \bar{v})$. The informed bidder's strategy is

$$b_I^* = \begin{cases} 0 & \text{if } v < R \\ R & \text{if } v \geq R \text{ but } v \text{ not too big} \\ g(v,z) & \text{otherwise} \end{cases}$$

- where g is continuous and monotone.

- Use the concept of winner's curse: non-neighbor firms have an incentive to bid more conservatively to avoid the problem of bidding too high when the tract yields negative profits.

Conclusions

- Results

Formal model predictions	Evidence
1. $\Pr[b_I^* = 0] < \Pr[b_U^* = 0]$	True. 17% < 32%
2. I wins with prob $\geq \frac{1}{2}$	~ but not astonishing. 62%
3. $E[\pi_U^*] = 0$, positive if $b_I^* > 0$, negative otherwise	True
4. For $c \approx 0$, bid distributions for I and U the same	Not tested
5. b_I^* independent of N	True. Regressing neighbor bids on N brings no significance.
6. $E[\pi_I^*] > 0$	True
7. $g(v,z)$ monotone increasing	Not tested

- Conclude that the formal auction model seems to be consistent with the bidding attitudes shown by the firms in the context of public offshore drainage lease auctions.

10.4 *Information Impact and Allocation Rules in Auctions with Affiliated Private Values*, Kagel, Harstad, Levin

Objective and Hypothesis

This paper aims to test theoretical predictions of auction theory in a laboratory experiment. The predictions are:

(1) In first price sealed bid auctions with affiliated values, the symmetric equilibrium for risk neutral agents will yield higher revenue if a public signal on values is revealed.

(2) The expected revenue in English auctions is higher than in first-price auctions.

(3) Second-price sealed bid auctions result in the same dominant strategy as in the English auction

The authors compare their findings with other "rule of thumb" models. Rules of thumb for first price auctions:

(1) Linear discounting to the value

(2) Sophisticated markdown bidding: complicated, but I think the point is that the distribution is taken into consideration and the bidder is strategic but bids as if he has the highest value.

(3) Nash equilibrium with risk neutrality

(4) Nash equilibrium with risk aversion

Experimental Design

In each experiment, there are six participants.

First-price auction

First, a mean x_0 is drawn from some distribution. Then participants receive private values randomly drawn from $[x_0 - \varepsilon, x_0 + \varepsilon]$, where $\varepsilon \in \{\$6, \$12, \$24\}$ is known (but varies across treatments). Agents are asked to submit bids.

Public signals include

(1) Another random draw from the distribution

(2) The value of x_0.

English auction and second-price auction

Second price auction was carried out the same way as the first price auction. The English auction had decreasing bid increments towards the end (when few bidders remained).

Results

(1) First price auctions: Nash equilibrium wins over the other rules of thumb.

(2) Large doses of public information raised revenue, but less than theory predicts (probably risk-aversion) and not very reliable (explained by individual mistakes).

(3) English auctions are as theory predicts (after brief learning period), but second-price auctions are not and raise more revenue.

(4) The revenue-gains from English/second-price auctions compared to first-price (with affiliated values) are comprised by risk aversion, and more severely in English auctions (not better for $\varepsilon \in \{\$12, \$24\}$.)

10.5 *Using Field Experiments to Test Equivalence Between Auction Formats*, Lucking-Reiley

This paper tests the revenue equivalence theorem (between second-price sealed-bid and first-price English auctions) and finds that collectible trading cards sold over the internet get 30% higher revenues using the Dutch auction format as opposed to the first-price auction format to which it is strategically equivalent.

Vickrey's revenue equivalence result requires risk-neutrality and independent private values. With risk-averse bidders, the first-price auction will receive larger revenues. With risk-neutral bidders with "affiliated" values ("affiliated values" is from Milgrom-Weber 1982 and it means that if my value is high then it is likely that your value is high, too), the first-price auction should yield lower revenues than the Dutch auction.

Lucking-Reiley runs a *field experiment* in which the "subjects" are real card collectors who do not know they are participating in the experiment - to empirically test revenue equivalence. He purchased $1500 worth of playing cards and sold them via auctions on the internet. Interestingly, he actually made a small return on the sale of the set of Magic cards. (He purchased the set for $1500 and received $2000 by selling the cards individually and in pairs.)

The key expression is the following (standard errors in parenthesis):

$$REVDIFF = \underset{(0.13)}{0.25} + \underset{(0.18)}{0.13} \cdot FD$$

FD is a dummy for whether the playing card was sold in a first-price auction first and then a Dutch auction, or vice-versa (this is statistically significant, which is good, since the order should not matter, but the point estimate is still large, which might be a small concern). The intercept represents the difference in revenue between the Dutch auction and the first-price auction. It is statistically significantly, even though the revenue equivalence theorem suggests that it should be zero.

11 Patents and Technology Diffusion

11.1 *Patent Buy-Outs: A Mechanism for Encouraging Innovation,* Kremer

Distortions associated with patents and direct government support of research, need for alternate mechanisms to encourage innovation. Patent buyout: the private value of patents are determined using an auction.

Historical examples

- There really isn't any data source in this paper, but Kremer provides examples of distortions with current system and also historical examples of patent buyouts.
- His examples of distortions are from the pharmaceutical and medical industries.
- In 1837, Daguerre invented photography by developing the Daguerreotype process. In July 1839, the French government purchased the patent, put it in the public domain except in England, and the method was rapidly adopted (but this is different from what we use these days: it follows from the colloidotype process, which was invented twenty years later).
- Whitney sold the rights to the cotton gin to South Carolina (1802)

Mechanism

- Patent holder decides to apply mechanism
- A sealed-bid, second-price auction
- Government offers to buy at markup private value
- If patent holder accepts offer, government randomizes and sells to the highest bidder with probability p.

Afterwards, Kremer discusses the informational advantage for the inventor (e.g. private information about the value of the patent, or the inventor being a low-cost producer) and argues that the incentives for marketing and development will be higher under patent buyouts. He then outlines necessary rules to deal with substitute or complementary patents, mechanisms to deal with collusion. ([The summarizer] doesn't think we have to worry about this section: it's more like a collection of related, but random, topics.)

Conclusion

- Patent buyouts involve more government discretion than the current patent system, but less than government funding of research through NSF or NIH.
- Potentially increase incentives for original invention closer to their social value; reduce incentives for wasteful research; eliminate monopoly pricing distortions
- Could cause a number of problems, e.g. collusion to raise buyout prices.
- First try on a limited basis: pharmaceutical industry.

11.2 *On the Division of Profit in Sequential Innovation*, Green, Scotchmer

Basic Setup

The authors are studying the interplay between optimal patent breadth and length. The consider a two-step process of innovation:

- Step 1: firm 1 invents a product with quality x
- Step 2: firm 2 has an idea for an incremental product with quality $x + y$

The key question is if the patent protects all incremental innovations with quality $\leq x + y^*$, how should y^* be chosen? The timing of the interactions between the firms is the following:

1. Firm 1 invents x at cost c_1
2. Firm 2 has idea an idea for an incremental product with quality $x + y$, but cost c_2
3. If the idea would infringe on the patent, firm 2 can either bargain with firm 1 ex ante (before sinking investment cost) or ex post (after sinking investment cost)

Main points

The main points made in the paper are:

- Ex ante licensing agreements are better than ex post, since ex post firm 2 has already sunk its cost, and wouldn't be able to negotiate enough of the surplus to justify making the investment.
- You don't want to set the patent breadth to $y^* = 0$
- Patent length T should be longer than that necessary to achieve exactly zero profits when you have innovations in more than one firm
- If there is no uncertainty about the value and the cost of the incremental idea, $y^* = \infty$ is best
- If there is uncertaintly about the value, but not about the cost of the incremental idea, then $y^* < \infty$ is best.
- Optimal anti-trust policy has the following trade off: if you want to prevent collusion by not allowing licensing, you have to have a longer patent life to still give firms incentive to invest. Thus you trade off the distortion from collusion with the distortion of long patent life.

11.3 *Innovation, Variety, and Patent Breadth*, Hopenhayn, Mitchell

This article sets up a model in which the social planner maximizes the welfare gains from patents by choosing their breadth, lenght, and a use fee for the patent. The main results are that, under certain conditions, there should be no fee for patents since the breadth and length mechanisms are sufficient to make users sort by type of invention.

Model:

- Ideas $\theta \in \Theta = [1, ..., J]$ arrive with probability $g(\theta)$, where θ indexes the "type" of the idea. These cost c to implement.
- B is breadth (the set of objects that at any time may be prevented by the patentholder) of patent, and T is the length of the patent.
- Firm profits are $\Pi(B, T, \theta)$: increasing in B and T
- Society's benefit from an idea is $S(B, T, \theta)$: decreasing in B and T.
- Society can charge $F(\theta)$ to obtain a patent. Crucially, this does not affect welfare in the planner's problem, which is:

$$
\begin{aligned}
\max_{B(\theta),T(\theta),F(\theta)} \sum_{\Theta} & S(B(\theta), T(\theta), \theta)\, g(\theta) \\
\text{subjet to: } \Pi[B(\theta), T(\theta), \theta] - c - F(\theta) &\geq 0 \quad \text{(IR)} \\
\Pi[B(\theta), T(\theta), \theta] - c - F(\theta) &\geq \Pi[B(\hat{\theta}), T(\hat{\theta}), \theta] - c - F(\hat{\theta}) \quad \text{(IC)} \\
F(\theta) &\geq 0
\end{aligned}
$$

What we have here is essentially a revelation problem: The planner tries to get the inventors to reveal their types so that it can assign them the lowest B and T possible while still providing them with incentives to invent. It turns out that if the following sorting conditions hold:

$$
\frac{\partial^2 \Pi}{\partial B \partial \theta} > 0, \; \frac{\partial^2 \Pi}{\partial T \partial \theta} < 0, \; \Pi \text{ monotonic in } \theta
$$

the optimal user fee $(F(\theta))$ *is zero.* Interpreting θ as profitability, this means that very profitable ideas values breadth more and duration less. The intuition for this is clear: The planner can achieve full sorting with just the B and T instruments, and $F(\theta)$ just tightens the IR constraint.

Note that it is not clear why we should believe these conditions: one would think that both breadth and duration might both be increasingly valuable with profitability (in this case a user fee is optimal). The authors then specify two models in which this might actually be expected to hold:

1. **Product Ladder:** Innovations arrive (randomly) in different markets and improve upon the existing technology. If the improvement is greater than B, the previous product becomes obsolete. If θ indexes the rate of innovation in a particular market, then high θ inventions will prefer greater B protection (and not care so much about T) and the sorting conditions hold.
2. **Horizontal Differentiation/Imitiability:** Once a good is invented it can be imitated with cost $m(\theta) > c$ which is located B units away in competition on a line. Here $m'(\theta) < 0$ so θ indexes how competitive the market is. Thus firms with higher θ prefer great B to get higher profits from more differentiation. Conversely, T does not matter so much if competition is high. (A few conditions apply...)

The paper concludes by advocating a system of patent breadth sorting rather than a system of patenting fees. This may seem difficult in practice (since breadth depends a lot on the context) but the authors provide one (Glenn-friendly) example of a case in which the Supreme Court legislated that Congress can define a menu of patents. The example is a case where one guy patented a dye filtration system that works for PH of 6-9, and another guy tried to patent one for PH 5.

11.4 *Rules of Thumb for Social Learning*, Ellison, Fudenberg

Overview

This paper is primarily concerned with the diffusion of technology when decisions are based on rules of thumb using the experience of one's neighbors. The learning environment has the following three features: "First, agents observe both their neighbor's choices and the payoffs that these choices generate. Second, agents periodically reevaluate their decisions, as opposed to making a once-and-for-all choice. Third, [the authors] consider the possibility that players may be sufficiently heterogeneous that under full information they would not all make the same choice." In class, we dealt with the simple model in which all individuals are homogeneous. This summary will focus on that model.

The second half of the paper concerns itself with heterogeneous populations and nonlinear technologies and how the two interact with the use of popularity weighting rules of thumb for making decisions about technology adoption. [The summarizer] did not understand what was going on here.

Simple model with no popularity weighting

Suppose there are two possible technologies, f and g. Everyone receives the same payoffs u_t^f and u_t^g from using each technology at time t. These payoffs are such that

$$u_t^g - u_t^f = \theta + \varepsilon_t,$$

where θ is unknown and $\varepsilon_t \overset{iid}{\sim} H$. Let $p = \Pr\left[u_t^g - u_t^f \geq 0\right] = 1 - H(-\theta)$. Suppose initially, a fraction x_0 of the players are using technology g. Each period, a fraction α of the players are selected at random to re-evaluate their choice of technology. Suppose the decision is made as follows. Each individual who is re-evaluating his/her choice of technology chooses whichever technology had the higher expected return the previous period. Using such a choice rule, the distribution of technology adoption follows a simple markov chain

$$x_{t+1} = \begin{cases} (1-\alpha)x_t + \alpha & \text{with probability } p \\ (1-\alpha)x_t & \text{with probability } 1-p \end{cases},$$

where x_t is the fraction of players using technology g at time t.

A simple theorem in markov processes shows that $\operatorname{plim} x_t = p = 1 - H(-\theta)$. In particular, we see that, even though one technology may be better than the other, the inferior one will still be adopted with positive probability in the long-run.

Simple model with popularity weighting

As in the previous model, there are two technologies f and g, and payoffs are given by u_t^f and u_t^g. Suppose each period, a fraction α of the players are to re-evaluate their choice of technology. Further, suppose they use the following rule of thumb: choose technology g if

$$u_t^g - u_t^f \geq m(1 - 2x_t),$$

where x_t is the fraction of players using technology g at time t. Here, we will have that the probability that the individual chooses technology g is given by

$$\Pr[\theta + \varepsilon_t \geq m(1-2x_t)] = \Pr[\varepsilon_t \geq m(1-2x_t) - \theta] = 1 - H(m(1-2x_t) - \theta).$$

Now, the stochastic process $\{x_t\}$ describing the fraction of players using technology g over time is given by the time-inhomogeneous markov chain

$$x_{t+1} = \begin{cases} (1-\alpha)x_t + \alpha & \text{with probability } 1 - H(m(1-2x_t) - \theta) \\ (1-\alpha)x_t & \text{with probability } H(m(1-2x_t) - \theta) \end{cases}.$$

If we assume that $\varepsilon_t \sim U[-\sigma, \sigma]$, the convergence results of this markov chain depend on the value of m chosen.

Proposition 1 (1) *If* $m = \sigma$, *then for any* x_0, $\operatorname{plim} x_t = 1$ *if* g *is better than* f. *(i.e. if* $\theta > 0$*)*

(2) *If* $m > \sigma$, *then we have that if* $|\theta| \geq m - \sigma$, *for all* x_0, $\operatorname{plim} x_t = 1$ *if* g *is better than* f. *If* $|\theta| < m - \sigma$, *then if* x_0 *is sufficiently high,* $\operatorname{plim} x_t = 1$, *and if* x_0 *is sufficiently low,* $\operatorname{plim} x_t = 0$.

(3) *If* $m < \sigma$, *the system converges to a nondegenerate ergodic distribution. (i.e. everything has positive probability.)*

By allowing for the right amount of popularity weighting, efficient technology choice can be reached with probability one. If everyone places too much emphasis on popularity, then the initial conditions matter. Consider the situation where everyone ONLY cares about popularity (i.e. $m = +\infty$). Then $H(m(1 - 2x_t) - \theta)$ is equal to 1 if $1 - 2x_t > 0$ (since $H(+\infty) = 1$). Thus, we see that if $x_0 < \frac{1}{2}$, since all people care about is which technology other people are taking, the system will eventually converge to $\underline{x} = 0$. Similarly, if $x_0 > \frac{1}{2}$, we will have convergence to $\bar{x} = 1$. Finally, if not enough emphasis is placed on popularity, the system will drift forever, and we will never reach a steady state. The extreme case of this is when $m = 0$, which corresponds to the previous section on "no popularity weighting."

Remark 2 *If players are allowed to account for the autocorrelation in the decision structure, (i.e. account for "trends") then convergence can occur much more quickly.*

12 Bounded Rationality

12.1 *Bounded Rationality in Industrial Organization*, Ellison

Simon - consider limited computational and informational processing capacities of organisms - satisficing rather than maximizing. (For a firm, satisficing means that a firm just wants to achieve profits above some threshold instead of achieving maximum profits.)

Early literature - focus on firms rather than consumers being irrational.

Rule of thumb approach - instead of fully rational model, assume that consumers follow simple rules of thumb; Smallwood and Conlisk (see lecture notes) - this model shows that rule of thumb approaches can lead to socially optimal outcomes.

Team theory - many individuals with private information working together; the best way to coordinate may be a decentralized mechanism; try to provide information and cost based microfoundations to guide rules of thumb; Radner examines firm with limited managerial resources to devote to cost reduction.

Empiricism - Joskow estimates probit model of profits, but doesn't really need to assume profit maximizing foundation; Ellison and Chevalier - empirically determine how consumers choose mutual funds, then solve firm's problem (same with consumers in Green and Porter)

Learning - Mobius - competition in local telephone service - several islands, everyone interacts disproportionately with others on their island, but occasionally switch; predicts rise and fall of independents; incorporates social networks, would be impossible with fully rational model.

Complexity - Fershtman and Kalai - firms avoid some actions if they make strategy space too complex.

Della Vigna and Malmendier - health clubs offer two part tariff; up-front payment L and per visit charge p; distort p below marginal cost (allows sophisticated consumers to go more often, increases gap between expected benefit of naive consumers and what they actually get); people must value commitment or overestimate, since they could get cheaper contracts.

Heidues and Koszegi: loss aversion $\Rightarrow$ firms should charge constant prices

Some thoughts - can regard the things a firm commits to in a contract as quality choices (i.e. per visit charge from health clubs affects quality of membership); monopolists will distort product characteristics along whatever dimensions increase irrational consumers' willingness-to-pay - distortion not eliminated by competition (but surplus returned to consumers)

Ellison and Fudenberg - two technologies, fraction alpha consider switching depending on market share; popularity weight determines whether play converges to one technology in equilibrium or if both have positive shares.

Ellison and Fudenberg - people ask k others what technology they use; if k is small, herding on inefficient technology is possible; when k is large, everything might survive in the long-run

Baye and Morgan - in an ε equilibrium, we can get big departures from Bertrand competition (if ε is 1% of the monopoly profit, such an equilibrium can yield 26% of monopoly profit for each firm)

Obfuscation - Gabaix and Laibson - obfuscation increases uncertainty about product's valuation; Spiegler - many dimensions of each product, firms evaluate it along one; firms randomize across dimensions making the product good along some and bad along others; obfuscation increases with more firms and more attractive outside options.

Performance evaluation - Sobel: candidate is successful if achieves a standard achieved by a fraction of previous successful candidates; if measurement changes over time, quality declines in equilibrium.

Ellison - evolving standards; overconfident editors assume more weight being placed on packaging of articles, which leads this to be realized.

12.2 *Contract Design and Self-Control*, Della Vigna, Malmendier

Summary

Note: This paper was really long, so [the author of this summary] left some extensions out that, according to [his/her] best judgment, are not likely to come up in Glenn's questions.

Objective: model rational firm behavior in the presence of consumers with time inconsistent preferences, focusing on contract design. The pricing of two types of goods is considered: leisure goods (immediate benefit, delayed costs) and investment goods (immediate costs, delayed benefits).

General evidence that consumers are being "manipulated": Back-loaded fees, contracts with automatic renewal/artificial costs of switching/cancellation costs.

Model

About quasi-hyperbolic discounting

In general, quasi-hyperbolic discounters evaluate their future utility in each period ad $U = u_0 + \beta \sum_{t=1}^{\infty} \delta^t u_t$, $\beta < 1$. Thus, they have an additional discount factor between now and $t+1$, but have a higher constant discounting factor between periods after that. All consumers are aware of this extra discount factor today. In addition, all consumers have beliefs $\hat{\beta}$, which represents their belief about their future discount rate between t and $t+1$.

Time consistent (exponential) individuals have $\beta = \hat{\beta} = 1$. Naive hyperbolic discounters have beliefs $\hat{\beta} = 1$, while their true β is less than 1. Sophisticated individuals have $\beta = \hat{\beta} < 1$. Partially naive agents have $\beta \neq \hat{\beta}$ and both are less than 1.

The paper's model

It is assumed that consumers are homogeneous. The firm is a monopolist (the qualitative results are unaffected when extended to the competitive case).

$t = 0$: Firm proposes two-part tariff (L, p). Consumer accepts/rejects (reservation utility $\bar{u}$, 0 profits in this case)

$t = 1$ (if accepted): Consumer pays L and learns cost c of consumption drawn from some distribution F. The consumer chooses to consume (C) or not to consume (NC). If the agent consumes, she gets $c - p$. Otherwise, she gets 0.

$t = 2$: Consumer gets benefit b if chose C and 0 if choose NC.

For consumption goods, $c, b > 0$. For leisure goods, $c, b < 0$.

Consumer overconfidence is defined as the difference between what the agent thinks she will consume in the future and what she ends up consuming. Time consistent individuals are not overconfident. Neither are sophisticated hyperbolic discounters - although they do not consume the optimal amount, they are aware of what the outcome will be. Naive individuals are overconfident - they think they will consume less of the leisure good and more of the investment good than they actually do.

A monopolist will set p below marginal cost for an investment good and set it above marginal cost for a leisure good (the standard result in a two-part tariff with known demand curve is to price at marginal cost). The fixed fee will be correspondingly higher and lower, relative to the base case. This also holds with perfect competition.

If c is known at the time of the decision, the result is a little more complicated, but skewed pricing is still a possibility.

Welfare

If consumers are time inconsistent but sophisticated, the contract design is optimal. Contracts for investment goods lower the marginal costs to consumers, inducing them to "invest" (i.e. go to the gym) more. Contracts for leisure goods increase the costs, which makes them get less leisure than they otherwise would have (they would have overconsumed if costs were lower).

If consumers are time inconsistent and naive, welfare is lower, because firms exploit their naivete and extract extra surplus from them. Welfare is also lower under partial naivete. Government intervention may be helpful, but the authors are skeptical of the government's ability to know the parameters of the model and suggest that educating the naive individuals may be a better strategy. Intervention may still not fix the adverse effects of naivete.

Empirical and Anecdotal Evidence

Health Clubs

The authors surveyed 64 health club companies in the Boston area (which operate 97 clubs). They asked about the contracts they provide and the cancellation policies. Most clubs had monthly, annual, and pay-per-visit options. The monthly and annual contracts had no per-visit fees. The authors computed the marginal cost of a visit (providing towels, personnel, fixing machines) to be between $3 and $6, thus showing below-cost pricing. Historically, health clubs charged per-visit fees. The authors suggest that firms learned the optimal contract in the industry over time.

Vacation time-sharing

The authors view planning a vacation as a cost and the vacation itself as a benefit. Thus, buying a time-share is like buying an investment good. They cite the large fixed costs and low per-vacation price of using the resort as evidence of strategic contract design.

Las Vegas

Going to Las Vegas is a leisure good - the low initial costs of a cheap hotel room or vacation package are followed by high "costs" of gambling - i.e. consumers end up gambling more than they had planned. The good that is being priced above cost is gambling, since the odds are not fair.

Credit cards

The authors cite evidence of 10 major credit card companies that provide cheap credit (no annual fee, despite the fact that credit cards are not costless to provide) and low introductory interest rates followed by high interest rates 6-9 months after the credit card account is first opened. The competing explanation for this is that the cost of default is high, but the fact that credit card debt is resold at a 20% premium makes the authors think this is unlikely. There is evidence of above-cost pricing (what's being priced above cost is borrowing - clearly, consumers who pay their balance back are unaffected and the company loses money on them).

Cell phones

Cell phones are also an example of a leisure good - the initial package that includes X minutes is priced relatively lower than the per-minute fee after the X minutes are used up. Naive consumers do not anticipate how much they will use the cell phone when they choose plans.

Extensions

Switching costs: allowing consumers to choose whether to continue with the firm costlessly (and making the same C-NC decision) or switch to another firm at a cost has distortionary effects when there are naive consumers who underestimate their renewal rate. Time-consistent and sophisticated individuals anticipate their renewal rates correctly, even though they may have different renewal rates.

Welfare of switching costs: under perfect competition, the per-period welfare is the same as in the basic model. Under monopoly, the firm can make more money from partially and fully naive consumers.

Heterogeneity: the authors do not model it formally, but they briefly consider the implications of two types of consumers (who could have any degree of time-inconsistency). Firms will adjust their contracts to account for heterogeneity. Switching costs can now have efficiency effects.

Evidence: automatic renewal of credit cards at the end of the introductory period makes cancellation costly. Rational switching cost models would not predict the back-loaded structure of the charges (high interest rates later on); same with mail order, health club, newspapers, and life insurance.

12.3 *The Market for Quacks*, Spiegler

Quick Summary

This is a bounded rationality model where n quacks play a price competition game facing a continuum of patients who recover with probability α. (this is like getting utility $1 - p_i$ with prob α and $-p_i$ with prob$1 - \alpha$). Patients are not rational and follow an anecdotal rule: asks one person who has used each of the products and the maximize $x_i - p_i$ over the anecdotes x_i that you hear - the true playoff is $\alpha - p_i$. With this assumption, this is an active market for quacks (they do not go out of business). Here, rules of thumb do NOT lead to socially optimal behavior even if we allow for things like 1) having a more qualified (but still profit maximizing) healer in the group and 2) allowing quacks to disclose their success rates.

Theoretical results

1. There is a unique MSNE in prices set by quacks. The expected price is strictly decreasing in α, the probability that the treatment works - eg, they charge higher prices if the condition is more incurable (since there are likely to be fewer successes in the sample, this weakens price competition)

2. The relationship between welfare loss and number of quacks is non-monotonic - more competition can increase welfare loss - this is because more quacks lead to lower prices but also to a higher probability of hearing a success that thus spurs demand.

3. Introducing a good doctor (with higher success probability α_e does not get rid of the quacks. This is because the good doctor draws clients away from the default and not from the quacks and ends up exploiting patients too.

4. Allowing doctors to disclose their success rates. This does not happen because in this model of decision making, the consumer's anecdotal decision procedure does not interpret not disclosing as a sign of incompetence so quacks all pool at not disclosing.

12.4 *Price Dispersion in the Lab and on the Internet*, Baye, Morgan

This paper has both a theoretical and an empirical contribution. The theory part develops a simple model that shows that modifying textbook Bertrand competition to include firms which only maximize profits to within ε (where ε can be arbitrarily small) can result in surprisingly large profits (achieved through mixed strategies).

The theory starts with a textbook Bertrand model (homogeneous products, identical costs). It is well known that the only pure strategy NE is that the firms will price at cost. Using the ε equilibrium concept (that firms will only deviate if they can make more than ε), the authors show that no pure strategy equilibrium can result in profits greater than ε.

The surprising theoretical result is that there is a symmetric mixed strategy ε equilibrium in which each player earns an expected profit that exceeds ε - for even small ε, the profits can be large fractions of the monopoly profits. As an example, with an ε of 1% of monopoly profits (so firms only deviate if they make more than 1% of the ex ante monopoly profits), the mixed-strategy equilibrium results in industry profits (per firm) of 26.4% of monopoly profits.

The theory therefore shows that a small amount of bounded rationality can result in (1) significant price dispersion (since the firms are playing mixed strategies) and (2) much larger profits that can be large fractions of the monopoly profits.

[The summarizer] thinks that this is a truly fascinating model, but it is hard for [the summarizer] to think about whether it is a good way to think about price dispersion. We don't really think firms randomize. We think that mixed strategies are a "theoretical shortcut" that really represent small amounts of imperfect information (i.e. Harsanyi's purification interpretation of mixed-strategies).

The empirical part of the paper attempts to fit the data to the standard model of Bertrand competition and the model of price dispersion that they introduce in the theory section. The data is taken from two independent laboratory experiments and a leading internet price comparison site. All of the data sets show that the bounded rationality model of price dispersion fits the data better. As they write in the discussion " a little bounded rationality goes a long way towards rationalizing the price dispersion observed in homogeneous product settings." The data sets included different treatments, subject nationalities, and different experimental designs. All data supported the models of bounded rationality more than alternative models that the authors test.

13 Advertising

13.1 *A Memory Jamming Theory of Advertising*, Shapiro

Coming before generals.

13.2 *The Effect of Price Advertising on Prices*, Milyo, Waldfogel

Quick Summary

They collect pre and post data (1995-1997) surrounding the Supreme Court decision to lift the ban on liquor advertising in Rhode Island. They have 3 controls for RI retail prices: MA retail prices, RI wholesale, and MA wholesale prices and basically run a Diff-Diff across the policy change. They find that advertising stores cut the prices of only the products they were advertising a lot, and of the products rival stores were advertising a little, but that non-advertising stores did not cut prices really. Advertising (at least newspaper advertising) does seem to increase quality as proxied by lottery ticket sales. There was no evidence that price dispersion across stores fell (which is sometimes a theoretical prediction). There is a small and insignficant drop in prices overall. We do not know anything about volume sold.

Data

Panel data on 115 different liquor stores in RI and MA, observations on 33 beverages, visited stores about quarterly. Also data on wholesale prices from trade publications, and state lottery ticket sales. Prices were collected secretly, sometimes by memorizing. Wholesale data did not include quantity discounts. Geographic area was southern RI, northwest Boston suburbs, and RI/MA border.

Regressions and Tests

1. Test if advertising stores already had lower prices before the ban.

$$pricesbefore = productdum + timedum + advdum.$$

Finds prices were lower at advertising stores and uses this to conclude that advertising can act as a signal to consumers.

2. Test for aggregate effect on prices.

$$price\_product_d\_state_s\_time_t = storeFE + productFE + product * RI + post + post * RI$$

The interaction on post and treatment is basically 0.

3. Prices of products at stores that do not advertise "nonadvertising store effect" - find near 0 for this
4. Non currently advertised products at stores that currently advertise other products "advertising store effect" - find that they are not discounted
5. Currently advertised prices of products at stores that advertise them "advertised products effect" - are steeply discounted about 20 percent off.
6. Do advertisers sell more? Look at lottery sales and conclude yes.
7. Is there less price dispersion - no.

13.3 *Persuasive Advertising with Sophisticated but Impressionable Consumers*, Lauga

Lesson #1 from paper: Get your job market paper professionally edited. While you may not think grammar is very important, it sure makes a paper hard to read and understand.

This paper builds on Shapiro (2006) "A 'Memory Jamming' Theory of Advertising" by allowing consumers to be aware of being exposed to persuasive advertising and prices to be strategically chosen by firms (signaling through both price and advertising quantities).

Some major points of the paper:

- Firms engage in persuasive advertising even if consumers are fully aware of the manipulative nature of this advertising (in contrast with some previous models, where the advertising equilibrium would break down if consumers were less naive)
- The relationship between type of advertising (informative v. persuasive) is not not necessarily monotonic with respect to price or quality
  - Depending on market conditions, persuasive advertising can signal a high, intermediate, or low quality range
  - Firms use both price and advertising to signal quality
- A monopolist may actually prefer to withhold information and instead engage in persuasive advertising

Basic Model assumptions:

- Monopolist with quality $q \in [0,1]$ from distribution f with support $\left[\underline{q}, \bar{q}\right] \subset (0,1]$; WLOG $c = 0$
- Unit mass of consumers that receive signal of quality; utility $q - p$; f known; receive signal $s \in \{0 = bad, 1 = good\}$
- $\Pr(s = 1|q) = \alpha(q) \in (0,1)$ is an increasing function of the true product quality
- Worst quality generates few good signals

Glenn highlighted these points in class (from notes):
Changes:

- Firms can set prices
- Firms are ex-ante heterogeneous
- Consumers observe advertising levels

Observations:

- Advertising signals low quality in replacement model.
  - return to memory replacement more valuable to low firm (since consumers have more bad memories)
- Advertising can improve rent extraction by increasing the concentration of good memories for the marginal type (consumer with marginal signal of quality about firm's product)