14.121 Lecture 4: Producer Theory and Monopoly

Parag Pathak
MIT

September 2019

Producer Theory

There are several branches of theory about firm behavior

1) Theory of the firm (internal organization, effects on behavior)

2) Monopoly pricing (includes price discrimination)

3) Imperfect competition (small number of firms, deadweight loss)

4) Neoclassical producer theory

The first 3 are very active and important today. What is 4?

Chapter 5 of MWG is a good summary.

One slight difference from undergraduate courses is production sets versus production functions.

Let $y = \{(y_1, ..., y_L) \in \mathbb{R}^L |$ net use of each is feasible$\}$.

Writing positive and negatives this way gives an elegant formulation of the **profit maximization problem**:

$$\pi(\mathbf{p}) = \max_{y \in Y} y \cdot \mathbf{p}.$$

The traditional way of writing is to have one output, y_L, and use a **production function**:

$$f(y_1, ..., y_{L-1}) = \{y_L \in \mathbb{R}^+ | (y_1, y_2, ..., y_L) \in Y\}$$

The other big problem is the **cost minimization problem**:

$$c(p, q) \equiv \min_{z \in \mathbb{R}^{L+1}} p \cdot z, \qquad \text{s.t. } f(z_1, ..., z_{L-1}) = q$$

Straightforward to establish theorems for profit:

a) Profit is homogenous of degree 1 in p
b) Profit convex in p
c) Output homogenous of degree 0

Or about the cost functions:

a) It is homogenous of degree one in p, and non-decreasing in q
b) It is concave in p
c) It is continuous in p

This is done in MWG 5.B-5.C (required reading for this part)

Many courses (e.g., Varian) starts with the producer side because it is easier: don't have to deal with income effects

"Producers are just like consumers, but they maximize profit instead of utiliy."

Why cover in class?

The standard reasons are

a) active area / need to know to build on

b) learn techniques by doing it

c) can give useful comments on methodology

In my view, none apply to MWG's treatment of production, so I'll ask you to learn this material on your own.

1) Demand theory is still of great deal of interest to people who estimate demand curves for prediction and welfare.

   Cost functions were even more widely estimated in the 1950s and 1960s as people built econometric models of economies.

   Today, this is rarely done.

   Part is distrust of cost data due to cost allocation depreciation, and other accounting conventions.

2) Techniques are just like consumer theory

3) Comments on Cobb-Douglas, use just like in utility theory

Hence, we will pursue an alternate approach to introduce some new methodology.

FOC Approach to Comparative Statics

$F : X \times T \to \mathbb{R}$ where $X, T \subset \mathbb{R}$.

$$X^*(t) = \arg\max_{x \in X} F(x, t)$$

FOC approach applies implicit function theorem to FOCs.

Assume:

- Smoothness: F is twice continuously differentiable
- Convexity of X
- Strict concavity: $F_{xx} < 0$ (to get unique maximizer: $X^*(t) = \{x(t)\}$)
- Interiority: For each t, $x(t)$ is in the interior of X

Unique maximizer $x^*(t)$ solves:

$$F_x(x(t), t) = 0$$

Implicit function theorem amounts to differentiating FOC with respect to t:

$$F_{xx}(x(t), t)x'(t) + F_{xt}(x(t), t) = 0$$

Hence,

$$x'(t) = -\frac{F_{xt}(x(t), t)}{F_{xx}(x(t), t)}$$

Potential issues:

- F may not exactly known - want to be robust to specifications
- F and/or X may not satisfy assumptions
- May not isolate qualitative features of F needed

Robust Comparative Statics

Idea: avoid use of calculus and find conditions which can be used even if functions are nondifferentiable, discontinuous, or even defined on discrete state spaces.

Why bother?

One reason is to generalize, i.e., what happens if not convex or do not have local nonsatiation, or differentiability?

Often, we do not care. Some of the theory about making corrections. If differentiable function makes point more easily, then most economists assume it.

Other virtues:

- More easily make assumptions on primitives
- Better intuition
- Need to be able to read others' papers

Some recent papers:

- Costinot, A. "An Elementary Theory of Comparative Advantage." *Econometrica*, 2009. [International]
- Acemoglu, D. "When Does Labor Scarcity Encourage Innovation?" *Journal of Political Economy*, 2010. [Growth / Innovation]
- Kircher, Phil and Jan Eeckhout. "Sorting and Decentralized Price Competition." *Econometrica*, 2010. [Labor Theory]
- Segal, Ilya and Michael Whinston. "Property Rights." Chapter for Handbook of Organizational Economics, 2011. [Organizational Econ]
- Acemoglu, D. and M. Jensen. "Robust Comparative Statics in Large Dynamic Economies." *Journal of Political Economy*, 2015. [Dynamic Macro]

Athey-Milgrom-Roberts starts with an example to illustrate where the old theory may be misleading.

Example: Cost-benefit problem

Suppose

$$x^*(\theta) = \arg\max_{x \in X} B(x, \theta) - c(x)$$

When is $x^*(\theta)$ increasing in θ?

A typical theorem one might see in a micro book is

Proposition

Suppose $X = \mathbb{R}$. If c is differentiable and convex and B is concave in x, differentiable in θ and has $\frac{\partial^2 B}{\partial x \partial \theta} > 0$, then $x^*(\theta)$ is increasing in θ.

Intuition: marginal benefit = marginal cost

$\frac{\partial^2 B}{\partial x \partial \theta} > 0$ means that $\frac{\partial B}{\partial x}$ higher when θ higher

FOC $\Rightarrow$ when θ higher, $\frac{\partial C}{\partial x}$ must be higher

convexity of C (C' increasing) $\Rightarrow$ when θ higher, x higher

Athey, Milgrom and Roberts argue that this is misguided. The problem is that the theorem is true even if C is **not convex**, and B is **not concave** (and x not continuous and functions not differentiable).

Our supposed intuition cannot be identifying the right reason if it makes us think convexity of cost is key. Concave costs might make it go in the other direction.

Increasing Differences

Consider a similar problem. Suppose

$$\pi(x, \theta) = f(x, \theta) + g(x)$$

Let

$$x^*(\theta) = \sup \arg\max_x \pi(x, \theta)$$

The sup is used to make this a function instead of a correspondence.

When is $x^*(\theta)$ non-decreasing in θ? To answer this, think about two possible choices x' and x'' with $x' < x''$. When could an increase in θ make you switch from x'' to x'?

Advantage of x'' relative to x' is

$$\begin{aligned}\pi(x'', \theta) - \pi(x', \theta) &= f(x'', \theta) + g(x'') - \big(f(x', \theta) + g(x')\big) \\ &= (f(x'', \theta) - f(x', \theta)) + (g(x'') - g(x'))\end{aligned}$$

Definition

$f(x, \theta)$ satisfies **increasing differences** if $\forall x', x''$ with $x' < x''$, $f(x'', \theta) - f(x', \theta)$ is weakly increasing in θ.

- incremental gain from choosing a higher x (x'' rather than x') is greater when θ is higher

- special case of f being **supermodular**

Proposition

If f is twice continuously differentiable, then f has increasing differences if and only if $\frac{\partial^2 f}{\partial x \partial \theta} \geq 0$, for all x, θ.

Theorem (Topkis' Monotonicity Theorem)

Let $x^*(\theta)$ be as above. If $f(x, \theta)$ has increasing differences, then $x^*(\theta)$ is weakly increasing in θ.

- Let

$$\begin{aligned}\pi(x, \theta) &= f(x, \theta) + g(x) \\ &= B(x, \theta) - c(x)\end{aligned}$$

  Cost-benefit result did not need differentiability and c convex assumptions.

- With our additively separable objective function:

$$\pi(x, \theta) = f(x, \theta) + g(x),$$

  only have to verify that f is ID.

  g trivially has ID, and the sum of two ID functions is ID.

  Provides sense in which comparative statics are robust to additive perturbations

Is increasing differences the "right" condition for this problem?

People who study robust comparative statics think it is. They argue for this by pointing to a converse with a $\forall g$ in the statement.

Proposition

Define

$$x^*(\theta) = \sup\{\arg\max_x f(x,\theta) + g(x)\}.$$

Then $x^*(\theta)$ is weakly increasing in θ **for all g** *if and only if f has increasing differences.*

Usually we like sufficiency results that say conclusion C holds if and only if H_1 and H_2 are satisfied – this is not what this says

Rather it says that conclusion C may not hold in all environments

Applications in Producer Theory

Example 1: Capital choice versus prices in cost minimization problem (CMP)

Consider a firms choice of capital in the cost minimization problem

$$\max_{k \in K} -rk - wl, \quad \text{s.t. } F(k, l) \geq q$$

When does the level of capital decrease with an increase in the cost of capital r?

Proposition

The choice of capital $k^*(r, w, q)$ is weakly decreasing in r.

Example 2: Cross price effects in PMP

How does k change with w? Will firms always substitute away from labor and use more capital or can things go the other way for some reason?

Definition

k and l are **substitutes** in production if $F(k,l)$ has decreasing differences. k and l are **complements** in production in $F(k,l)$ has increasing differences.

CES: $f(k,l) = (\alpha_1 k^\rho + \alpha_2 l^\rho)^{1/\rho}$

$$\frac{\partial f}{\partial k} = (1/\rho)\rho\alpha_1 k^{\rho-1}\left(\alpha_1 k^\rho + \alpha_2 l^\rho\right)^{1/\rho-1}$$

$$\frac{\partial^2 f}{\partial k \partial l} = \alpha_1 k^{\rho-1}\alpha_2 l^{\rho-1}\left(1/\rho - 1\right)\rho\left(\alpha_1 k^\rho + \alpha_2 l^\rho\right)^{1/\rho-2} > 0 \text{ when } \rho < 1$$

Cobb-Douglas: $f(k, l) = k^{\alpha} l^{\beta}$

$$\frac{\partial^2 f}{\partial k \partial l} = \alpha\beta k^{\alpha-1} l^{\beta-1} > 0$$

Note: This is a little different from standard treatments that use prices in the definition. Intuition for complements is increasing differences $\approx \frac{\partial}{\partial l}\frac{\partial F}{\partial k} > 0$ i.e. increasing l increases returns to capital.

Consider the standard PMP:

$$\max_{k,l} pF(k, l) - rk - wl$$

Usually solved in steps:

- Solve for $\hat{l}(k, w)$ given w at capital k (holding r constant)
- Plug in, and solve for $k^*(r, w)$ as function of (r, w)
- Plug in k^* into $\hat{l}(k, w)$ to get $l^*(r, w)$

Can also solve as follows:

- Solve for $\hat{k}(l, r)$ given r at labor l (holding w constant)
- Plug in, and solve for $l^*(r, w)$ as function of (r, w)
- Plug in l^* into $\hat{k}(l, r)$ to get $k^*(r, w)$

Proposition (Cross-Price Effects)

a) *If k and l are complements, then $k^*(r, w)$ is weakly decreasing in w and $l^*(r, w)$ is weakly decreasing in r.*

b) *If k and l are substitutes, then $k^*(r, w)$ is weakly increasing in w and $l^*(r, w)$ is weakly increasing in r.*

Robustness to Feasible Set

Suppose we want to characterize function $f : X \times T \to \mathbb{R}$ where the conclusion of Topkis's theorem holds for any feasible set $\bar{X} \subset X$.

Turns out we can get away with a weaker condition

A function $f : \mathbb{R}^2 \to \mathbb{R}$ is **single-crossing** in (x, θ) if for all $x' > x$ and $\theta' > \theta$,

(i) $f(x', \theta) - f(x, \theta) \geq 0$ implies that $f(x', \theta') - f(x, \theta') \geq 0$,

(ii) $f(x', \theta) - f(x, \theta) > 0$ implies that $f(x', \theta') - f(x, \theta') > 0$.

- if f has increasing differences then it is also single crossing, but the opposite is not true.

- unlike ID, this condition is not symmetric in variables (x, θ)

Suppose we have strict single crossing (at most one point):

$$f(x', \theta) - f(x, \theta) \geq 0 \Rightarrow f(x', \theta') - f(x, \theta') > 0$$

Write

$$x^*(\theta, \bar{X}) = \arg\max_{x \in \bar{X}} f(x, \theta),$$

where we make dependence on the constraint set $\bar{X}$ explicit.

Proposition (Milgrom-Shannon Monotonicity Theorem)

If f is strict single-crossing in (x, θ), then $x^*(\theta, \bar{X})$ is weakly increasing in θ for all $\bar{X} \subset X$. Moreover, if $x^*(\theta, \bar{X})$ is non-decreasing in θ for all finite sets $\bar{X} \subset \mathbb{R}$, then f is single-crossing in (x, θ).

Also called the **Monotonic selection** theorem

If we want comparative statics which generate strictly increasing maximizers, then the results depend on differentiability and being in the interior - see Edlin and Shannon (1998)

Multivariate Comparative Statics

Choice variable so far is one dimensional. Consider

$$\max_{(x_1,x_2)\in X\subset\mathbb{R}^2} F(x_1, x_2, t)$$

Earlier univariate version of Topkis's theorem implies

- If F has ID in (x_1, t), then the optimal value of x_1 *holding* x_2 *fixed* is non-decreasing in t.
- If F has ID in (x_2, t), then the optimal value of x_2 *holding* x_1 *fixed* is non-decreasing in t.

Here, both variables are chosen simultaneously, so there are indirect effects ("feedback") arising from the interaction between x_1 and x_2.

Intuitively, if F has ID in (x_1, x_2), the indirect effects work in the same direction as the direct effects
e.g., if x_2 optimally increases in response to an increase in t further increases incentive to raise x_1

When F has ID *all pairs of variables*, all indirect effects reinforce the direct effects

ID in all pairs of variables motivates the condition of **supermodularity**

For $x, y \in \mathbb{R}^n$, define operations **meet** and **join** as follows:

$$x \wedge y = (\min\{x_1, y_1\}, ..., \min\{x_n, y_n\})$$
$$x \vee y = (\max\{x_1, y_1\}, ..., \min\{x_n, y_n\})$$

Sometimes these operators are called the "greatest lower bound" and "least upper bound" of $\{x, y\}$, respectively

Definition

A set $X \subset \mathbb{R}^n$ is a **sublattice** if for all $x, y \in X$ we have $x \wedge y \in X$ and $x \vee y \in X$.

Graphically, when $X \subset \mathbb{R}^2$, this property means that if you take two non-ordered corners of a rectangle whose edges are parallel to the axes are in X, then the other two corners are also in X

Examples:

- Any product set $X = X_1 \times ... \times X_n$
- Any set described by an inequality $x_i \leq g(x_j)$ where g is an increasing function

(It can be shown that any sublattice in $\mathbb{R}^n$ can be described as an intersection of sets of these two types – see Topkis for details)

More generally, a sublattice is a subset of a lattice – since we work with ordering $\geq$ on subsets of $\mathbb{R}^n$ this will be unimportant for us

Definition

A function $f : X \to \mathbb{R}^n$ on a sublattice X is **supermodular** if for all $x, y \in X$,

$$F(x) + F(y) \leq F(x \wedge y) + F(x \vee y)$$

Suppose X is a sublattice of $\mathbb{R}^2$. If $x \leq y$ or $x \geq y$ then $\{x, y\} = \{x \wedge y, x \vee y\}$, so condition has no bite.

Suppose instead that $y_1 < x_1$ but $x_2 < y_2$. Supermodularity means:

$$F(x_1, x_2) + F(y_1, y_2) \leq F(y_1, x_2) + F(x_1, y_2)$$

or

$$F(x_1, x_2) - F(y_1, x_2) \leq F(x_1, y_2) - F(y_1, y_2)$$

Change from increasing first argument from y_1 to x_1 can only go up when the second argument goes up from x_2 to y_2.

Hence, ID implies supermodularity. Generally, if X is a product set in $\mathbb{R}^n$, supermodularity is characterized by ID in each pair of variables holding the others fixed (see ps4)

Now we can state the multivariate version of Topkis' theorem

Theorem (Topkis)

If X is a sublattice, $T \subset \mathbb{R}$ and $F : X \times T \to \mathbb{R}$ is supermodular, then for all $t, t' \in T$ such that $t' > t$, and all $x \in X^*(t)$, $x' \in X^*(t')$,

$$x \wedge x' \in X^*(t) \quad \textit{and} \quad x \vee x' \in X^*(t')$$

- In special case, if $X^*(t) = \{x\}$ and $X^*(t') = \{x'\}$, the statement says $x \vee x' = x$ and $x \wedge x' = x'$, which means that $x \leq x'$ or the maximizer is nondecreasing in t

- Sometimes the conclusion of theorem says that $X^*(t)$ is non-decreasing in t in the **strong set order**: $A \leq B$ in the strong set order if for all $a \in A, b \in B$, $a \wedge b \in A$ and $a \vee b \in B$

The LeChatelier Principle

Samuelson in *Foundations of Economic Analysis* (1947) suggested that a firm would react more to input price changes in the long-run than in the short-run because it has more inputs it can adjust

We treat "short run" as meaning with k fixed, and "long run" as meaning with k allowed to vary.

That is,

$$
\begin{aligned}
l_{LR}(w) &= \sup_l \arg\max_l \max_k pF(k, l) - rk - wl \\
l_{SR}(w_0, w) &= \arg\max_l pF(k_{LR}(w_0), l) - rk_{LR}(w_0) - wl
\end{aligned}
$$

Samuelson's argument was based on small price changes in the neighborhood of the long run price

Consider large price changes. Let the output price be 1. Suppose

$$F(k,l) = \begin{cases} 10 & \text{if } l \geq 2 \text{ or } k,l \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Input price is $\mathbf{w} = (3,2)$.

  Firm maximizes profit by using two units of labor

- Suppose price of labor rises to 6, so $\mathbf{w}' = (3,6)$.
  - SR: firm cannot make positive profit since $2 \cdot 6 > 10$, so it shuts down and factor demands are zero
  - LR: use a unit of both capital and labor, generating profit of 1.

When price of labor rises from 3 to 6, demand for labor changes from 2 to 0 (in short-run) but then is 1 in the long-run $\Rightarrow$ SR impact larger than LR

Can modify example to make production function less unusual

There is an interesting set of economic models where the long-run responses to price changes are larger than short run responses.

Key idea: need a "positive feedback" argument

- Start with $f(k, l)$ where inputs are substitutes ($f_{kl} \leq 0$)
- If capital is fixed, when wage increases, firm uses less labor in short-run and long-run
- Since two inputs are substitutes, increased wage implies an **increased** use of capital in the long-run
- Since $f_{kl} \leq 0$, additional capital used in long-run reduces marginal product of labor, so in the long-run the firm will use less labor

LR effect is larger than SR because in the SR firm only responds to a higher wage, but in LR, it responds to both higher wage and increased capital stock that reduces marginal product of labor – this is a positive feedback loop

Same intuition if complements

- Suppose two inputs are complements: $f_{kl} \geq 0$
- When wage increases, firm uses less labor in short-run and long run
- Since inputs are complements, increased wage implies **reduced** use of capital in long run
- Since $f_{kl} \geq 0$, reduced capital in long run will reduce marginal product of labor, so in the long-run firm uses less labor

Milgrom and Roberts (1996) develop positive feedback argument for two inputs. Suppose $X, Y = \mathbb{R}$. Let

$$x(y,t) = \arg\max_{x \in X} F(x,y,t)$$

$$y(t) = \arg\max_{y \in Y} F(x(y,t),y,t)$$

Proposition (Milgrom and Roberts)

Suppose $F : X \times Y \times \mathbb{R} \to \mathbb{R}$ *is supermodular,* $t' \geq t$*, and the maximizers described below are unique for parameter values* t *and* t'*. Then*

$$x(y(t'), t') \geq x(y(t), t') \geq x(y(t), t)$$

and

$$x(y(t'), t') \geq x(y(t'), t) \geq x(y(t), t)$$

Let x be labor input and y be capital. Let $t = -w_x$, where w_x is the price of labor. Firm's objective is to maximize:

$$F(x, y, t) = pf(x, y) - w_x x - w_y y$$

- If capital and labor are complements, then the firm's objective is supermodular in $(x, y, -w_x)$ since it satisfies all pairwise supermodularity conditions
- If capital and labor are substitutes then the firm's objective function is supermodular in $(x, -y, -w_x)$

Proposition (LeChatelier Principle)

If production is $f(k, l)$ and wage $w > w_0$ increases, then

a) *if k and l are complements, then*

$$l_{LR}(w) \le l_{SR}(w_0, w) \le l_{LR}(w_0)$$

b) *if k and l are substitutes, then*

$$l_{LR}(w) \le l_{SR}(w_0, w) \le l_{LR}(w_0)$$

Can't have sometimes complements, sometimes substitutes.

In example, capital is a complement to labor when labor is abdundant, but substitute when labor is scarce. Why?

- When we go from zero units of labor to one, marginal product of capital goes up from 0 to 10
- When we go from one unit of labor to 2, marginal product of capital goes back down from 10 to 0.

Monopoly Pricing

The profit-maximization problem in producer theory is usually unrealistic in that p is taken to be independent of q. The "monopoly" pricing problem is a more reasonable alternate formulation.

$$\max_q q \cdot P(q) - c(q)$$

Assume $\pi(q)$ is differentiable, concave, and has an interior solution.

The FOC for profit maximization is the familiar $MR = MC$:

$$q^m \cdot P'(q^m) + P(q^m) = c'(q^m)$$

Proposition

The percentage markup in the monopoly model is:

$$\frac{P(q^m) - c'(q^m)}{P(q^m)} = -\frac{1}{\epsilon},$$

where $\epsilon = \frac{dQ}{dP} \cdot \frac{P}{Q}$ *is the price elasticity of demand.*

$\frac{P(q^m)-c'(q^m)}{P(q^m)}$ also known as Lerner index

This result is sometimes known as the inverse elasticity rule.

Intuition: Monopolist is more wary of the perverse effect of a high price on demand when consumers react to a price increase by greatly reducing their demand.

How do we fit DWL into earlier thinking about welfare?

The loss to the consumer from the price increase from p^c to p^m is

$$CV = -\int_{p^c}^{p^m} x_i^h(p, u_0)dp$$

Note this depends on the Hicksian demand, not Marshallian.

Also, we need to trade off profits versus consumer utility - an even more problematic issue than cross consumer comparisons.

Here's the standard way to make sense of this. Let good 1 be the good with monopoly pricing.

Assumptions:

A1) Assume indirect utility has form:

$$
\begin{aligned}
v_i(p, w) &= \max_{\mathbf{q}, p\cdot\mathbf{q} \leq w} u_i(q_1, ..., q_L) \\
&= \max_{q_1} f_i(q_1) + [w - p_1 q_1]
\end{aligned}
$$

A2) $w_i = w_{0i} + s_i \pi(p)$

Recall that $(A1)$ let's us write $x(p)$ instead of $x(p, w)$ and lets us make comparisons across people.

$(A2)$ means the firms are owned by the consumers, so there are no "firm utilities" to worry about.

s_i is the fraction of the profits i obtains.

Proposition

Given any demand curve $P(q)$ and cost function $c(q,\theta)$, define

$$q^m = \sup \arg\max_q qp(q) - c(q,\theta).$$

If $c(q,\theta)$ has increasing differences, then q^m is weakly decreasing in θ.

Proposition (Monopolistic underprovision)

Let

$$q^{FB} = \sup \arg\max_q \int_0^q P(s)ds - c(q)$$

and

$$q^m = \sup \arg\max_q qP(q) - c(q).$$

If $P'(q) \leq 0$ *then* $q^m \leq q^{FB}$*.*

Monopoly and Product Quality

DWL is not the only problem with monopoly people talk about. For example, common complaints about Windows are not that it is so expensive, but that it does not work well.

Suppose monopolist chooses quality s as well as quantity.

$$W(q,s) = \int_0^q P(x,s)dx - c(q,s)$$
$$\pi(q,s) = qP(q,s) - c(q,s)$$

One comparative static theorem is

Proposition

Let $s^{FB} = \sup\arg\max_s W(q,s), s^m(q) = \sup\arg\max_s \pi(q,s)$*. If* $P(q,s)$ *has decreasing differences, then* $s^{FB}(q) \geq s^m(q)$*.*

Intuition: If $P(q,s)$ has decreasing differences, then the value of quality is less on the marginal unit sold then on the average unit.

Monopolist only cares about price of the marginal unit. Social planner cares about (value of quality) on average unit. Known as **Spence distortion**

Note: this theorem gives one reason why monopoly quality too low. IO literature notes a force that goes in the opposite direction: monopolist sells fewer units $\Rightarrow$ monopolist marginal consumer has a **higher value**, than social planner's marginal consumer.

In richer models, if one considers choices **not holding q fixed**, this can offset the first effect and give $s^m > s^{FB}$.