Theory of the Firm

Based on Alexander Wolitzky (MIT – Microeconomics 1) and Leonardo Felli (LSE – Advanced Microeconomics)

Neoclassical Producer Theory in One Sentence

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

We expand on this just slightly, and show how main results of producer theory follow from results from consumer theory.

The Profit Maximization Problem (PMP)

Choose production plan $y \in \mathbb{R}^n$ from production possibilities set $Y \subseteq \mathbb{R}^n$ to maximize profit $p \cdot y$:

$$\max_{y \in Y} p \cdot y$$

- Some prices can be negative.

Lets us model inputs and outpus symmetrically.

- Inputs have negative prices (firm pays to use them).
- Outputs have positive prices (firm makes money by producing).
- Neoclassical firm is price taker.
- No market power.
- Study of firms with market power is a topic in industrial organization.
- Firm's objective is profit maximization.
- In reality, firm is organization composed of individuals with different goals.
- Study of internal behavior and organization of firms is topic in organizational economics.

The PMP and the EMP

For our purposes, producer theory leaves everything interesting about firm behavior to other areas of economics, and reduces firm's problem to something isomorphic to consumer's expenditure minimization problem.

PMP is

$$\max_{y \in Y} \; p \cdot y.$$

Letting $S = \{x \in \mathbb{R}^n : u(x) \geq u\}$, EMP is

$$\min_{x \in S} p \cdot x.$$

Up to flipping a sign, PMP the same as EMP.

EMP: consumer chooses bundle of goods x to minimize expenditure, subject to x lying in set S.

PMP: firm chooses bundle of goods y to minimize net expenditure (maximize net profit), subject to y lying in set Y.

The PMP and the EMP

Solution to EMP: Hicksian demand $h(p)$.

Value function for EMP: expenditure function $e(p)$.

(omitting u because we hold it fixed)

Solution to PMP: optimal production plan $y(p)$.

Value function for EMP: profit function $\pi(p)$.

Our treatment of producer theory consists of recalling facts about Hicksian demand and expenditure function, and translating into language of optimal production plan and profit function.

Properties of Hicksian Demand/Optimal Production Plans

Theorem

Hicksian demand satisfies:

1. Homogeneity of degree 0: for all $\lambda > 0, h(\lambda p) = h(p)$.
2. Convexity: if S is convex (i.e., if preferences are convex), then $h(p)$ is a convex set.
3. Law of demand: for every $p, p' \in \mathbb{R}^n, x \in h(p)$, and $x' \in h(p')$, we have $(p' - p)(x' - x) \leq 0$.

Theorem

Optimal production plans satisfy:

1. Homogeneity of degree 0 : for all $\lambda > 0, y(\lambda p) = y(p)$.
2. Convexity: if Y is convex, then $y(p)$ is a convex set.
3. Law of supply: for every $p, p' \in \mathbb{R}^n, y \in y(p)$, and $y' \in y(p')$, we have $(p' - p)(y' - y) \geq 0$.

Properties of Expenditure Function/Profit Function

Theorem

The expenditure function satisfies:

1. Homogeneity of degree 1: for all $\lambda > 0, e(\lambda p) = \lambda e(p)$.
2. Monotonicity: e is non-decreasing in p.
3. Concavity: e is concave in p.
4. Shephard's lemma: under mild conditions (see Lectures 2-3), e is differentiable, and $\frac{\partial}{\partial p_i} e(p) = h_i(p)$.

Theorem

The profit function satisfies:

1. Homogeneity of degree 1: for all $\lambda > 0, \pi(\lambda p) = \lambda \pi(p)$.
2. Monotonicity: π is non-decreasing in p.
3. Convexity: π is convex in p.
4. Hotelling's lemma: under mild conditions, π is differentiable, and $\frac{\partial}{\partial p_i} \pi(p) = y_i(p)$.

Production Plan

- Both outputs and inputs (measured in terms of flow) are services and commodities.
- y_j^o = quantity of commodity j produced by the firm as output,
- y_j^i = quantity of commodity j used as input,
- $z_j = y_j^o - y_j^i$ net output/input depending on whether the sign of z_j is positive/negative.
- Production plan = vector of net outputs and/or inputs of all available commodities

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_L \end{pmatrix}$$

Net Inputs and Outputs

- Without loss of generality we assume that:
- the first h commodities are net inputs
- while the remaining $L - h$ commodities are net outputs.
- Define:

$$x_1 = -z_1, \dots, x_h = -z_h, y_1 = z_{h+1}, \dots, y_{L-h} = z_L$$

Production Plan and Possibility Set

- A production plan is:

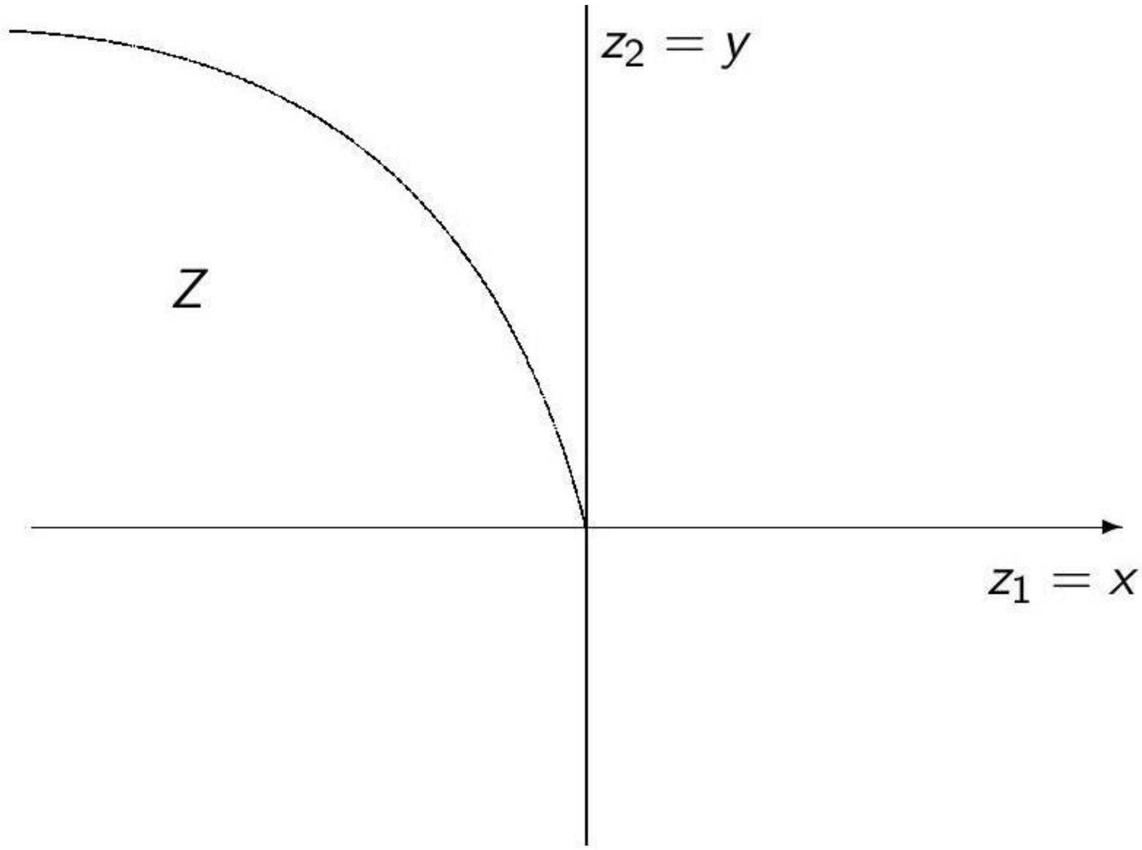
$$z = \begin{pmatrix} -x_1 \\ \vdots \\ -x_h \\ y_1 \\ \vdots \\ y_{L-h} \end{pmatrix}$$

Definition

The Production Possibility Set $Z \subset \mathbb{R}^L$ (PPS) is the set of all technologically feasible production plans: the set of all vectors of inputs and outputs that are technologically feasible. The PPS Z provides a complete description of the technology identified with the firm.

One Input One Output

Example of one input x and one output y production plan $z = \binom{-x}{y}$



Short-run PPS

- Sometime is possible to distinguish between:
- immediately technologically feasible production plans $Z(\bar{x}_1, \dots, \bar{x}_h)$;
- and eventually technologically feasible production plans Z.
- Consider

$$Z(\bar{x}_1) = \left\{ z = \begin{pmatrix} -x_1 \\ -x_2 \\ y \end{pmatrix} \mid x_1 = \bar{x}_1 \right\}$$

- For example if the input x_1 is fixed at the level $\bar{x}_1$ then we can define a short-run or restricted production possibility set.

Input Requirement Set and Isoquant

- A special feature of a technology is the input requirement set:

$$V(y) = \left\{ x \in \mathbb{R}_+^h \mid \begin{pmatrix} -x \\ y \end{pmatrix} \in Z \right\}$$

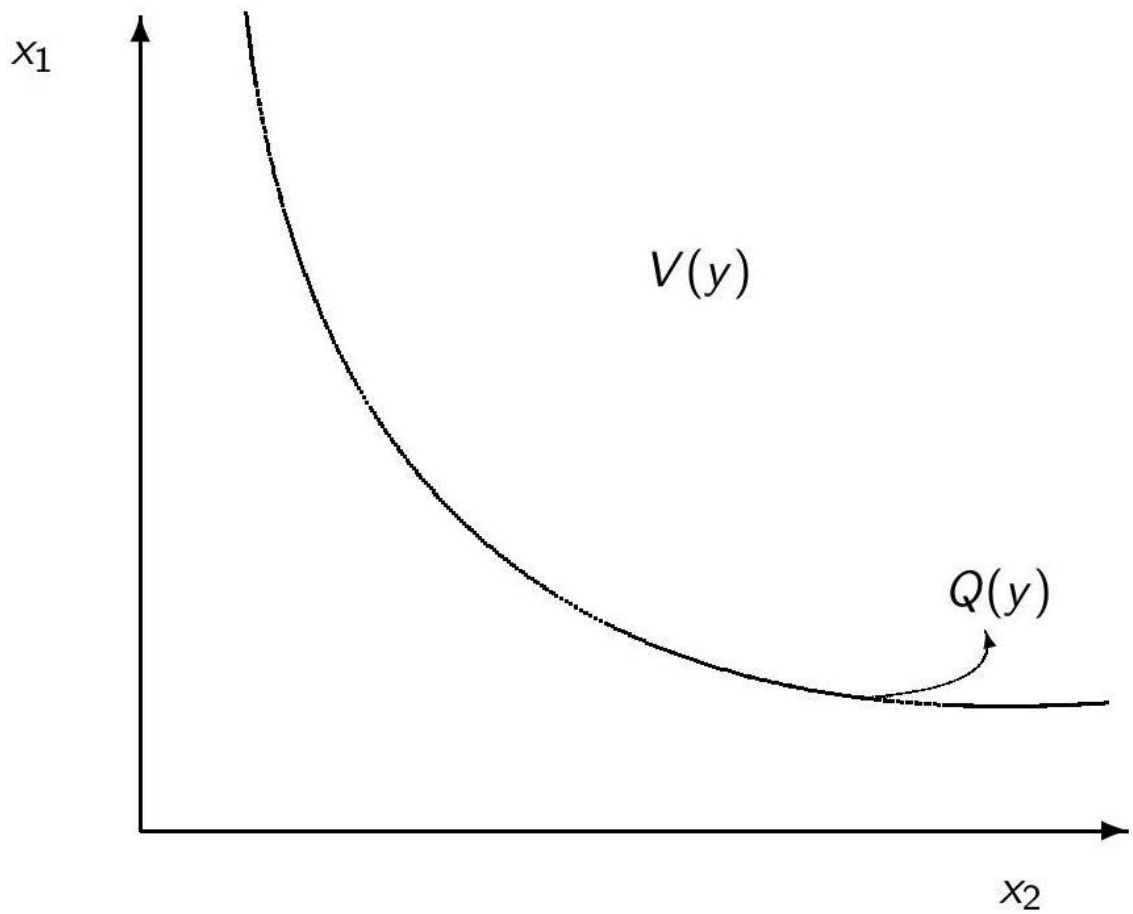
the set of all input bundles that produce at least y units of output.

- We define also the isoquant to be the set:

$$Q(y) = \{ x \in \mathbb{R}_+^h \mid x \in V(y) \text{ and } x \notin V(y'), \forall y' > y \}$$

all input bundles that allow the firm to produce exactly y.

Two Inputs One Output



Production Function

Consider a technology with only one output

Definition

Production function in the case of only one output:

$$f(x) = \sup_{y'} \left\{ \begin{pmatrix} -x \\ y' \end{pmatrix} \in Z \right\}$$

the maximal output associated with the input bundle x.

Technologically Efficient

In general, we can define the technologically efficient production plan z as:

Definition

The general production plan $z = \begin{pmatrix} -x \\ y \end{pmatrix} \in Z$ is technologically efficient, if and only if there does not exist a production plan $z' = \begin{pmatrix} -x' \\ y' \end{pmatrix} \in Z$ such that $z' \geq z (z'_i \geq z_i \forall i)$ and $z' \neq z$.

If z is efficient it is not possible to produce more output with a given input or the same output with less input (sign convention).

Example: Cobb-Douglas Technology

- We definite the Cobb-Douglas technology as

$$f(x_1, x_2) = x_1^\alpha x_2^\beta,\ \alpha > 0, \beta > 0$$

or

$$Z = \left\{ \begin{pmatrix} -x_1 \\ -x_2 \\ y \end{pmatrix} \mid y \leq x_1^\alpha x_2^\beta \right\}$$

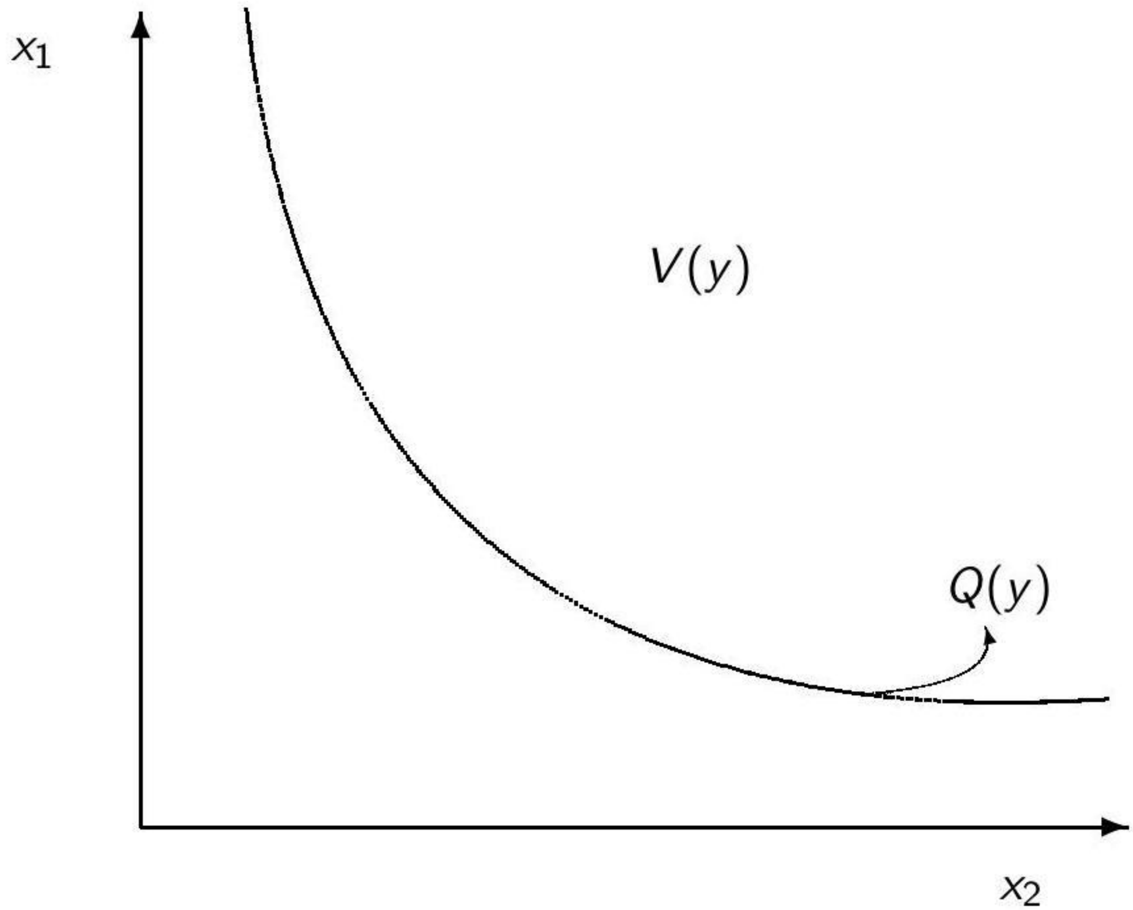
- with isoquant:

$$Q(y) = \left\{ (x_1, x_2) \in \mathbb{R}_+^2 \mid y = x_1^\alpha x_2^\beta \right\}$$

- and input requirement set:

$$V(y) = \left\{ (x_1, x_2) \in \mathbb{R}_+^2 \mid y \leq x_1^\alpha x_2^\beta \right\}$$

Cobb-Douglas Isoquant



Example: Leontief Technology

- We definite the Leontief technology as

$$f(x_1, x_2) = \min\{ax_1, bx_2\},\ a > 0, b > 0$$

or

$$Z = \left\{ \begin{pmatrix} -x_1 \\ -x_2 \\ y \end{pmatrix} \mid y \leq \min\{ax_1, bx_2\} \right\}$$

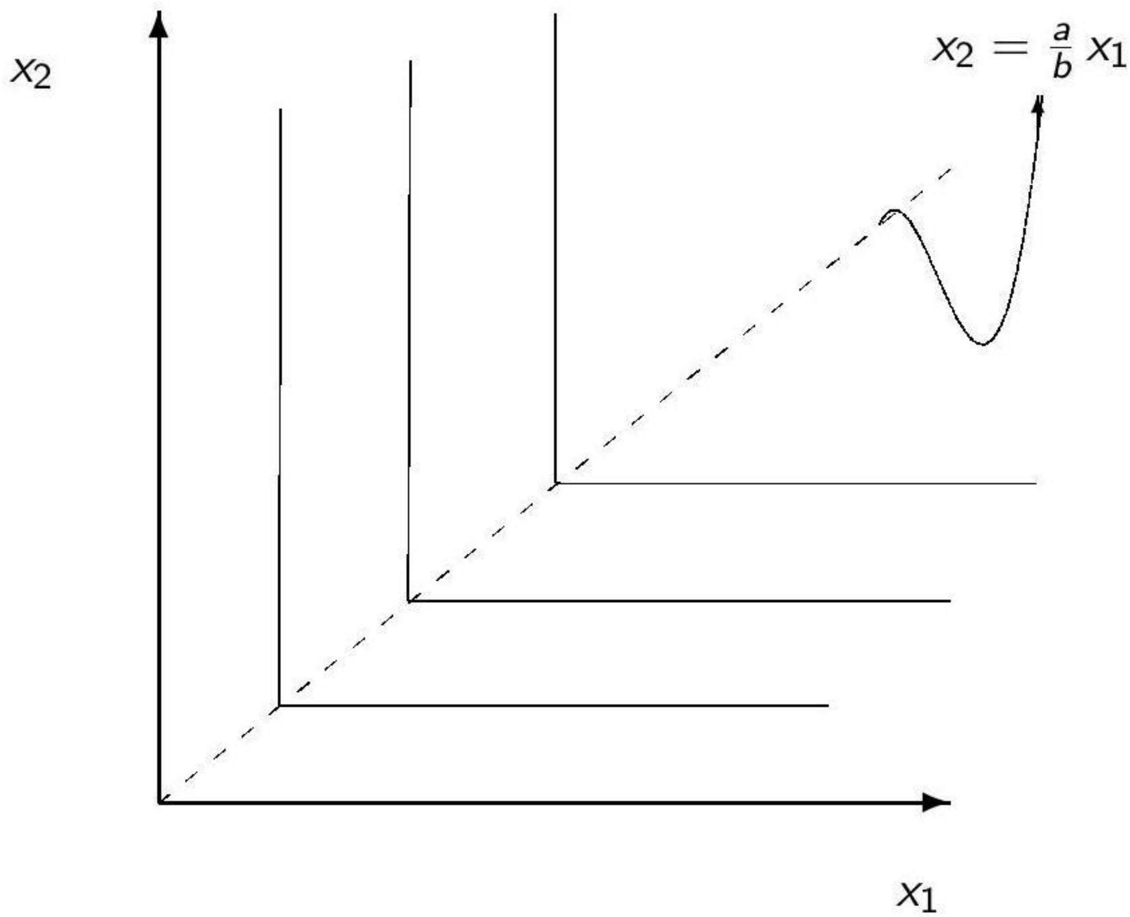
- with isoquant:

$$Q(y) = \{(x_1, x_2) \in \mathbb{R}_+^2 \mid y = \min\{ax_1, bx_2\}\}$$

- and input requirement set:

$$V(y) = \{(x_1, x_2) \in \mathbb{R}_+^2 \mid y \leq \min\{ax_1, bx_2\}\}$$

- where efficiency imposes $x_1 = \frac{y}{a}$, $x_2 = \frac{y}{b}$ Leontief Isoquants



Example: Perfect Substitutes

- We definite the technology where inputs are perfect substitutes as

$$f(x_1, x_2) = ax_1 + bx_2,\ a > 0, b > 0$$

or

$$Z = \left\{ \begin{pmatrix} -x_1 \\ -x_2 \\ y \end{pmatrix} \mid y \leq ax_1 + bx_2 \right\}$$

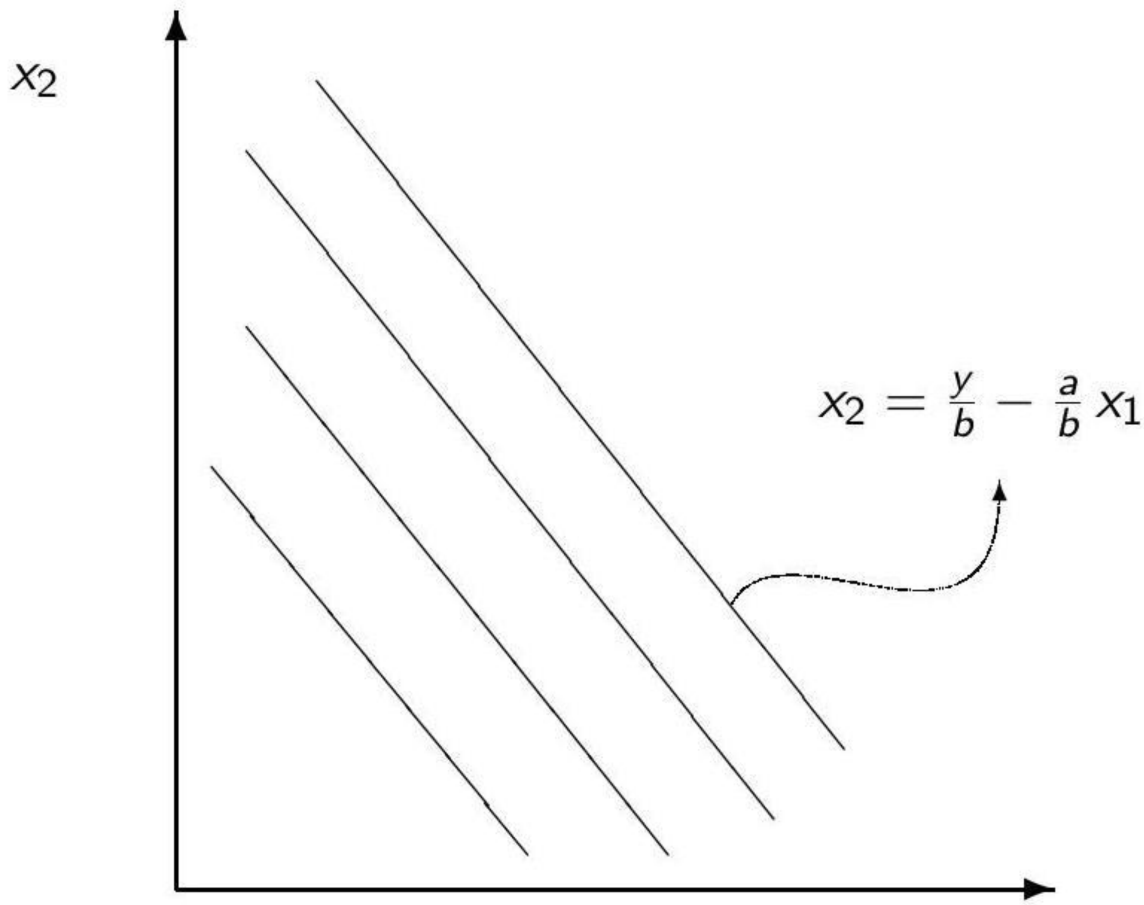
- with isoquant:

$$Q(y) = \{(x_1, x_2) \in \mathbb{R}_+^2 \mid y = ax_1 + bx_2\}$$

- and input requirement set:

$$V(y) = \{(x_1, x_2) \in \mathbb{R}_+^2 \mid y \leq ax_1 + bx_2\}$$

Perfect Substitutes Isoquants



x_1

Assumptions on PPS

Common assumptions on the PPS are:

1. Z is closed (it contains its boundaries). It is an important property for the definition of a production function:

sup is a max.

2. $0 \in Z$. Shut-down property. It is an uncontroversial property in the long run, not necessarily in the short run (inputs used with no outputs).

3. Free disposal, monotonicity: if $z \in Z$ and $z' \leq z$ then $z' \in Z$. Alternatively: if $x \in V(y)$ and $x' \geq x$ then $x' \in V(y)$.

Given a feasible production plan if either one increases the quantity of inputs or reduces the quantity of output the new production plan is still feasible.

4. Additivity: if $z, z' \in Z$ then $z + z' \in Z$ (stronger condition). For $f(x)$ this property implies $f(x^1 + x^2) \geq f(x^1) + f(x^2)$.

5. Convexity of $V(y)$: if $x, x' \in V(y)$ then $tx + (1-t)x' \in V(y)$ for every $0 \leq t \leq 1$ which means that $V(y)$ is convex set. (rescaling of production processes).

5'. A similar condition may (or may not) be imposed on the Z : if $z, z' \in Z$ then $tz + (1-t)z' \in Z$ for every $0 \leq t \leq 1$, or Z is a convex set.

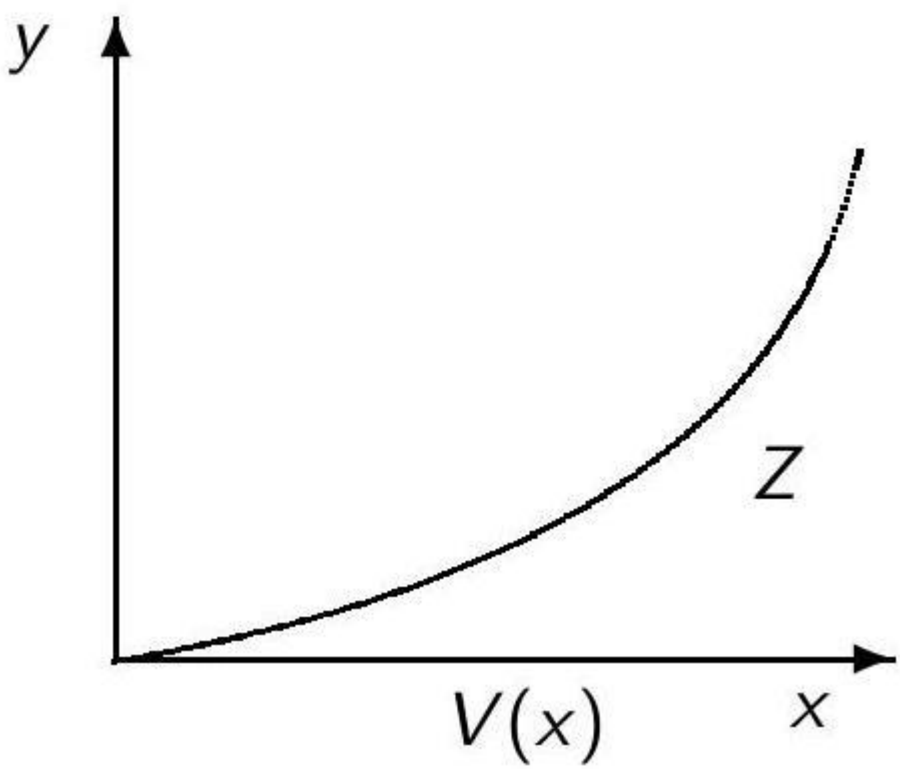
Notice that the latter condition is stronger than the former.

Some Results

Result

The convexity of Z implies the convexity of $V(y)$. The opposite implication does not hold.

Proof: It follows from the convexity of Z, the definition of $V(y)$ and the following counter-example of a one-input x and one output y technology.



Result

The convexity of $V(y)$ implies that the $f(x)$ is quasi-concave.

Proof: It follows from the convexity of $V(y)$ and the definition of a quasi-concave $f(x)$.

Definition

The function $f(\cdot)$ is quasi-concave if and only if the set $\{x \mid f(x) \geq k\}$ is convex for every $k \in \mathbb{R}$.

Notice that if you choose $y = k$ the set above is $V(y)$.

Result

The convexity of Z implies that $f(x)$ is (weakly) concave.

Proof: Consider

$$z = \begin{pmatrix} -x \\ f(x) \end{pmatrix} \in Z, \; z' = \begin{pmatrix} -x' \\ f(x') \end{pmatrix} \in Z$$

Convexity of Z implies that for every $0 \leq t \leq 1$

$$tz + (1-t)z' = \begin{pmatrix} -(tx + (1-t)x') \\ tf(x) + (1-t)f(x') \end{pmatrix} \in Z$$

By definition of $f(x)$ this means:

$$tf(x) + (1-t)f(x') \leq f(tx + (1-t)x')$$

for every $0 \leq t \leq 1$. This is the definition of a concave $f(x)$.

Returns to Scale

- Decreasing Returns to Scale: (DRS) if $z \in Z$ then $tz \in Z$ for every $0 \leq t \leq 1$ (graph.).
- Increasing Returns to Scale: (IRS) if $z \in Z$ then $tz \in Z$ for every $t \geq 1$ (graph.).
- Constant Returns to Scale: (CRS) if $z \in Z$ then $tz \in Z$ for every $t \geq 0$ (graph.).

More Results

Result

Assumptions $0 \in Z$ and Z convex imply DRS.

Proof: It follows from the definition of convexity applied at $z' = 0$.

Result

A technology exhibits CRS if and only if the production function $f(x)$ (if available) is homogeneous of degree 1.

Proof: Assume CRS. This implies that if $z \in Z$ then $tz \in Z$, for all $t \geq 0$.

By definition of $Z, z \in Z$ means

$$y \leq f(x)$$

further $tz \in Z$ means

$$ty \leq f(tx).$$

By definition of production function choose z, and hence x and y, so that $y = f(x)$.

We can re-write the latter condition as:

$$tf(x) \leq f(tx)$$

We need to prove that the equality holds.

Suppose it does not. Then there exists y' such that

$$tf(x) < y' < f(tx)$$

Now $y' < f(tx)$ implies by definition of Z that

$$\begin{pmatrix} -tx \\ y' \end{pmatrix} \in Z$$

By CRS we get

$$\frac{1}{t}\begin{pmatrix} -tx \\ y' \end{pmatrix} \in Z \text{ or } \begin{pmatrix} -x \\ \frac{1}{t}y' \end{pmatrix} \in Z$$

which means

$$(1/t)y' \leq f(x)$$

or

$$y' \leq tf(x)$$

the latter inequality contradicts $tf(x) < y'$.

The opposite implication is an immediate consequence of the definition of homogeneity of degree 1 .

Conditions for DRS and IRS

Weaker conditions apply for DRS and IRS technology.

Result

Consider a technology characterized by a homogenous of degree $\alpha < 1 (\alpha > 1)$ production function. This technology exhibits DRS (IRS). The opposite implication does not hold.

Result

Assume that $f(0) = 0$ then we can prove:

- $f(x)$ concave implies DRS;
- $f(x)$ convex implies IRS;
- $f(x)$ concave and convex implies CRS.

Some Definitions

We can now introduce few definitions:

Definition

The marginal product of input x_i is

$$MP = \frac{\partial f(x)}{\partial x_i}$$

Definition

The average product of input x_i is

$$AP = \frac{f(x)}{x_i}$$

Definition

The marginal rate of technical substitution between input x_i and x_j is

$$\left|\frac{dx_i}{dx_j}\right| = \frac{\partial f(x)/\,\partial x_j}{\partial f(x)/\,\partial x_i}$$

this is the absolute value of the slope of the isoquant.

The set of output bundles that are efficient for a given technology:

Definition

The Production Possibility Frontier:

$$\text{PPF}\,(x) = \left\{y \mid \nexists z' \in Z \text{ s.t. } z' \geq z = \begin{pmatrix} -x \\ y \end{pmatrix}\right\}$$

Definition

The Marginal Rate of Transformation between output y_m and y_n is

$$MRT = \frac{dy_m}{dy_n}$$

as the slope of the PPF.

The Competitive Firm

- Assume that input and output prices are taken parametrically (no influence on such prices).
- As you have seen in consumer theory what this means is that whatever each firm decides in term of production does not affect the market.
- In other words, either firms are very small with respect to the market.
- Alternatively we are assuming that firms are not strategic: they do not realize that their choices trigger reactions in other firms in the market,
- or any of their potential choices would be taken into account by competitors when making their own choices.

Profit Maximization

The basic producer problem is than profit maximization:

$$\max_{\{x,y\}} py - \sum_{i=1}^{h} w_i x_i$$
$$\text{s.t.} \ \begin{pmatrix} -x \\ y \end{pmatrix} \in Z$$

where p and w_i are taken as parameters.

Let:

- the h-dimensional vector of input prices be $w = (w_1, \dots, w_h)$;
- the L - h-dimensional vector of output prices be $p = (p_1, \dots, p_{L-h})$.

We can re-write the producer's problem as:

$$\max_{\{z\}} \quad \hat{p}z$$
$$\text{s.t.} \quad z \in Z$$

where $\hat{p} = (w, p)$ and $z = \begin{pmatrix} -x \\ y \end{pmatrix}$.

This is what we’ve seen before, with a slightly different notation: z instead of y.

Profit Maximization with only one output: differentiable case

In the case of a technology that produces only one output the profit maximization problem may be written as:

$$\max_{\{x,y\}} pf(x) - wx$$

The necessary first order conditions of this problem are:

$$p\nabla f(x^*) \leq w$$

or

$$\frac{\partial f(x^*)}{\partial x_i} \leq \frac{w_i}{p}, \; \forall i = 1, \dots, h$$

and

$$\left[\frac{\partial f(x^*)}{\partial x_i} - \frac{w_i}{p}\right] x_i^* = 0, \; \forall i = 1, \dots, h.$$

Profit Maximization: Second-order conditions

In the event that the production possibility set is convex (the production function is concave) the first order conditions are both necessary and sufficient.

- In other case, the following set of sufficient conditions for a local maximum have to be verified.
- The Hessian matrix of the production function has to be negative definite at the point x^*.

This condition can be checked by the sufficient determinant condition according to which the leading principal minors have to alternate sign starting from the negative one.

For the case of two variables the first order conditions are for $i = 1,2$:

$$\frac{\partial f(x^*)}{\partial x_i} \leq \frac{w_i}{p}$$

and

$$\left[\frac{\partial f(x^*)}{\partial x_i} - \frac{w_i}{p}\right] x_i^* = 0$$

while the second order conditions are:

$$H = \begin{pmatrix} \frac{\partial^2 f(x^*)}{\partial x_1^2} & \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x^*)}{\partial x_2^2} \end{pmatrix} \text{ negative definite}$$

Which is implied by:

$$\frac{\partial^2 f(x^*)}{\partial x_i^2} < 0$$

and

$$\frac{\partial^2 f(x^*)}{\partial x_1^2}\frac{\partial^2 f(x^*)}{\partial x_2^2} - \left(\frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2}\right)^2 > 0$$

Unconditional Factor Demand and Supply Function

The solution to the profit maximization problem if it exists provides the unconditional factor demands:

$$x(p, w) = x^*$$

By substitution it is possible to obtain the supply function of the producer:

$$y(p, w) = f(x(p, w)).$$

Comparative statics results obtained by differentiating the FOC (they are identities in (p, w)).

Properties of Factor Demands

1. Non-positive own factor demands price effects (SOC) (generalizes to h inputs):

$$\frac{\partial x_1}{\partial w_1} \leq 0 \; \frac{\partial x_2}{\partial w_2} \leq 0$$

2. Symmetry (generalizes to h inputs):

$$\frac{\partial x_1}{\partial w_2} = \frac{\partial x_2}{\partial w_1}$$

3. Complementary inputs (generalizes to h inputs):

$$\frac{\partial x_1}{\partial w_2} = \frac{\partial x_2}{\partial w_1} < 0$$

4. Substitutability of inputs (it does not generalize to several inputs):

$$\frac{\partial x_1}{\partial w_2} = \frac{\partial x_2}{\partial w_1} > 0$$

5. Finally positive output price effects (generalizes to h inputs):

$$\frac{\partial x_1}{\partial p} > 0 \; \frac{\partial x_2}{\partial p} > 0$$

(If x_1 and x_2 are complementary inputs.)

Some comparative static results obtained differentiating the supply function of the firm:

$$y(p, w) = f(x(p, w))$$

6. Own price effect non-negative:

$$\frac{\partial y}{\partial p} \geq 0$$

7. Symmetry:

$$-\frac{\partial x_i}{\partial p} = \frac{\partial y}{\partial w_i}$$

for $i = 1,2$.

Summary of the Properties

Such comparative statics properties can be summarized as:

$$\begin{pmatrix} \frac{\partial y}{\partial p} & \frac{\partial y}{\partial w_1} & \frac{\partial y}{\partial w_2} \\ -\frac{\partial x_1}{\partial p} & -\frac{\partial x_1}{\partial w_1} & -\frac{\partial x_1}{\partial w_2} \\ -\frac{\partial x_2}{\partial p} & -\frac{\partial x_2}{\partial w_1} & -\frac{\partial x_2}{\partial w_2} \end{pmatrix} \text{ s.t. } \begin{pmatrix} + & a & b \\ a & + & c \\ b & c & + \end{pmatrix}$$

Other Properties

8. Both $x(p, w)$ and $y(p, w)$ are homogeneous of degree 0 .

Proof: If you increase both input and output prices by a factor $t > 0$ you obtain:

$$\max_x (tp)f(x) - (tw)x = \max_x t[pf(x) - wx]$$

which clearly is solved by the same vector $x(p, w)$ that solves:

$$\max_x pf(x) - wx$$

Further, by definition of supply function:

$$y(tp, tw) = f(x(tp, tw)) = f(x(p, w)) = y(p, w) \quad \square$$

Profit Function

Definition

The following is defined as the profit function

$$\pi(p, w) = \max_x pf(x) - wx = pf(x(p, w)) - wx(p, w)$$

Properties:

1. Price effects:

$$\frac{\partial \pi}{\partial w_i} \leq 0 \; \frac{\partial \pi}{\partial p} \geq 0$$

2. The profit function $\pi(p, w)$ is homogeneous of degree 1 in (p, w).

Proof: It follows from the homogeneity of degree 0 of $x(p, w)$ and, for any scalar α,

$$\begin{aligned}
\pi(\alpha p, \alpha w) \quad &= \alpha p f(x(\alpha p, \alpha w)) - \alpha w x(\alpha p, \alpha w) \\
&= \alpha p f(x(p, w)) - \alpha w x(p, w) \\
&= \alpha[pf(x(p, w)) - wx(p, w)] \\
&= \alpha \pi(p, w)
\end{aligned}$$

3. Hotelling Lemma (which proves property 1):

$$\frac{\partial \pi}{\partial p} = y(p, w) \geq 0 \;\text{ and }\; \frac{\partial \pi}{\partial w_i} = -x_i(p, w) \leq 0$$

Proof: It follows by Envelope Theorem applied to

$$\pi(p, w) = \max_x pf(x) - wx = pf(x(p, w)) - wx(p, w)$$

4. The profit function $\pi(p, w)$ is convex in (p, w).

Proof: Consider the two price vectors (p, w) and (p', w') and for every scalar $\lambda \in (0,1)$ let

$$p'' = \lambda p + (1 - \lambda)p'$$

and

$$w'' = \lambda w + (1 - \lambda)w'$$

- Then:

$$
\begin{aligned}
\pi[p'', w''] = \quad & p'' f\big(x(p'', w'')\big) - w'' x(p'', w'') \\
= \quad & \lambda\big[p f\big(x(p'', w'')\big) - w x(p'', w'')\big] \\
& \quad + (1 - \lambda)\big[p' f\big(x(p'', w'')\big) - w' x(p'', w'')\big] \\
\leq \quad & \lambda \pi(p, w) + (1 - \lambda)\pi(p', w')
\end{aligned}
$$

which proves convexity of $\pi(p, w)$.

Monotone Comparative Statics: Motivation

Comparative statics are statements about how solution to a problem changes with parameters.

Core of most applied economic analysis.

Last twenty years or so:

revolution in how comparative statics are done in economics.

Traditional approach: differentiate FOC using implicit function theorem.

New approach:

monotone comparative statics.

Example: Traditional Approach

Consider problem:

$$\max_{x \in X} b(x, \theta) - c(x)$$

- x is choice variable
- θ is parameter
- $b(x, \theta)$ is benefit from choosing x given parameter θ
- c is cost of choosing x

Example: Traditional Approach

$$\max_{x \in X} b(x, \theta) - c(x)$$

If $X \subseteq \mathbb{R}$ and b and c are differentiable, FOC is

$$b_x(x^*(\theta), \theta) = c'(x^*(\theta)).$$

If b and c are twice continuously differentiable and $b_{xx}(x^*(\theta), \theta) \neq c''(x^*(\theta))$, implicit function theorem implies that solution $x^*(\theta)$ is continuously differentiable, with derivative

$$\frac{d}{d\theta} x^*(\theta) = \frac{b_{x\theta}(x^*(\theta), \theta)}{c''(x^*(\theta)) - b_{xx}(x^*(\theta), \theta)}.$$

If c is convex, b is concave in x, and $b_{x\theta} > 0$, can conclude that $x^*(\theta)$ is (locally) increasing in θ. Intuition: FOC sets marginal benefit equal to marginal cost. If $b_{x\theta} > 0$ and θ increases, then if b is concave in x and c is convex, x must increase to keep the FOC satisfied.

What's Wrong with This Picture?

Unnecessary assumptions: as we'll see, solution(s) are increasing in θ even if b is not concave, c is not convex, b and c are not differentiable, and choice variable is not continuous or real-valued.

Wrong intuition: Intuition coming from the FOC involves concavity of b and convexity of c.

This can't be the right intuition.

We'll see that what's really needed is an ordinal condition on b-the single-crossing property-which is a more meaningful version of the assumption $b_{x\theta} > 0$.

Why Learn Monotone Comparative Statics?

Three reasons:

1. Generality: Cut unnecessary convexity and differentiability assumptions.

2. Analytical power: Often, can't assume convexity and differentiability.

(Traditional approach doesn't work.)

3. Understanding: By focusing on essential assumptions, help to understand workings of economic models.

(Don't get confused about what drives what.)

Fourth reason: need to understand them to read other people's 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, P. and J. Eeckhout. "Sorting and Decentralized Price Competition." Econometrica, 2010. [Labor]

- Segal, I. and M. Whinston. "Property Rights." Chapter for Handbook of Organizational Economics, 2011.

[Organizational Econ]

- Acemoglu, D. and A. Wolitzky. "The Economics of Labor Coercion." Econometrica, 2011. [Political Economy]

MCS with 1 Choice Variable and 1 Parameter

Start with simple case: $X \subseteq \mathbb{R}, \Theta \subseteq \mathbb{R}$.

Interested in set of solutions $X^*(\theta)$ to optimization problem

$$\max_{x \in X} f(x, \theta)$$

Under what conditions on f is $X^*(\theta)$ increasing in θ ?

The Strong Set Order

What does it mean for set of solutions to be increasing?

Relevant order on sets: strong set order.

Definition

A set $A \subseteq \mathbb{R}$ is greater than a set $B \subseteq \mathbb{R}$ in the strong set order (SSO) if, for any $a \in A$ and $b \in B$,

$$\begin{aligned} \max\{a, b\} &\in A, \text{ and} \\ \min\{a, b\} &\in B. \end{aligned}$$

$X^*(\theta)$ greater than $X^*(\theta')$ if, whenever x is solution at θ and x' is solution at θ', either

1. $x \geq x'$, or
2. both x and x' are solutions for both parameters.

Increasing Differences

Simple condition on f that guarantees that $X^*(\theta)$ is increasing (in SSO) : increasing differences.

Definition

A function $f: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has increasing differences in (x, θ) if, whenever $x^H \geq x^L$ and $\theta^H \geq \theta^L$, we have

$$f(x^H, \theta^H) - f(x^L, \theta^H) \geq f(x^H, \theta^L) - f(x^L, \theta^L).$$

Return to choosing a higher value of x is non-decreasing in θ.

Form of complementarity between x and θ.

Increasing Differences: Differential Version

Theorem

If f is twice continuously differentiable, then f has increasing differences in (x, θ) iff

$$\frac{\partial^2 f(x,\theta)}{\partial x \, \partial \theta} \geq 0 \text{ for all } x \in X, \theta \in \Theta.$$

Increasing differences generalizes condition on cross-partial derivatives used to sign comparative statics in traditional approach.

Topkis' Monotonicity Theorem

Simplest MCS theorem:

Theorem (Topkis)

If f has increasing differences in (x, θ), then $X^*(\theta)$ is increasing in the strong set order.

Back to Example

$$\max_{x \in X} b(x,\theta) - c(x)$$

If b has increasing differences in (x, θ), then $X^*(\theta)$ is increasing in the strong set order.

No assumptions about convexity or differentiability of anything.

Necessity

Want to find minimal assumptions for given comparative statics result to hold.

Is increasing differences minimal assumption?

No: increasing differences is cardinal property, but property that $X^*(\theta)$ is increasing is ordinal.

What's ordinal version of increasing differences?

Single-Crossing

Definition

A function $f: X \times \Theta \to \mathbb{R}$ is single-crossing in (x, θ) if, whenever $x^H \geq x^L$ and $\theta^H \geq \theta^L$, we have

$$f(x^H, \theta^L) \geq f(x^L, \theta^L) \Longrightarrow f(x^H, \theta^H) \geq f(x^L, \theta^H)$$

and

$$f(x^H, \theta^L) > f(x^L, \theta^L) \Rightarrow f(x^H, \theta^H) > f(x^L, \theta^H).$$

Whenever choosing a higher x is better at a low value of θ, it's also better at a high value of θ.

Increasing differences implies single-crossing, but not vice versa.

Milgrom-Shannon Monotonicity Theorem

Theorem (Milgrom and Shannon)

If f is single-crossing in (x, θ), then $X^*(\theta)$ is increasing in the strong set order.

Conversely, if $X^*(\theta)$ is increasing in the strong set order for every choice set $X \subseteq \mathbb{R}$, then f is single-crossing in (x, θ).

Strictly Increasing Selections

A stronger set order: for $\theta < \theta'$, every $x \in X^*(\theta)$ is strictly less than every $x' \in X^*(\theta')$.

(Every selection is strictly increasing.)

When is every selection strictly increasing?

Strictly increasing differences: whenever $x^H > x^L$ and $\theta^H > \theta^L$, we have

$$f(x^H, \theta^H) - f(x^L, \theta^H) > f(x^H, \theta^L) - f(x^L, \theta^L).$$

Theorem (Edlin and Shannon)

Suppose f is continuously differentiable in x and has strictly increasing differences in (x, θ).

Then, for all $\theta < \theta', x^* \in X^*(\theta) \cap \text{int } X$, and $x^{*\prime} \in X^*(\theta')$, we have $x^* < x^{*\prime}$.

MCS with n Choice Variables and m Parameters

Previous theorems generalize to $X \subseteq \mathbb{R}^n$ and $\Theta \subseteq \mathbb{R}^m$.

Two main issues in generalization:

1. What's "max" or "min" of two vectors?
2. Need complementarity within components of x, not just between x and θ.

Once clear these up, analysis same as in 1-dimensional case.

Meet and Join

Relevant notion of min and max are component-wise min and max, also called meet and join:

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

Definition

A set $A \subseteq \mathbb{R}^n$ is greater than a set $B \subseteq \mathbb{R}^n$ in the strong set order if, for any $a \in A$ and $b \in B$,

$$a \vee b \in A, \text{ and}$$
$$a \wedge b \in B.$$

A lattice is a set $X \subseteq \mathbb{R}^n$ such that $x \wedge y \in X$ and $x \vee y \in X$ for all $x, y \in X$.

Ex. A product set $X = X_1 \times \dots X_n$ is a lattice.

Increasing Differences

Definition of increasing differences in (x, θ) same as before: $x^H \geq x^L, \theta^H \geq \theta^L \Longrightarrow$

$$f(x^H, \theta^H) - f(x^L, \theta^H) \geq f(x^H, \theta^L) - f(x^L, \theta^L)$$

(Note: x and θ are vectors. What does $x^H \geq x^L$ mean?)

Increasing differences in (x, θ) no longer enough to guarantee $X^*(\theta)$ increasing.

Issue: what if increase in θ_1 pushes x_1 and x_2 up, but increase in x_1 pushes x_2 down?

Need complementarity within components of x, not just between x and θ.

27 This is called supermodularity of f in x.

Supermodularity

Definition

A function $f: X \times \Theta \to \mathbb{R}$ is supermodular in x if, for all $x, y \in X$ and $\theta \in \Theta$, we have

$$f(x \vee y, \theta) - f(x, \theta) \geq f(y, \theta) - f(x \wedge y, \theta).$$

Differential Versions

Theorem

If $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is twice continuously differentiable, then f has increasing differences in (x, θ) iff

$\frac{\partial^2 f(x,\theta)}{\partial x_i \partial \theta_j} \geq 0$ for all $x \in X, \theta \in \Theta, i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$,

and f is supermodular in x iff

$$\frac{\partial^2 f(x,\theta)}{\partial x_i \partial x_j} \geq 0 \text{ for all } x \in X, \theta \in \Theta, i \neq j \in \{1, \dots, n\}.$$

Topkis' Theorem

Theorem

If $X \subseteq \mathbb{R}^n$ is a lattice, $\Theta \subseteq \mathbb{R}^m$, and $f: X \times \Theta \to \mathbb{R}$ has increasing differences in (x, θ) and is supermodular in x, then $X^*(\theta)$ is increasing in the strong set order.

There are also multidimensional versions of the Milgrom-Shannon and Edlin-Shannon theorems.

Application 1: Comparative Statics of Input Utilization

Suppose firm has production function $f: \mathbb{R}^n \to \mathbb{R}$, output price p, input price vector q :

$$\max_{y \in \mathbb{R}^n_+} pf(y) - q \cdot y$$

Assume f non-decreasing and supermodular.

f non-decreasing $\Longrightarrow$ objective has increasing differences in $(y, (p, -q))$.

Theorem

Suppose a competitive firm's production function is increasing and supermodular in its inputs. If the price of the firm's output increases and/or the price of any of its inputs decreases, then the firm increases the usage of all of its inputs.

31 (Formally, $Y^*(p, q)$ increases in the strong set order.)

Application 1.5: The Law of Supply

$$\max_{y \in \mathbb{R}^n_+} pf(y) - q \cdot y$$

Can use Topkis' theorem to give alternative proof of law of supply, without any assumptions on f.

Let

$$\begin{aligned} x &= f(y) \\ c(x) &= \min_{y \in \mathbb{R}^n_+ : f(y) \geq x} q \cdot y \end{aligned}$$

Rewrite problem as

$$\max_{x \in \mathbb{R}} px - c(x)$$

Problem has increasing differences in (x, p), so $x^*(p)$ increasing in strong set order.

(And every selection from $x^*(p)$ is increasing: see pset.)

Application 2: The LeChatelier Principle

"Firms react more to input price changes in the long-run than the short-run."

Suppose inputs are labor and capital, and capital is fixed in short run.

Seems reasonable that if price of labor changes, firm only adjusts labor slightly in short run, stuck with its old capital usage.

In long run, will adjust labor more, once can choose "right" capital usage.

We give example that shows LeChatelier Principle doesn't always apply, and then use Tokpis to formulate rigorous version of the principle.

Example

Firm can produce \$10 of output by using either

1. 2 units of L.

2. 1 unit each of L and K.

Can also shut down and produce nothing.

Initial prices: \$2 per unit of L, \$3 per unit of K.

Firm produces using 2 units of L.

Suppose price of L rises to \$6, K fixed in short run.

In short run, firm shuts down.

In long run, firm produces using 1 unit each of L and K.

In short run, demand for L drops from 2 to 0 .

In long run, goes back up to 1 .

LeChatelier principle fails.

What went wrong?

1 unit of L is complementary with 1 unit of K, but 2 units of L are substitutable with 1 unit of K.

L usage drops from 2 to 0 makes 1 unit of K more valuable ("substitution"), but when K usage rises from 0 to 1 this makes 1 unit of L more valuable ("complementarity").

Suggests LeChatelier principle failed because inputs switched from being complements to substitues at different usage levels.

LeChatelier Revisited

Let

$$\begin{aligned} x(y,\theta) &= \arg \max_{x\in X} f(x,y,\theta) \\ y(\theta) &= \arg \max_{y\in Y} f(x(y,\theta),y,\theta) \end{aligned}$$

$x(y,\theta)$ is optimal "short-run" x (i.e., holding y fixed).

$y(\theta)$ and $x(y(\theta),\theta)$ are optimal "long-run" choices.

Theorem

Suppose $f: X \times Y \times \Theta \to \mathbb{R}$ is supermodular, $\theta \geq \theta'$, and maximizers below are unique. Then

$$x(y(\theta),\theta) \geq x(y(\theta'),\theta) \geq x(y(\theta'),\theta').$$

Corollary (LeChatelier Principle)

Suppose a firm's problem is

$$\max_{K,L\in\mathbb{R}_+} pf(K,L) - wL - rK$$

with either $f_{KL} \geq 0$ for all (K,L) or $f_{KL} \leq 0$ for all (K,L), and suppose K is fixed in the short-run, while L is flexible.

Then, if the wage w increases, the firm's labor usage decreases, and the decrease is larger in the long-run than in the short-run.