Microeconomics

Chapter 2
Profit maximization

Fall 2023

Introduction

Profit = Revenues - Costs

Note that all costs must be included in the calculation of the (economic) profits. For instance, a small businessman should count his own salary, his interest payments for borrowed capital, etc. as costs.

We will generally assume that **firms aim to maximize profit**.

Introduction

Let revenues R depend upon n number of inputs: $R(x_1, ..., x_n) = R(\mathbf{x})$.

Let costs C also depend upon n number of inputs: $C(x_1, ..., x_n) = C(\mathbf{x})$.

Inputs can take many forms, which are often broadly categorized as capital or labor.

A firm uses inputs to maximize profits π:

$$\max_{\mathbf{x}} \pi(\mathbf{x}) = R(\mathbf{x}) - C(\mathbf{x}).$$

Introduction

To maximize profits, the firm should set the derivative of profits $\pi(\mathbf{x})$ towards x_i to zero. It should do this for each i. In particular, a vector containing the optimal inputs $\mathbf{x}^* = (x_1^*, ..., x_n^*)$ is such that:

$$\frac{\partial R(\mathbf{x}^*)}{\partial x_i} - \frac{\partial C(\mathbf{x}^*)}{\partial x_i} = 0, \quad \forall i.$$

The equations above are called **first-order conditions** (FOCs). We can write the FOCs as:

$$\frac{\partial R(\mathbf{x}^*)}{\partial x_i} = \frac{\partial C(\mathbf{x}^*)}{\partial x_i}, \quad \forall i.$$
$$MR_i = MC_i$$

This tells us that **optimal choices are made at the margin**. The intuition is that the firm should use more of input i if $MR_i > MC_i$, and it should use less of input i if $MR_i < MC_i$.

Introduction

For now, we will assume that firms are **price takers**: prices are exogenous variables to the profit maximization problem. That is, prices of inputs and outputs are fixed numbers, and only the level of inputs (and outputs) are chosen by the firm.

When can we expect firms to be price takers? Imagine the following conditions:

- Well informed consumers (students)
- Homogeneous product (beer)
- Large number of firms (student bars in Bairro Alto)

Price taking firms are often referred to as **competitive firms**. We will discuss perfect competition in Chapter 13.

Profit maximization

Let **p** be a row vector of (fixed) prices and let **y** be a column vector of net outputs, then $\pi(\mathbf{p})$ is called the **profit function**:

$$\pi(\mathbf{p}) = \max_{\mathbf{y}} \mathbf{py}$$

Hence, the **profit function** $\pi(\mathbf{p})$ gives us the maximum profits as a function of the prices: For each **p** it uses the feasible technologically **y** that maximizes profits. Note that we can write this as:

$$\begin{aligned}\pi(\mathbf{p}) = \mathbf{py} = (p_1, \dots, p_n) \begin{pmatrix} y_1^o - y_1^i \\ \vdots \\ y_n^o - y_n^i \end{pmatrix} &= \sum_{i=1}^{n} p_i(y_i^o - y_i^i) \\ &= \sum_{i=1}^{n} p_i y_i^o - \sum_{i=1}^{n} p_i y_i^i \\ &= \text{Revenues} - \text{Costs}\end{aligned}$$

Profit maximization

Now imagine that outputs are never used as inputs and inputs are never used outputs. We can write the net output vector $\mathbf{y}$ as $(\mathbf{y}, -\mathbf{x})$:

(1) for good n with $y_n^o > 0$ we know that $y_n^i = 0$, and we relabel y_n^o as y_n,
(2) for good n with $y_n^i > 0$ we know that $y_n^o = 0$, and we relabel y_n^i as x_n.

Then if we have n' outputs and n'' inputs, with $n = n' + n''$, we can write the profit function as:

$$\pi(\mathbf{p}, \mathbf{w}) = \mathbf{py} = \Big(\underbrace{p_1, \ldots, p_{n'}}_{=\mathbf{p}}, \underbrace{w_1, \ldots, w_{n''}}_{=\mathbf{w}}\Big) \begin{pmatrix} \left.\begin{matrix} y_1 \\ \vdots \\ y_{n'} \end{matrix}\right\} =\mathbf{y} \\ -\left.\begin{matrix} x_1 \\ \vdots \\ x_{n''} \end{matrix}\right\} =\text{-}\mathbf{x} \end{pmatrix}$$

$$= \mathbf{py} - \mathbf{wx}$$

$$= \sum_{i=1}^{n'} p_i y_i - \sum_{i=1}^{n''} w_i x_i$$

$$= \text{Revenues} - \text{Costs}$$

Profit maximization

The special case we will mostly analyze is a firm with $n' = 1$ output and $n'' > 1$ inputs. We can write:

$$\begin{aligned}\pi(p, \mathbf{w}) =& py - \mathbf{wx}, \\ =& py - \sum_{i=1}^{n} w_i x_i,\end{aligned}$$

where we relabeled n'' as n. Note that we can now substitute for the production technology of the firm, $y = f(\mathbf{x})$:

$$\begin{aligned}\pi(p, \mathbf{w}) =& pf(\mathbf{x}) - \mathbf{wx} \\ =& pf(\mathbf{x}) - \sum_{i=1}^{n} w_i x_i.\end{aligned}$$

Profit maximization

Recall that the **first-order condition** (FOC) for profit maximization was to set the derivative of $\pi(p, \mathbf{w})$ towards x_i to zero, for each i.

Hence, the FOCs for profit maximizing behavior are:

$$\frac{\partial pf(\mathbf{x})}{\partial x_i} - \frac{\partial \sum_{i=1}^{n} w_i x_i}{\partial x_i} = 0, \quad \forall i.$$

Which can be written as:

$$\begin{aligned} p\frac{\partial f(\mathbf{x})}{\partial x_i} &= w_i, \quad \forall i \\ MR_i &= MC_i. \end{aligned}$$

The vector of inputs $\mathbf{x}$ for which the above FOCs hold is referred to as $\mathbf{x}^*$. Hence, if the firm chooses her inputs to be $\mathbf{x}^* = (x_1^*, ..., x_n^*)$ it maximizes profits.

Profit maximization

Let us introduce the **gradient** $\mathbf{D}f(\mathbf{x})$: the vector of partial derivatives of f with respect to each of its inputs,

$$\mathbf{D}f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}.$$

Using the gradient evaluated at $\mathbf{x}^*$, we can now write the FOCs for all i more efficiently:

$$p\mathbf{D}f(\mathbf{x}^*) = \mathbf{w}.$$

Exercise

Show that

$$p\mathbf{D}f(\mathbf{x}^*) = \mathbf{w},$$

implies the following:

$$p\frac{\partial f(\mathbf{x}^*)}{\partial x_i} - w_i = 0, \quad \forall i.$$

How do you interpret the equations for each *i* above?

Profit maximization graphically

Consider now the case that the firm produces 1 output with $n = 1$ input. We will analyze this special special case because it helps us to introduce and understand **second-order conditions** (SOCs). We can write:

$$\pi(p, w) = py - wx.$$

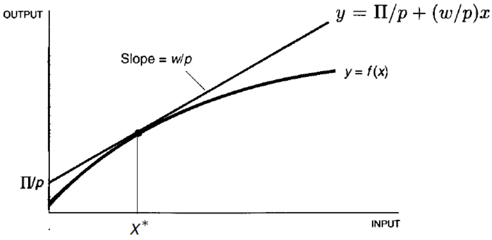
The **isoprofit** line is the level set for a single value of profit Π:

$$L(\Pi) = \{(y, x) : y = \left(\frac{\Pi}{p}\right) + \left(\frac{w}{p}\right) x\}$$

Hence, the isoprofit line reflects all combinations (y, x) that generate profit level Π.

The intercept of the isoprofit line (Π/p) gives the profit measured in terms of the price of output. Note that, since prices are fixed, a higher intercept implies a higher profit.

Profit maximization graphically



The profit-maximizing firm wants to find a **point on the production function** $y = f(x)$ **with the maximal profit**: this is a point where the intercept of the isoprofit line ($\frac{\Pi}{p}$) is maximal. This point x^* is characterized by the slopes of the two lines being equal, which is the optimally condition we have seen before:

$$\frac{\partial f(x^*)}{\partial x} = \frac{w}{p}.$$

Second-order condition

In this two dimensional case, the **second-order condition** (SOC) for profit maximization is intuitive, namely:

$$\frac{\partial^2 f(x^*)}{\partial x^2} \leq 0.$$

Why is this a SOC? With it we can be certain that the tangent line of the production function at x^* lies above the production function.

Tangent line (of $f(x)$ at point $x = x^*$): a straight line that (1) passes through the point $(f(x^*), x^*)$ and (2) has slope $\frac{\partial f(x^*)}{\partial x}$. Hence, the tangent line at x^* is:

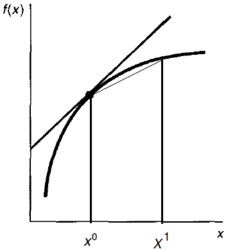
$$f(x^*) + f'(x^*)(x - x^*)$$

Hence, **the isoprofit line is the tangent line at** x^*. So with the SOC we can be certain that the isoprofit line lies above the production function at x^* . This is what we need for the x^* to be optimal.

Concave functions, second derivatives and tangent lines

The following are **useful relationships**:

- $f(x)$ is concave $\leftrightarrow f(x^t) \geq tf(x^0) + (1-t)f(x^1)$ for $0 \leq t \leq 1$.
  (a function f is concave if for every pair of points on its graph, the straight line connecting them lies weakly below f)
- $f(x)$ is concave $\leftrightarrow f''(x) \leq 0$.
- $f(x)$ is concave $\leftrightarrow f(x) \leq f(x^0) + f'(x^0)(x - x^0)$.
  (for a concave function f, the tangent line lies weakly above f)
- If $f(x)$ is concave, and $f'(x^*) = 0$, then x^* maximizes the function.

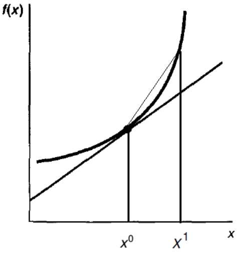


Concave

Convex functions, second derivatives and tangent lines

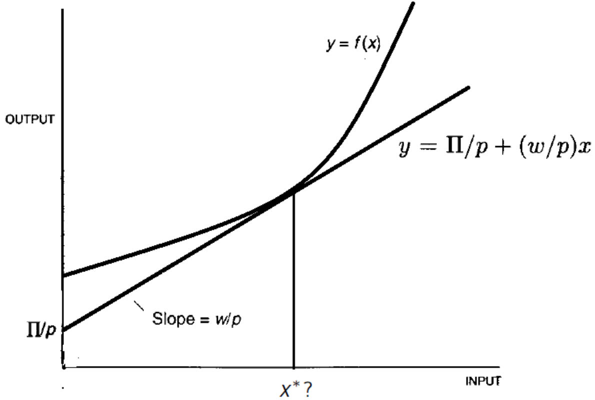
The following are **useful relationships**:

- $f(x)$ is convex $\leftrightarrow f(x^t) \leq tf(x^0) + (1-t)f(x^1)$ for $0 \leq t \leq 1$.
  (a function f is convex if for every pair of points on its graph, the straight line connecting them lies weakly above f)
- $f(x)$ is convex $\leftrightarrow f''(x) \geq 0$.
- $f(x)$ is convex $\leftrightarrow f(x) \geq f(x^0) + f'(x^0)(x - x^0)$.
  (for a convex function f, the tangent line lies weakly below f)
- If $f(x)$ is convex, and $f'(x^*) = 0$, then x^* minimizes the function.



Convex

Profit maximization with FOC, but without SOC



Imagine that $y = f(x)$ is a convex function with $\frac{\partial^2 f(x)}{\partial x^2} = f''(x) \geq 0$, then profit maximization is meaningless.

Second-order condition with $n > 1$*

With $n > 1$ inputs, so that we have $(y, \mathbf{x})$, we can define the **hessian $\mathbf{D}^2 f(\mathbf{x})$**: a matrix of second partial derivatives of f with respect to each of its inputs,

$$\mathbf{D}^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2}, \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2}, \dots, \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1}, \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2}, \dots, \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1}, \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2}, \dots, \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}$$

In the case with $n > 1$ inputs, the SOC for profit maximization is that the hessian of the production function evaluated at $\mathbf{x}^*$ must be **negative semidefinite**:

$$\mathbf{h}^T \mathbf{D}^2 f(\mathbf{x}^*) \mathbf{h} \leq 0 \quad \text{for all vectors } \mathbf{h} = (h_1, ..., h_n) \neq 0.$$

Factor demand and supply function

Factor demand function $\mathbf{x}(p, \mathbf{w})$: a function that gives us the optimal choice of inputs as a function of the prices.

How to get this function? From the FOC we can write $\mathbf{x}$ in terms of $(p, \mathbf{w})$. This gives us a function $\mathbf{x}(p, \mathbf{w})$ so that the FOC of profit maximization holds for every $(p, \mathbf{w})$. Hence, an optimal choice.

Supply function $y = f(\mathbf{x}(p, \mathbf{w}))$: a function that gives us the optimal choice of outputs as a function of the prices.

How to get this function? Substitute $\mathbf{x}(p, \mathbf{w})$ into $y = f(\mathbf{x}) = f(\mathbf{x}(p, \mathbf{w}))$.

Profit function* $\pi(p, \mathbf{w})$: a function that gives us the maximum profits as a function of the prices.

How to get this function? Substitute $\mathbf{x}(p, \mathbf{w})$ in
$\pi = pf(\mathbf{x}) - \mathbf{wx} = pf(\mathbf{x}(p, \mathbf{w})) - \mathbf{wx}(p, \mathbf{w}) = \pi(p, \mathbf{w})$.

Partial and total derivative

Note that the supply function $f(\mathbf{x}(p, \mathbf{w}))$ depends upon p and $\mathbf{w}$ via $\mathbf{x}$, but once you substituted for $\mathbf{x}(p, \mathbf{w})$ it is a function in terms of p and $\mathbf{w}$,

$$f(\mathbf{x}(p, \mathbf{w})) = f(p, \mathbf{w}).$$

We can use the chainrule to write:

$$\frac{\partial f(p, \mathbf{w})}{\partial p} = \frac{\partial f(\mathbf{x})}{\partial x} \frac{\partial x(p, w)}{\partial p} = \frac{df(\mathbf{x}(p, \mathbf{w}))}{dp}.$$

Note that taking the partial derivative towards p, after we substituted for $\mathbf{x}(p, \mathbf{w})$, is similar to taking the total derivative towards p, as we are taking into account that p affects f via $\mathbf{x}$.

Note that formally we have that $\frac{\partial f(\mathbf{x})}{\partial p} = 0$: the function f only depends upon p via $\mathbf{x}$, so that a change in p while keeping $\mathbf{x}$ constant does not change f.

Exercise

Consider a firm with the following production function $f(x) = x^{\alpha}$. The price is fixed at p and the cost for input x is fixed at w. The firm wants to maximize it profits $\pi(p, w) = pf(x) - wx$.

1. Show that the second-order condition requires that $\alpha \leq 1$.
2. Show that $\alpha > 1$ implies IRTS, $\alpha = 1$ implies CRTS, and $\alpha < 1$ implies DRTS.
3. What is the profit maximizing choice of x when $\alpha = 1$?
4. Derive the factor demand function and supply function when $\alpha < 1$.
5. Take the derivative of the factor demand function towards w and p. Interpret the signs of the derivatives.
6. Use the factor supply function derived in question 4 to show that $\frac{\partial f(p,w)}{\partial p}$ is identical to $\frac{\partial f(x)}{\partial x} \frac{\partial x(p,w)}{\partial p}$

Comparative statics

Comparative statics: comparing an equilibrium situation before something changed to an equilibrium situation after something changed.

With equilibrium situation we mean that all agents have been able to make optimal choices, so that no agent has an incentive to deviate.

For instance, a comparative statics exercise is: What happens to the factor demand if we multiply all prices by a factor $t > 0$? The factor demand gives us the optimal choice of inputs given the prices, and since its optimal no firm will have an incentive to deviate.

Comparative statics: example 1

What happens to the factor demand if we multiply all prices by a factor $t > 0$? So we want to know how $x_i(tp, t\mathbf{w})$ and $x_i(p, \mathbf{w})$ are related.

Factor demand functions are homogeneous of degree zero: If we multiply all prices by $t > 0$, the factor demand functions do not change. That is:

$$x_i(tp, t\mathbf{w}) = t^0 x_i(p, \mathbf{w}) = x_i(p, \mathbf{w}).$$

This result follows directly from the FOCs for $\pi(tp, t\mathbf{w}) = tpf(\mathbf{x}) - t\mathbf{wx}$:

$$\begin{aligned} tp\frac{\partial f(\mathbf{x})}{\partial x_i} &= tw_i \quad \forall i, \\ p\frac{\partial f(\mathbf{x})}{\partial x_i} &= w_i, \end{aligned}$$

which are identical to the FOCs for $\pi(p, \mathbf{w}) = pf(\mathbf{x}) - \mathbf{wx}$:

$$p\frac{\partial f(\mathbf{x})}{\partial x_i} = w_i, \quad \forall i.$$

Comparative statics: example 2

How does the factor demand $x_i(p, \mathbf{w})$ respond to an increase in the input price w_i? So we want to know $\frac{\partial x_i(p,\mathbf{w})}{\partial w_i}$.

The factor demand function has a negative slope: when w_i goes up, $x_i(p, \mathbf{w})$ goes down.

Lets consider a firm with $n = 1$ input, and substitute the factor demand function $x(p, w)$ into the FOC:

$$p\frac{\partial f(x(p, w))}{\partial x} = w.$$

Take the derivative of the FOC towards w so that we can write:

$$\begin{aligned} p\frac{\partial^2 f(x)}{\partial x^2}\frac{\partial x(p, w)}{\partial w} &= 1, \\ \frac{\partial x(p, w)}{\partial w} &= \frac{1}{pf''(x)} \leq 0. \end{aligned}$$

Since $f''(x) \leq 0$ by SOC, the factor demand function must have a negative slope: $\frac{\partial x(p,w)}{\partial w} \leq 0$. Hence, when w goes up, $x(p, w)$ goes down.

Comparative statics: example 3

We can also do comparative statics if we imagine observing data on a firm's actual (real life) behavior.

This avoids we have to use theoretical factor demand (and supply) functions. Hence, it avoids we have to make assumptions about the production technology (such as $f''(\cdot) \leq 0$) to conduct comparative statistics.

Consider you observe the following data for a firm: price row vectors $\mathbf{p}^t$ and net output column vectors $\mathbf{y}^t$ for $t = 1, ..., T$. Think of t as time periods, so that we have data for a firm over time t. Note that with this data we can calculate profits for every time period t:

$$\pi^t = \mathbf{p}^\mathbf{t}\mathbf{y}^t, \quad \forall t.$$

Weak Axiom of Profit Maximization

Weak Axiom of Profit Maximization (WAPM): a necessary condition for profit maximization is that,

$$\underbrace{\mathbf{p^t y}^t}_{\text{actual } \pi} \geq \underbrace{\mathbf{p^t y}^s}_{\text{potential } \pi}, \quad \forall t \text{ and } s \neq t.$$

If the firm is maximizing profits, then the observed net output choice $\mathbf{y}^t$ at price $\mathbf{p}^t$ must have a level of profit at least as great as the potential profit at any other net output that the firm could have chosen. We do not know all other choices that are feasible, but we do know some, namely $\mathbf{y^s}$ for $s \neq t$.

Note that WAPM only follows from the definition of profit maximization. We did not make any assumptions on the production technology. Hence, only with data on a firm's prices $\mathbf{p}^t$ and net outputs $\mathbf{y}^t$ across time t you may conclude that the firm makes choices that do not maximize profits.

Exercise

Imagine you observe the following price and net output data across two years for firm A.

Year	price output (p)	price input (w)	output (y)	input (x)
1	2	1	5	5
2	1	2	0	0

1. Does WAPM hold for firm A?
2. What can you conclude about the profit maximizing behavior of firm A?

Homework exercises

Exercises: 2.2, 2.7, and exercises on the slides