Microeconomics

Chapter 2
Profit maximization

Fall 2024

Introduction

Profit = revenues - costs

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

Side note: all costs must be included in the calculation of the (economic) profits. For instance, a businesswoman with a small business should count her own salary, her interest payments for borrowed capital, etc. as costs.

Both revenues and costs depend on actions taken by the firm (e.g., production activities, purchases of factors, advertising,...)

The firm can engage in n such actions

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 x_i 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. Why?
- It should do this for each i. 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:

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

$$MR_i = MC_i$$

$\Rightarrow$ To maximize profits, the firm should use more of input i if $MR_i > MC_i$, and it should use less of input i if $MR_i < MC_i$. In other words, **optimal choices are made at the margin**.

Introduction

What does $MR_i = MC_i$ mean for the firm? The firm makes different choices, for example:

Choosing output

- The firm decides how much to produce.
- The condition says: produce up to the point where the extra revenue from selling one more unit of output equals the extra cost of producing that unit.
- If revenue is greater than cost, produce more. If cost is greater, produce less.

Choosing inputs (like labor)

- The firm also decides how much of each input (like labor, capital, etc.) to use.
- The condition says: hire labor until the extra revenue generated by one more worker equals the extra cost of hiring that worker (the wage).
- Again, if extra revenue is higher than wage, hire more. If it is lower, hire fewer.

Example: Hiring Workers

Setup:

- Production function: $Q(L) = 10L$.
- Output price: $p = 20$.
- Wage per worker: $w = 150$.

1: Revenue and Cost

$$R(L) = p \cdot Q(L) = 20 \cdot 10L = 200L, \quad C(L) = w \cdot L = 150L$$

2: Profit

$$\pi(L) = R(L) - C(L) = 200L - 150L = 50L$$

3: Derivatives

$$MR_L = \frac{\partial R(L)}{\partial L} = 200, \quad MC_L = \frac{\partial C(L)}{\partial L} = 150$$

Decision:

- If $MR_L > MC_L$, hire more workers. ($200 > 150$, so hire.)

Constraints in Profit Maximization

The firm **could** maximize profits by choosing optimal prices for inputs and outputs, as well as optimal levels.

But the firm cannot freely set prices or activity levels, it faces two kinds of constraints.

1. Technological constraints

- Concern the feasibility of the production plan.
- Determined by the production function and available technology (formalized in Chapter 1).

2. Market constraints

- Depend on the actions of other agents in the market.
- Consumers may only be willing to pay certain prices for certain quantities.
- Suppliers may only accept certain prices for their inputs.

Implication: When determining optimal actions, the firm must take into account *both technological and market constraints*.

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 given / fixed numbers / not chosen by the firm.
- **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, homogeneous product, and large number of firms. What happens if a firm deviates?
- Price taking firms are often referred to as **competitive firms**. We will discuss perfect competition in Chapter 13.

Profit maximization

Let us consider the problem of a firm that takes prices as given in both its output and inputs.

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 production plan **y** (in the production possibilities set Y) that maximizes profits. Note that we can write this as:

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

$$= \sum_{i=1}^{n} p_i y_i^o - \sum_{i=1}^{n} p_i y_i^i$$

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

Profit maximization

Consider that outputs are never used as inputs and vice versa. Recall from Chapter 1 that now we can write the net output vector $\mathbf{y}$ as $(\mathbf{y}, -\mathbf{x})$.

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

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

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

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

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

Profit maximization

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

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

where we relabeled n'' as n. We can now substitute for the production technology of the firm, $y = f(\mathbf{x})$, analyzed in Chapter 1:

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

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,

$$\mathbf{D}f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \quad \frac{\partial f(\mathbf{x})}{\partial x_2}, \quad \cdots, \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right).$$

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

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

Value of the marginal product of each factor must be equal to its price (special case of MR = MC).

The Gradient

- **Gradient as a vector:** For two inputs,

$$Df(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2} \right).$$

- **Length of the gradient:** The length (magnitude or norm) is

$$\|Df(x)\| = \sqrt{\left(\frac{\partial f(x)}{\partial x_1} \right)^2 + \left(\frac{\partial f(x)}{\partial x_2} \right)^2}.$$

- **Geometric meaning:**
  The gradient points in the direction of *steepest increase* of the function.
  Its length shows how steep the increase is.
  Intuitively:

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 can write:

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

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

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

  - 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 a higher intercept implies a higher profit (price p is fixed).

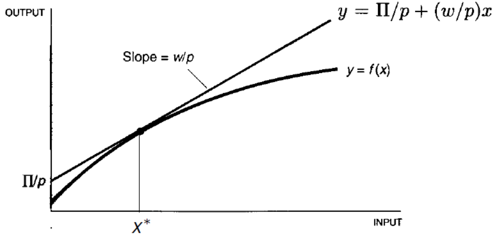
Profit maximization graphically

The profit-maximizing firm wants to find a **point on the production function** $y = f(x)$ **with 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 optimality condition we saw before:

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



Second-order condition

- In this two dimensional case, the **second-order condition** (SOC) is:

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

- When satisfied, it we can be certain that the **isoprofit line** at x^* will always lie above the production function (locally concave). This is what we need for x^* to be optimal.
- Note that the **isoprofit line** is the **tangent line** of $f(x)$ at point $x = x^*$.
- **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:

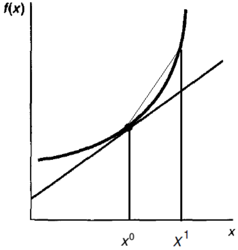
$$f(x^*) + f'(x^*)(x - x^*).$$

Concavity: Equivalent Definitions

- f is concave $\leftrightarrow f(x^t) \geq y^t$ for $0 \leq t \leq 1$.
- f is concave $\leftrightarrow$ points on and below the graph form a convex set.
- f is concave $\leftrightarrow$ upper contour set is a convex set.
- f is concave $\leftrightarrow f''(x) \leq 0$.
- f is concave $\leftrightarrow f(x) \leq f(x^0) + f'(x^0)(x - x^0)$.
- If $f(x)$ is concave, and $f'(x^*) = 0$, then x^* maximizes the function.

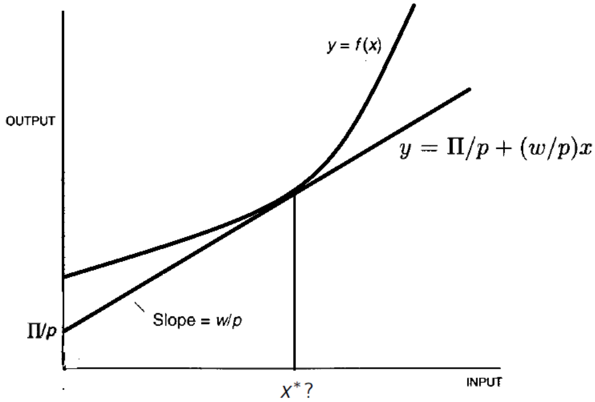
Convexity: Equivalent Definitions

- f is convex $\leftrightarrow f(x^t) \leq y^t$ for $0 \leq t \leq 1$.
- f is convex $\leftrightarrow$ points on and above the graph form a convex set.
- f is convex $\leftrightarrow$ lower contour set is a convex set.
- f is convex $\leftrightarrow f''(x) \geq 0$.
- f is convex $\leftrightarrow f(x) \geq f(x^0) + f'(x^0)(x - x^0)$.
- If f 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 $f''(x) \geq 0$, then profit maximization is meaningless. The FOC holds in point x^*, but the SOC does not.

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 this case, the SOC for profit maximization is that the hessian of the production function must be **negative semidefinite** at the optimal point:

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

Second Order Conditions for Two-Variable Functions

In the 2-variable case, checking for the second order derivatives is not enough. For finding a max, you need to check:

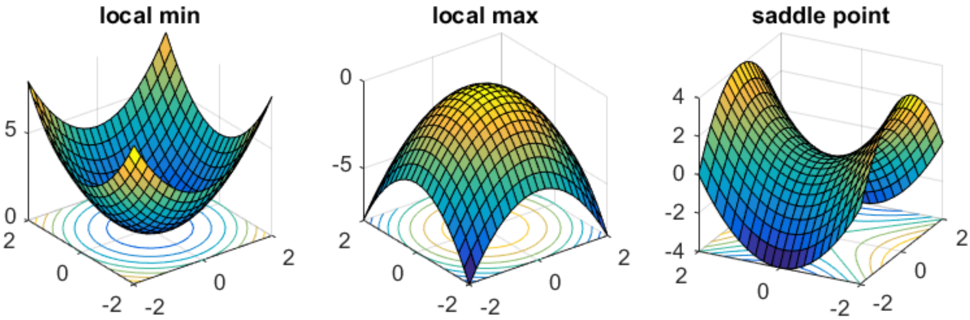
1. $f_{xx} < 0$
2. $f_{yy} < 0$
3. $f_{xx} f_{yy} - f_{xy}^2 > 0$ (determinant of the matrix > 0)

Question: Why do you need the third condition?
Answer: if you only check the sign second-order derivatives f_{xx} and f_{yy}, you cannot be sure whether you have found a max or a saddle point.

Second Order Conditions for Two-Variable Functions (continues)

- $f_{xx} f_{yy} > f_{xy}^2$: minimum (if $f_{xx} > 0$ and $f_{yy} > 0$)
- $f_{xx} f_{yy} < f_{xy}^2$: saddle point
- $f_{xx} f_{yy} = f_{xy}^2$: the test is inconclusive

Visually



At home (optional): Check the three conditions for the following equation $f(x, y) = x^2 + y^2 + 2x$. You can also plot it on WolframAlpha to see it visually and to better understand why the second order conditions alone are not enough.

Exercise

Consider the following function:

$$f(x_1, x_2) = x_1(20 - 5x_1) + x_2(4 - 2x_2) - 2x_1x_2 - 4$$

Maximize $f(x_1, x_2)$ with respect to x_1 and x_2.

1. Compute the first-order conditions (FOCs).
2. Solve FOCs to find the candidate optimum (x_1^*, x_2^*).
3. Check second-order conditions.

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 FOC, write $\mathbf{x}$ in terms of $(p, \mathbf{w})$. Gives 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:

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

That is, 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 $\mathbf{x}$.

In contrast, taking the partial derivative towards p without having substituted for $\mathbf{x}(p, \mathbf{w})$ gives us $\frac{\partial f(\mathbf{x})}{\partial p} = 0$: 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 IRS, $\alpha = 1$ implies CRS, and $\alpha < 1$ implies DRS.
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? i.e., $\frac{\partial x_i(p,\mathbf{w})}{\partial w_i}$.

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 perform comparative statics directly from **observed data** on a firm's behavior and choices.
- This avoids the need to rely solely on **theoretical demand or supply functions**.
- In particular, we do not have to impose strong assumptions about the production technology (e.g. concavity, $f''(\cdot) \leq 0$).
- Suppose we observe, for each period $t = 1, \ldots, T$:
  - a vector of prices $\mathbf{p}^t$ (row),
  - a vector of net outputs $\mathbf{y}^t$ (column).
- With this data, we can compute the firm's profit in each period:

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

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 feasible net outputs $\mathbf{y}$ (in the production possibilities set Y) that the firm could have chosen, 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