ECON5110: Microeconomics

Lecture 2: Sept, 2017

Contents

1 Overview

Much of theoretical economics loosely divides the economy into two groups: the supply side and the demand side, or consumers and producers. We will begin our investigation into economics with production. We will take a very reduced form approach to modeling a firm: we will think of a firm is an entity that takes inputs and creates outputs.

Is this necessarily a good way of representing firms? It abstracts away from many important features of real-life firms including employment, research and development, political involvement, etc. These aspects can be added back in once we understand the basic features of production choices. We will also assume that a firm's ability to turn input into outputs is given by a fixed and known function called its production technology. This makes the firm's decision rather straight forward; it simply maximizes its objective function (usually profit) according its technology. Of course, in reality, a firm's ability to produce is certainly not so clear cut.

2 Production Technology

Assume that there are n goods. We can represent every possible combination of such goods by the points in $\mathbb{R}^n$. A vector $\boldsymbol{x}$ represents x_1 worth of good 1, x_2 of good 2, etc. A **production technology** is a subset of $Y \subseteq \mathbb{R}^n$. A production technology states what allocations are feasible. For a given point, if x_i is negative it is called an **input** if it is positive it is called an **output**. What are some reasonable properties for Y to have?

no free lunch	iff $Y \cap \mathbb{R}^n_+ = 0$,
monotonicity	iff $\boldsymbol{y} \in Y$ implies $\boldsymbol{y}' \in Y$ for all $\boldsymbol{y}' \leq \boldsymbol{y}$,
possibility of inaction	iff $0 \in Y$,
nonincreasing returns to scale	iff $\boldsymbol{y} \in Y$ implies $a\boldsymbol{y} \in Y$ for $a \geq 1$,
nondecreasing returns to scale	iff $\boldsymbol{y} \in Y$ implies $a\boldsymbol{y} \in Y$ for $a \in [0, 1]$,
constant returns to scale	iff $\boldsymbol{y} \in Y$ implies $a\boldsymbol{y} \in Y$ for $a \geq 0$,
additive	iff $\boldsymbol{y}, \boldsymbol{y}' \in Y$ implies $\boldsymbol{y} + \boldsymbol{y}' \in Y$,
convex	iff Y is convex.

The interpretation of outputs is that they are things which can be sold for profit, and therefore as valued. As such, it is inefficient to produce less outputs, or use more inputs, than is possible. A production possibility $\boldsymbol{y} \in Y$ is **efficient** if there is no distinct $\boldsymbol{y}' \in Y$ with $\boldsymbol{y}' \geq \boldsymbol{y}$. We will often assume the existence of a **transformation function**, $T : \mathbb{R}^n \to R$ with the property that $T(\boldsymbol{y}) \leq 0$ for all $\boldsymbol{y} \in Y$ with equality if and only if $\boldsymbol{y}$ is efficient.

Production functions. The use of a production technology is very general, and lets us consider many different complex notions of production. However, we will usually be interested in production technologies that have specific inputs and outputs (and often only 1 output), and which can be described by a production function. Assume there are is 1 output, given by y and $n-1$ inputs given by a vector $\boldsymbol{x}$. Then a **production function**, $f : \mathbb{R}^{n-1}_+ \to R$, which dictates how much output is produced by a given set of inputs. Each production function defines a production technology via:

$$Y = \{(y, -\boldsymbol{x}) \in \mathbb{R}^n \mid y \leq f(\boldsymbol{x})\}.$$

Theorem 1. *Let Y be defined by the production function f. Then Y satisfies monotonicity iff f is monotone, constant returns to scale iff f is h.d.1, and convexity iff y is concave.*

Proof. We will prove the middle claim. Assume Y is CRS. Fix $a \in \mathbb{R}$ Assume $(y, -\boldsymbol{x}) \in Y$ with $y = f(\boldsymbol{x})$. Then by CRS $(ay, -a\boldsymbol{x}) \in Y$ so $ay = af(\alpha) \leq f(a\boldsymbol{x})$. Now let $y = f(\boldsymbol{x}) + \frac{\epsilon}{a}$ for small positive ϵ. Then, $(y, -\boldsymbol{x}) \notin Y$. So $(ay, -a\boldsymbol{x}) \notin Y$ (since then we could multiply by $\frac{1}{a}$ violating our original assumption). So $af(\boldsymbol{x}) + \epsilon > f(a\boldsymbol{x})$; passing to the limit we have $af(\boldsymbol{x}) \geq f(a\boldsymbol{x})$. ∎

Remark 2. *If Y is defined by a production function f then $T : \mathbb{R}^n \to \mathbb{R}$ defined by $T(y, \boldsymbol{x}) = y - f(\boldsymbol{x})$ is a transformation function.*

For a given production function, we can define the **iso-quant** of f by the function $Q(y) = f^{-1}(y)$ (note that f need not be invertible, we allow Q to be set valued). Alternatively, if Y is the associated production technology, $Q(y) = \{\boldsymbol{x} \in \mathbb{R}_+^{n-1} \mid (y, -\boldsymbol{x}) \in Y, \quad y' > y \implies (y' - \boldsymbol{x}) \notin Y\}$. The iso-quant is the set of all inputs which will produce y. These are the level sets of the function f.

Example 1 (Cobb Douglas)**.** *There are two inputs and one output. The production function is given by $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ for $\alpha \in (0, 1)$. Therefore*

$$Y = \{(y, -x_1, -x_2) \in \mathbb{R}^3 \mid y \leq x_1^\alpha x_2^{1-\alpha}, -x_1 \geq 0, -x_2 \geq 0\}$$

and

$$Q(y) = \{(x_1, x_2) \in \mathbb{R}_+^2 \mid x_1^\alpha x_2^{1-\alpha} = y\}.$$

Example 2 (Leontief)**.** *Let $f(x_1, x_2) = \min\{x_1, x_2\}$. This could happen, for example when the inputs are capital and labor, and each machine needs an operator, and an operator without a machine is likewise idle. We can draw the iso-quants of such a firm.*

Notice that the Cobb-Douglas function is h.d.1 $f(a\boldsymbol{x}) = (ax_1)^\alpha (ax_2)^{1-\alpha} = a^\alpha a^{1-\alpha} x_1^\alpha x_2^{1-\alpha} = af(\boldsymbol{x})$. This implies that the CD function is constant returns to scale by Theorem 1. Moreover, by examination of the Hessian of f (which we will not do) we can see that f is concave, so that CD is a convex technology.

Whenever Y is convex then the **upper contour sets**, $UC(y) = \{\boldsymbol{x} \in \mathbb{R}_+^{n-1} \mid f(\boldsymbol{x}) \geq y\}$ is convex. This is immediate since the y coordinate is simply invariant under linear combinations. What is $UC(y)$? It is the set that lies above the iso-quant $Q(y)$. Convexity of $UC(y)$ is a important feature. It states essentially that the firm can always do at least as well by mixing together different inputs as it can by using the inputs individually. For example, if the firm can produce 1 unit with 1 unit of x_1 alone or with 1 unit of x_2 alone, (so that $(1,0),(0,1) \in UC(1)$) then it must be that using $\frac{1}{2}$ each produces at least 1 because then $(\frac{1}{2},\frac{1}{2}) \in UC(1)$.

This raises the question: when is $UC(y)$ convex? We know that it so when Y is convex, which happens whenever f is concave. But is there a more general solution. A theorem from last lecture, stating that the level sets of f are convex if and only if f is quasi concave provides the answer.

Substitution and Homotheticity. The iso-quant lines trace out bundles on inputs that all result in the same output. These are principally of interest because they provide insight into how one input can be transformed into another while keeping output fixed. This will give us insight into the demand for inputs, there relation to prices, etc, when we begin to think about the *choices* of a firm. But even without thinking about optimization yet, the tradeoffs between different inputs is an intuitive property to understand the larger production technology.

The **technical rate of substitution** or MRT is the amount by which input i changes in response to an change in input j, so as to keep the level of output constant. Consider the curve, $f(x_1,x_2) = y$. Under basic conditions such as monotonicity, we could think of x_2 as being a function of x_1 (if we choose x_1 then x_2 can be solved for). We are interested in $\frac{\partial x_2}{\partial x_1}$. Differentiating $f(x_1,x_2(x_1)) = y$ on both sides gives us $\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial x_1} = 0$, or

$$MRT = \frac{\partial x_2}{\partial x_1} = -\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}}$$

That is, the MRT is the ratio of the change output by changing the two different inputs. This has a clear geometric interpretation, the MRT is the slope of the iso-quant lines.

Example 1 (Cobb Douglas)**.** *What is the MRT of the CD production function? Well:*

$$\begin{aligned}\frac{\partial x_2}{\partial x_1} &= -\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} \\ &-\frac{\alpha x_1^{\alpha-1}x_2^{1-\alpha}}{(1-\alpha)x_1^{\alpha}x_2^{-\alpha}} \\ &-\frac{\alpha x_2}{(1-\alpha)x_1}\end{aligned}$$

A function is said to be **homothetic** if it is a positive monotone transformation of a h.d.1 function. That is, if there is a way to change the scale at which we measure inputs and outputs (without changing our notion of ordering) such that things look like they are constant returns to scale. Assume that $\boldsymbol{x}, \boldsymbol{x}' \in Q(y)$ so that both input bundles produce the same output. If f is h.d.1 then we know that $a\boldsymbol{x}, a\boldsymbol{y} \in Q(ay)$—this is another way of stating that the technology is constant returns to scale. If f is homothetic, it does not necessarily hold that $a\boldsymbol{x}$ or $a\boldsymbol{x}'$ produce ay but they still must produce the *same* amount. Why?

A homothetic functions MRT does not depend on the amount being produced. Recall from the previous lecture that a function that is h.d.1 has a partial derivative that is h.d.0.

3 Profit Maximization

In the last section, we saw different ways of representing and analyzing what a firm *could* do, that is what is physically or technologically possible. But what will the firm choose? We made a passing reference to the idea that a firm should choose an *efficient* point, but this raises two more questions: what assumptions ensure efficiency and which, of the many efficient points will be chosen.

We will now provide a behavioral foundation to the firm's choice, that it is trying to maximize profit. Profit is the value created by selling (or using, trading, etc) the outputs less the value expended to obtain the inputs (inputs can include non-tangibles such as time, effort, etc). The former is called revenue and the later costs. A firm can choose set of inputs $\boldsymbol{x} \in \mathbb{R}^n$, and we assume that the corresponding revenue and costs are given: $R(\boldsymbol{x})$ and $C(\boldsymbol{x})$.

Before we make more concrete assumptions about the flavor of R and C we can see the following: The firm faces the problem

$$\max_{Y} \pi(\boldsymbol{x}) = R(\boldsymbol{x}) - C(\boldsymbol{x}),$$

which, when everything is nicely behaved, tells us that at the optimal choice $\frac{\partial R}{\partial \boldsymbol{x}} = \frac{\partial C}{\partial \boldsymbol{x}}$. This is the adage, and cornerstone of undergrad economic wisdom: marginal revenue equals marginal cost.

But to get more out of the analysis, we must put more in. Let p denote the **price** of the output (which can be a vector if there are multiple outputs) and $\boldsymbol{c}$ denote the vector of **costs** for each input. Then

$$\pi(p, \boldsymbol{c}, y, \boldsymbol{x}) = py - \boldsymbol{c}\boldsymbol{x},$$

where p and $\boldsymbol{c}$ are parameters of the function. When Y is given by a production function f this problem becomes

$$\pi(p, \boldsymbol{c}, \boldsymbol{x}) = pf(\boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x}.$$

If we maximize the above we see that for each input, the $p\nabla f = \boldsymbol{c}$. The marginal contribution of each input on production (times price) must be equal to the marginal cost of obtaining more of that unit. We can solve the first order conditions to find the optimal inputs for a given price, that is $x^\star(p, \boldsymbol{c})$, called the **factor demand function** of the firm. Plugging these into the production function, $y^\star(p, \boldsymbol{x}) = f(\boldsymbol{x}^\star(p, \boldsymbol{c}))$, givens us the **supply function** of the firm. These together define the **optimal profit function** of the firm $\pi^\star(p, \boldsymbol{c}) = py^\star(p, \boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x}^\star(p, \boldsymbol{c})$.

Example 2. *Let the production function be given by $f(x_1, x_2) = a\ln(x_1) + b\ln(x_2)$. The profit function is*

$$p(a\ln(x_1) + b\ln(x_2)) - c_1x_1 - c_2x_2 \tag{π}$$

The corresponding first order conditions for each,

$$\frac{pa}{x_1} = c_1 \tag{π_1}$$

$$\frac{pb}{x_2} = c_2 \tag{π_2}$$

rearranging, we get $x_1^\star(p, c) = \frac{ap}{c_1}$ and $x_2^\star(p, c) = \frac{bp}{c_2}$. Plugging this into the production function, we get $y^\star(p, c) = a\ln(\frac{ap}{c_1}) + b\ln(\frac{bp}{c_2})$. The profit of the firm is can be figured out by plugging things into other things.

Notice we can now preform comparative statics on these functions. $\frac{\partial x_1^\star}{\partial c_1} = \frac{-ap}{c_1^2}$. So as we would intuit, the firm uses less input when the cost increases (and therefore also decreases production). $\frac{\partial x_1^\star}{\partial p} = \frac{a}{c}$ which is positive. This means as the price increases, the firm used more of each input and therefore production increases.

What happens if both the price of the outputs and the cost of the input increase the same factor? That is what is $x^\star(tp, tc)$ for some $t \geq 0$. Well, $x_1^\star(tp, tc) = \frac{atp}{tc_1} = \frac{ap}{c_1} = x_1^\star(p, c)$.

Example 5. *Let the production function be given by $f(x) = 10x - x^2$. Therefore the profit function is*

$$p10x - px^2 - cx. \tag{π}$$

The first order condition for maximization is

$$p(10 - 2x) = c. \tag{π_x}$$

which can be rearranged to solve for the factor demand: $x^\star(p, c) = 5 - \frac{c}{2p}$. We really need to check the second order conditions but it is clear the function is concave. Examining $x^\star$ we see *it is increasing in* p*, decreasing in* c *and positive (i.e., a valid solution) if* $10p \geq c$*. The supply function is ...*

4 Properties of Profit Maximization

If we were to look at a particular firm and ask "is this firm maximizing profit with respect to some demand function?" how could be answer the question. If we knew the profit function the answer would be as simple as to calculate the maximal allocation and compare, but often we do not have access to such information. More likely, we (as economists) observe the input and output decisions for a firm at a variety of prices and times. Can we ever falsify the profit maximization assumption from this data?

In the previous examples, we see that the factor demand of the CD production function (with a single input) is h.d.0. This makes intuitive sense: if everything becomes more expensive at the same rate (for example, under inflation) then revenue increases by t as does costs. We have shifted everything up by a constant factor, but this shouldn't change what allocation is most profitable. Such a line of reasoning, however, does not rely on the CD functional form at all, and, indeed, such an observation can be generalized.

Theorem 3. *Let* f *be a production function and let* $\pi = pf(\boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x}$*, Then the factor demand,* $x^\star(p, \boldsymbol{c})$ *is h.d.0.*

Lemma 1. *Let* $f : X \to \mathbb{R}$ *and let* $x^\star$ *denote the argmax of the function. Then for any monotone function* $h : \mathbb{R} \to \mathbb{R}$ *we have* $x^\star$ *is an argmax of* $h \circ f$*.*

Proof of Lemma. By definition, for all $y \in X$ $f(x^\star) \geq f(y)$. By the monotonicity of h, we have $h(f(x^\star)) \geq h(f(y))$, so x is a maximum. ∎

Proof of Theorem. Let $h : \mathbb{R} \to \mathbb{R}$ be the map $h : x \mapsto tx$. Then $h \circ \pi(p, \boldsymbol{c}) = t(pf(\boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x}) = tpf(\boldsymbol{x}) + t\boldsymbol{c}\boldsymbol{x} = \pi(tp, t\boldsymbol{c})$. Since h is monotone, we can apply the above lemma. So if $\boldsymbol{x}^\star$ is a maximizer of $\pi(p, \boldsymbol{c})$ it is also a maximizer $\pi(tp, t\boldsymbol{c})$. ∎

Theorem 3 can be immediately extended to the supply function of the firm, $y^\star$. This means that if we observe the decisions of a firm under two different price vectors, one a linear translation of the other it *must be* that the the factor demand and supply of the firm are unchanged *if* the firm is profit maximizing. Of course, this observation does not guarantee the firm is profit maximizing. That is: the h.d.0 supply is a necessary but not sufficient condition.

Properties of the Profit Function. Can we say anything about the profit function? Of course. For starters, the above Theorem extends cleanly the profit function. Since $\boldsymbol{x}^*$ is h.d.0., the profit function is h.d.1.

Theorem 4. *The optimal profit function $\pi^\star(p, \boldsymbol{c})$ is convex in $(p, \boldsymbol{c})$.*

Proof. Consider $(p, \boldsymbol{c})$, $(p', \boldsymbol{c}')$ and $(p^\alpha, \boldsymbol{c}^\alpha) = \alpha(p, \boldsymbol{c}) + (1-\alpha)(p', \boldsymbol{c}')$. Let $\boldsymbol{x}$, $\boldsymbol{x}'$ and $\boldsymbol{x}^\alpha$ denote the corresponding factor demands (eschewing the $\star$ in the name of notational cleanliness). By definition of an factor demands (i.e., that they optimize profit) we have

$$pf(\boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x} \geq pf(\boldsymbol{x}^\alpha) - \boldsymbol{c}\boldsymbol{x}^\alpha$$
$$p'f(\boldsymbol{x}') - \boldsymbol{c}'\boldsymbol{x}' \geq p'f(\boldsymbol{x}^\alpha) - \boldsymbol{c}'\boldsymbol{x}^\alpha$$

Multiplying the first equation by α and the second by $(1-\alpha)$ and summing we obtain

$$\alpha(pf(\boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x}) + (1-\alpha)(p'f(\boldsymbol{x}') - \boldsymbol{c}'\boldsymbol{x}') \geq p^\alpha f(\boldsymbol{x}^\alpha) - \boldsymbol{c}^\alpha(\boldsymbol{x}^\alpha)$$

or that $\alpha\pi^\star(p, \boldsymbol{c}) + (1-\alpha)\pi^\star(p', \boldsymbol{c}') \geq \pi^\star(\alpha(p, \boldsymbol{c}) + (1-\alpha)(p', \boldsymbol{c}'))$, which is what we set out to prove. ■

This is really just a particular case of the general principle that the max operator is convex. What is the intuition behind this result? If prices increase a fixed amount t, the firm could easily obtain exactly t times its profits. It can achieve this by doing nothing at all, producing exactly the same amount as before. Every dollar made will become t dollars. This is a linear increase. But the firm *might* be able to do better. By taking advantage of substitutions between inputs, the firm might be able to do better than t-fold increase. Hence, the profit function is super-linear, i.e., convex.

Another thing to notice in the previous examples was that the factor demands were decreasing in cost and the supply was increasing in price. We will momentarily show this holds in general, by exploiting the following fact about the profit function:

Theorem 5. *The profit function is non-decreasing in price and non-increasing in costs.*

Proof. Consider $(p, \boldsymbol{c})$, $(p', \boldsymbol{c}')$ such that $p \geq p'$ and $\boldsymbol{c}' \geq \boldsymbol{c}$. Let $\boldsymbol{x}$ and $\boldsymbol{x}'$ be the corresponding factor demands. We have

$$\begin{aligned}\pi(p,\boldsymbol{c}) &= pf(\boldsymbol{x}) - \boldsymbol{c}\boldsymbol{x} \\ &\geq pf(\boldsymbol{x}') - \boldsymbol{c}\boldsymbol{x}' \\ &\geq p'f(\boldsymbol{x}') - \boldsymbol{c}'\boldsymbol{x}' = \pi(p',\boldsymbol{c}')\end{aligned}$$

where the first inequality is by the definition of factor demand being the argmax and the second is by the assumption on prices and costs. ■

Theorem 5 is highly intuitive. It also allows us to talk about how factor demand and supply change as a function of the parameters, cost and price.

Theorem 6 (Hotelling's Lemma)**.** *The supply function of the firm is the derivative of profit with respect to price:* $y^\star = \frac{\partial \pi^\star}{\partial p}$.

Proof by Envelope Theorem. We have $\pi^\star(p,\boldsymbol{c}) = pf(\boldsymbol{x}^\star(p,\boldsymbol{c})) - \boldsymbol{c}\boldsymbol{x}^\star(p,\boldsymbol{c})$. By the envelope theorem, we have $\frac{\partial \pi^\star}{\partial p} = \frac{\partial \pi}{\partial p}$ evaluated at the optimum. I.e., $\frac{\partial \pi^\star}{\partial p} = f(\boldsymbol{x}^\star(p,\boldsymbol{c})) = y^\star(p,\boldsymbol{c})$. ■

Theorem 6 is less intuitive. The best understanding of this arises from the intuition behind the envelope theorem itself: when profit is maximizes, small changes in supply will have no effect on profits (i.e,. the first order condition). This means that a small change in the price, while it may lead to a slight adjustment in supply, will not effect profit via the supply channel. Thus, the only change in profit is simply the change that results from selling each unit of output for a slightly different amount. Of course the magnitude of this change is exactly how much is sold: the supply of the firm.

We can combine the above results to examine the effect of price on supply, by differentiating again.

Theorem 7 (Law of Supply)**.** *The supply function of the firm is an increasing function of price:* $\frac{\partial y^\star}{\partial p} \geq 0$.

Proof. From Theorem 6 we have $\frac{\partial y^\star}{\partial p} = \frac{\partial^2 \pi^\star}{\partial p^2} \geq 0$, where the later hold via the convexity of the profit function, as proven by Theorem 4. ■

Notice that all of these results follow directly from the assumption of profit maximization, without *any* addition assumptions on the shape of the production function (in fact, everything can be proven using a production technology Y, rather than a production function).

5 Cost Minimization

Often firms cannot simply produce *any* amount of a good. Perhaps there are limitations of demand, or they have been contracted to fill a particular order. It is often the case that a firm might want to produce a specific amount of a good, and to do so in the cheapest way possible (this is true even in the absence of competitive markets, making cost minimization theory a potentially more general way of looking at the problem.)

Say that a firm has production function, $f : \mathbb{R}^n \to \mathbb{R}$, and wants to produce $y \in R$ units of output. To minimize its cost, it solves the problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \boldsymbol{cx} \qquad \text{subject to } f(\boldsymbol{x}) \geq y$$

Just like the profit maximization, the firm's choice is the allocation of inputs. We will use the Lagrangian method to solve these problems. Whenever f is monotone in $\boldsymbol{x}$ and $\boldsymbol{c} > 0$ then we know that the firm will produce exactly y (make sure you understand exactly why this is, such an argument pops up often). Hence, we can replace the constraint with $f(\boldsymbol{x}) = y$ (to avoid the additional complications of KKT). The Lagrangian:

$$\mathcal{L} = \boldsymbol{cx} - \lambda(f(\boldsymbol{x}) - y),$$

with first order conditions given by:

$$c_i = \lambda \frac{\partial f}{\partial x_i} \tag{$\mathcal{L}_i$}$$

$$f(\boldsymbol{x}) = 0 \tag{$\mathcal{L}_\lambda$}$$

When there is a single input, the solution is trivial: solve the output constraint. When there are more than 1 output, we can divide the first order-conditions to obtain:

$$\frac{c_1}{c_j} = \frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial x_j}},$$

the right hand side of which is the MRT of j with respect to i. At the optimum, this must be: holding output fixed at y, the contribution of the two goods must be proportional to the costs.

Again, we can define the **conditional factor demand** $\boldsymbol{x}^\star(\boldsymbol{c}, y)$ that solves the problem and the **cost function** of the firm is $C^\star(\boldsymbol{c}, y) = \boldsymbol{cx}^\star(\boldsymbol{c}, y)$.

Example 6. *Consider a firm the the production function $f(x_1, x_2) = x_1^\alpha x_2^\beta$. Thus the firm* *wants to solve the problem:*

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} c_1x_1 + c_2x_2 \qquad \textit{subject to } x_1^\alpha x_2^\beta \geq y.$$

The associated Lagrangian is

$$c_1x_1 + c_2x_2 - \lambda(y - x_1^\alpha x_2^\beta) \tag{$\mathcal{L}$}$$

and first order conditions:

$$c_1 = -\lambda\alpha x_1^{\alpha-1}x_2^\beta \tag{$\mathcal{L}_1$}$$

$$c_2 = -\lambda\beta x_1^\alpha x_2^{\beta-1} \tag{$\mathcal{L}_2$}$$

$$y = x_1^\alpha x_2^\beta \tag{$\mathcal{L}_\lambda$}$$

Dividing ($\mathcal{L}_1$) by ($\mathcal{L}_2$) we get $\frac{c_1}{c_2} = \frac{\alpha}{\beta}\frac{x_2}{x_1}$ or $x_2 = \frac{c_1}{c_2}\frac{\beta}{\alpha}x_1$. Plugging this into ($\mathcal{L}_\lambda$) as simplifying

$$x_1^\star = (\frac{c_1}{c_2}\frac{\beta}{\alpha})^{\frac{-\beta}{\alpha+\beta}} y^{\frac{1}{\alpha+\beta}}$$

Example 7 (Leontief)**.** *Assume strictly positive costs. The function of a firm is trying to minimize is therefore:*

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} c_1x_1 + c_2x_2 \qquad \textit{subject to } \min\{\boldsymbol{x}_1, \boldsymbol{x}_2\} \geq y$$

What are the first order conditions of the corresponding maximization problem? Notice, $\frac{\partial \min\{x_1,x_2\}}{\partial x_1}$ *is 1 whenever* x_1 *is strictly less that* x_2*, 0 when it is strictly greater and undefined when they are equal. Calculus will not help us here. While this is unfortunate, we can use an alternative, more precious form of attack: common sense.*

If, at the optimum, we had $x_1 > y$ *then the firm could clearly lower costs by reducing* x_2 *since this does not affect output at all! An analogous argument rules out* $x_2 > y$ *and we are left with* $x_1 = x_2 = y$*. The cost is obvious.*

Example 8. *Imagine a firm has two plants with different technologies. The cost of producing* y *at the first plant is* $C_1^\star(y_1) = y^2$ *and at the second is* $C_2^\star(y_2) = y_2$*. The firm needs to produce* y *and can do this by splitting production across the two plants in any manner (so long as each produces a weakly positive amount). What is the cost function of the firm in terms of* y*?*

At the optimum its clear that the marginal cost at each plant must be the same. To see this notice that the firm is choosing how much to produce at plant 1 and therefore seeks to minimize: $C_1^\star(y_1) + C_2^\star(y - y_1)$*. The optimality condition is thus* $2y_1 = 1$ *or* $y_1 = \frac{1}{2}$*. Of course, this is* *only possible when $y \geq \frac{1}{2}$, else $y_1 = y$. The cost (when $y \geq \frac{1}{2}$) is therefore: $y - \frac{1}{4}$.*

Properties of the Cost Function. We can prove dual claims to the theorems of the previous section, using similar methods:

Theorem 8. *Conditional factor demand is h.d.0. The cost function is h.d.1 in $\boldsymbol{c}$ and if f is constant h.d.1 (so that Y is constant returns to scale) then it is also h.d.1 in p, i.e., $C^\star(\boldsymbol{c}, y) = yC(\boldsymbol{c}, 1)$.*

Proof. The first two claims follow similar arguments as under profit maximization. Towards the later: Let $\boldsymbol{x}$ denote $\boldsymbol{x}^\star(\boldsymbol{c}, y)$ so that $f(\boldsymbol{x}) = y$. Then $f(\frac{\boldsymbol{x}}{y}) = 1$ by h.d.1 of f. Further, $\boldsymbol{c}\frac{\boldsymbol{x}}{y} = \frac{C^\star(\boldsymbol{c},y)}{y}$. Hence, $C^\star(\boldsymbol{c}, 1) \leq \frac{C^\star(\boldsymbol{c},y)}{y}$ or $yC^\star(\boldsymbol{c}, 1) \leq C^\star(\boldsymbol{c}, y)$. Now let $\boldsymbol{x}'$ denote $\boldsymbol{x}^\star(\boldsymbol{c}, 1)$ so that $f(\boldsymbol{x}) = 1$. Then $f(y\boldsymbol{x}) = y$ by h.d.1 of f. Further, $\boldsymbol{c}y\boldsymbol{x} = yC^\star(\boldsymbol{c}, 1)$. Hence, $C^\star(\boldsymbol{c}, y) \leq yC^\star(\boldsymbol{c}, 1)$. ■

Theorem 9. *The cost function is non-decreasing in y and $\boldsymbol{c}$.*

Theorem 10. *The cost function of the is concave in $\boldsymbol{c}$.*

The proofs of these Theorems are in direct analogy to the proof of Theorem 4

Theorem 11 (Shepard's Lemma)**.** *The conditional factor demand of input i is the derivative of cost with respect to c_i: $x^\star = \frac{\partial C^\star}{\partial c_i}$*

Proof. We appeal to the constrained version of the envelope theorem:

$$\frac{\partial C^\star}{\partial c_i} = \frac{\partial \mathcal{L}}{\partial c_i} = \frac{\partial \boldsymbol{c}\boldsymbol{x}}{\partial c_i} - \lambda \frac{\partial f}{\partial c_i} = x_i^\star.$$

■

The above two results show that factor demand is downward sloping in price. The proof and intuition echo what we have already done; the derivative of factor demand is the second derivative of the concave cost function.

6 Duality

From the vantage of the above results, you might have the sneaking suspicion that that profit maximization and cost minimization are describing the same behavior. These are not two totally distinct management styles but are in fact two sides of the same coin. This is in fact something that can be made totally rigorous.

Theorem 12. *Let f be a production function, with supply function $y^\star(p, \boldsymbol{c})$ and factor demand $x^\star(p, \boldsymbol{c})$. Then $x^\star(p, \boldsymbol{c})$ solves the cost minimization problem $\min \boldsymbol{c}\boldsymbol{x}$ subject to $f(\boldsymbol{x}) \geq y^\star(p, \boldsymbol{c})$.*

Proof. By contradiction. Assume this was not the case so that some other vector $\boldsymbol{x}'$ incurred a strictly lower cost to produce $y^\star(p, \boldsymbol{c})$. Then

$$\begin{aligned} pf(\boldsymbol{x}') - \boldsymbol{c}\boldsymbol{x}' &= py^\star(p, \boldsymbol{c}) - \boldsymbol{c}\boldsymbol{x}' \\ &> py^\star(p, \boldsymbol{c}) - \boldsymbol{c}\boldsymbol{x}^\star(p, \boldsymbol{c}) = pf(\boldsymbol{x}^\star(p, \boldsymbol{c}) - \boldsymbol{c}\boldsymbol{x}^\star(p, \boldsymbol{c}), \end{aligned}$$

a contradiction to the assumption that $\boldsymbol{x}^\star$ was a maximizer. ■

What does this mean? A firm that is maximizing profits, must be minimizing costs at the given level of production. This makes sense when contemplated upon, even if it is a little surprising at first. We could think of profit maximization in two stages: choose the level of production, choose the allocations of inputs to produce that level most efficiently. Both of these stages must be maximized to fully maximize profit.