Microeconomics - Firm Theory

Renato E. Molina [*]

December 29, 2018

Contents

1 Primitive Notions

Very much like the hypothesis of utility maximization for consumers, profit maximization is the most robust and compelling assumption to examine and predict firm behavior. In any choice the firm makes, we impose this assumption, saying that its decision is guided by profit maximization. The set of actions that achieves that goal depends on the circumstances the firm faces, namely input markets and output markets. Overall firm behavior will depend on the firms objectives, which we assume remain constant, and its constraints, which are varied and depend on the market conditions.

2 Production

We define production as the process of transforming inputs into outputs. Define **technological feasibility** as the state of technology that allows the firm to produce according to its restrictions. Let the **production possibility set**, $Y \subset \mathbb{R}^m$, where each vector $\boldsymbol{y} = (y_1, ..., y_n) \in Y$ is a **production plan** whose components indicate the amount of various inputs and outputs. The common convention is to write elements of $\boldsymbol{y} \in Y$ so that $y_i < 0$ if resource i is used up in the production plan, and $y_i > 0$ if resource i produced in the production plan.

[*]Most of these notes follow the third edition of the book "Advanced Microeconomic Theory" by G. Jehle & P. Reny. All errors are my own.

Often we want to characterize the production possibility set as a firm producing only one output, which is characterized by the **production function**. We denote output by y, and the amount of input i by x_i, so that with n inputs, the entire vector of inputs is denoted as $\boldsymbol{x} = (x_1, .., x_n)$, where we require $y \geq 0$ and $\boldsymbol{x} \geq 0$. The production function, f, is therefore a mapping $\mathbb{R}^n_+ \to \mathbb{R}_+$. We write $y = f(\boldsymbol{x})$ with:

Assumption 2.1. ***Properties of the Production Function.*** *The production function $f : \mathbb{R}^n_+ \to \mathbb{R}_+$, is continuous, strictly increasing, and strictly quasiconcave on $\mathbb{R}^n_+$, and $f(\mathbf{0}) = 0$.*

Continuity ensures that small increases in the input vector result in small changes in output. Strictly increasing implies that more input will lead to more output. Strict quasiconcavity is mostly assumed to simplify scrutiny of the production behavior. Moreover, it implies that a positive amount of output requires positive amounts of some of the inputs.

When the production function is differentiable, its partial derivative, $\partial f(\boldsymbol{x})/\partial x_i$, is called the **marginal product** of input i, and provides the rate at which output changes per additional unit of input i. From Assumption (2.1) it follows that $\partial f(\boldsymbol{x})/\partial x_i > 0$ for almost all input vectors. For simplicity, however, we assume for the strict equality always holds.

For any fixed level of output, y, the set of input vectors producing y units of output is called the y-level **isoquant**. An isoquant is then the level set of f. We denote isoquants as $Q(y)$:

$$Q(y) \equiv \{\boldsymbol{x} > 0 | f(\boldsymbol{x}) = y\} \tag{1}$$

For an input vector $\boldsymbol{x}$, the **isoquant through $\boldsymbol{x}$** is the set of input vectors producing the same output as $\boldsymbol{x}$, namely, $Q(f(\boldsymbol{x}))$.

As an analogous concept for the marginal rate of substitution for consumers, we also have the **marginal rate of technical substitution (MRTS)** for firms. It measures the rate which one input can be substituted for another without changing the amount of output produced. Formally, the marginal rate of technical substitution of input j for input i, when the current input vector is $\boldsymbol{x}$, denoted $MRTS_{ij}(\boldsymbol{x})$, is defined as the ratio between marginal products:

$$MRTS_{ij}(\boldsymbol{x}) = \frac{\partial f(\boldsymbol{x})/\partial x_i}{\partial f(\boldsymbol{x})/\partial x_j} \tag{2}$$

In general, the MRTS between any two inputs depends on the amount of all inputs employed. Nevertheless, it is common to suppose that inputs can be classified into types. The degrees of substitution differ systematically between inputs of different types. Production functions of this type are called **separable**, and there are at least two major forms of separability:

Definition 2.1. Separable Production Functions. Let $N = \{1, ..., n\}$ index the set of all inputs, and suppose that these inputs can be divided into $S > 1$ mutually exclusive and exhaustive subsets, $N_1, .., N_S$. The production function is called weakly separable if the MRTS between two inputs within the same group is independent of inputs used in other groups:

$$\frac{\partial(f_i(\boldsymbol{x}/f_j(\boldsymbol{x}))}{\partial x_k} = 0, \quad \forall\, i, j \in N_S \;\wedge\; k \notin N_S \tag{3}$$

The MRTS is a local measure of substitutability between inputs in producing a given level of output. We can refine this measurement with the **elasticity of substitution**, σ. Holding all other inputs constant and the level of output constant, the elasticity of substitution of input j for input i can be defined as the percentage change in the input proportions, x_j/x_i, associated with a one percent change in the MRTS between them:

Definition 2.2. The Elasticity of Substitution. For a production function $f(\boldsymbol{x})$, the elasticity of substitution of input j for input i at the point $\boldsymbol{x}^0 \in \mathbb{R}_{++}$ is defined as:

$$\sigma_{ij} \equiv \left(\frac{d\ln[MRTS_{ij}(\boldsymbol{x})(r)]}{d\ln[r]} \bigg|_{r=x_j^0/x_i^0} \right)^{-1} \tag{4}$$

with $\boldsymbol{x}(r)$ as the unique vector of inputs $\boldsymbol{x} = (x_1, ..., x_n)$, such that (i) $x_j/x_i = r$, (ii) $x_k = x_k^0$ for $k \neq i, j$, and (iii) $f(\boldsymbol{x}) = f(\boldsymbol{x}^0)$.

The elasticity of substitution, $\sigma_{ij}(\boldsymbol{x}^0)$, is a measure of the curvature of the $i-j$ isoquant through $\boldsymbol{x}^0$ at $\boldsymbol{x}^0$. When the production function is **quasiconcave**, the elasticity of substitution can **never** be **negative**, so $\sigma_{ij} \geq 0$. In general, the closer it is to zero, the more difficult it is to substitute between inputs, while the opposite it is true when its value is large. To see this measure at work, consider the **constant elasticity of substitution (CES)** function:

$$y = (x_1^\rho + x_2^\rho), \; 0 \neq \rho < 1 \tag{5}$$

First we calculate its MRTS:

$$MRTS_{12} = \left(\frac{x_2}{x_2} \right)^{1-\rho} \tag{6}$$

Let $r = x_2/x_1$, so we can plug into our previous definition:

$$\begin{aligned} \frac{d\ln[MRTS_{12}(\boldsymbol{x}(r))]}{d\ln[r]} &= \frac{d\ln[r]^{1-\rho}}{d\ln[r]} \\ &= (1-\rho)\frac{d\ln[r]}{d\ln[r]} \\ &= (1-\rho) \\ \implies \sigma &= \frac{1}{1-\rho} \end{aligned} \tag{7}$$

which is a constant, hence the term constant elasticity of substitution.

With the CES, the degree of substitutability between inputs is always the same, regardless of output or input proportions. The closer ρ is to unity, the larger the value of σ. When ρ is equal to 1, σ is infinite, and the production function is linear, implying full substitutability between inputs.

Other popular production functions can be seen as special cases of the CES form. In particular:

$$y = \left(\sum_{i=1}^{n} \alpha_i x_i^\rho \right)^{1/\rho}, \quad \text{with} \; \sum_{i=1}^{n} \alpha_i = 1 \tag{8}$$

is a CES with $\sigma_{ij} = 1/(1-\rho)$ for all $i \neq j$. It can be shown that as $\rho \to 0$, $\sigma \to 1$ and the CES reduces to the linear **homogeneous Cobb-Douglas** function:

$$y = \prod_{i=1}^{n} x_i^{\alpha_i} \tag{9}$$

On the other hand, as $\rho \to -\infty$, $\sigma \to 0$, giving the **Leontief** form as a limiting case, where:

$$y = \min\{x_1, ...x_n\} \tag{10}$$

All CES production functions (including the Cobb-Douglas and the Leontief) are members of the **linear homogenous production functions**. We have that among other things, linear homogenous production functions will always be concave functions.

Theorem 2.1. ***(Shephard) Homogeneous Production Functions are Concave****. Let $f(\boldsymbol{x})$ be a production function satisfying Assumption (2.1), and suppose it is homogeneous of degree $\alpha \in (0, 1]$. Then $f(\boldsymbol{x})$ is a concave function.*

3 Varying Proportions and Returns to Scale

Very often we want to know how input responds as the amounts of different inputs are varied. For example, **the short run**. The period of time in which at least one input is fixed. This limitation implies that output can only be varied by adjusting the amount a limited number of inputs. As these variations take place, the proportion at which fixed and free inputs are used also changes. How output responds in this situation is referred to as **Returns to variable proportions**.

In the **long run**, when the firm is free to vary all inputs. Describing how output responds to marginal changes can be done through the **returns to scale**. More specifically, returns to scale refers to how output responds when all inputs are varied in the same proportion.

Measures of returns to varying proportion include the **marginal product**: $MP_i(\boldsymbol{x}) \equiv f_i(\boldsymbol{x})$, and the **average product**: $AP_i(\boldsymbol{x}) \equiv f(\boldsymbol{x})/x_i$ for each input. The **output elasticity of input** i, measures the percentage response of output to a one percent change in input i, and it is given by: $\mu_i \equiv f_i(\boldsymbol{x})x_i/f(\boldsymbol{x}) = MP_i/AP_i$.

Note that all these measures are local. For a production function to have **globally constant, increasing, or decreasing returns to scale** according to the definition below:

Definition 3.1. (Global) Returns to Scale. A production function $f(\boldsymbol{x})$ has the property of globally:

1. Constant returns to scale if $f(t\boldsymbol{x}) = tf(\boldsymbol{x}) \ \forall \ t > 0, \boldsymbol{x}$.
2. Increasing returns to scale if $f(t\boldsymbol{x}) > tf(\boldsymbol{x}) \ \forall \ t > 1, \boldsymbol{x}$.
3. Decreasing returns to scale if $f(t\boldsymbol{x}) < tf(\boldsymbol{x}) \ \forall \ t > 1, \boldsymbol{x}$.

It's important to note from these definitions, that a production function has constant returns to scale if it is a (positive) linear homogeneous function. Also, that every homogenous function of degree greater (less) than one must have increasing (decreasing) returns to scale, though the opposite need not to hold.

Logically, we can also derive a local measure of returns to scale. This measure is defined at a point, and it informs about the instantaneous percent change in output that occurs with a one percent change of all inputs. It is known as the **elasticity of scale** or the overall **elasticity of output**:

Definition 3.2. (Local) Returns to scale. The elasticity of scale at point $\boldsymbol{x}$ is defined as:

$$\mu(\boldsymbol{x}) \equiv \lim_{t \to 1} \frac{d \ln[f(t\boldsymbol{x})]}{d \ln[t]} = \frac{\sum_i f_i(\boldsymbol{x})x_i}{f(\boldsymbol{x})} \tag{11}$$

Returns to scale are locally constant, increasing or decreasing as $\mu(\boldsymbol{x})$ is equal to, greater than, or less than one. The elasticity of scale and the output elasticities of the inputs are related as follow:

$$\mu(\boldsymbol{x}) = \sum_{i=1}^{n} \mu_i(\boldsymbol{x}) \tag{12}$$

4 Cost

We define the firm's cost as the expenditure associated with acquiring inputs to produce a given output. If the objective of the firm is to maximize profits, it is necessary to choose the least costly, or **cost-minimizing**, production plan for every level of output. In this case we will assume that the firm is perfectly competitive on its input markets, and therefore it will take prices as given.

Let $\boldsymbol{w} = (w_1, ..., w_n) \geq \mathbf{0}$ be a vector of prevailing market prices at which the firm can buy inputs $\boldsymbol{x} = (x_1, ..x_n)$. Because the firm is maximizing profits, it will choose to produce some level of output requiring the input vector to have the least possible cost:

Definition 4.1. The Cost Function. The cost function, defined for all input prices $\boldsymbol{w} >> \mathbf{0}$ and all output levels $y \in f(\mathbb{R}^n_{++})$ is the minimum value function:

$$c(\boldsymbol{w}, y) \equiv \min_{\boldsymbol{x} \in \mathbb{R}^n_{++}} \boldsymbol{w} \cdot \boldsymbol{x}, \quad \text{s.t.} \quad f(\boldsymbol{x}) \geq y \tag{13}$$

If $\boldsymbol{x}(\boldsymbol{w}, y)$ solves the cost-minimization problem, then:

$$c(\boldsymbol{w}, y) = \boldsymbol{w} \cdot \boldsymbol{x}(\boldsymbol{w}, y) \tag{14}$$

Let's examine the cost function a little further. Since $f(\boldsymbol{x})$ is strictly increasing, we know that the restriction will hold with equality, so we have:

$$c(\boldsymbol{w}, y) \equiv \min_{\boldsymbol{x} \in \mathbb{R}^n_{++}} \boldsymbol{w} \cdot \boldsymbol{x}, \quad \text{s.t.} \quad f(\boldsymbol{x}) = y \tag{15}$$

Let $\boldsymbol{x}^*$ denote the solution to (15). Assume $\boldsymbol{x} >> \mathbf{0}$, and that f is differentiable at $\boldsymbol{x}^*$, with $\nabla f(\boldsymbol{x}^*) >> \mathbf{0}$. By the Lagrange's theorem. there is a λ^* such that:

$$w_i = \lambda \frac{\partial f(\boldsymbol{x}^*)}{\partial x_i}, \quad i = 1, ..., n \tag{16}$$

Because $w_i > 0$, $i = 1, ..., n$, we can take the ratio of the ith element with respect to the jth element:

$$\frac{\partial f(\boldsymbol{x}^*)/\partial x_i}{\partial f(\boldsymbol{x}^*)/\partial x_j} = \frac{w_i}{w_j} \tag{17}$$

Thus, cost minimization implies that the marginal rate of substitution between any two inputs is equal to the ratio of their prices. The solution $\boldsymbol{x}^*(\boldsymbol{w}, y)$ is denoted as the firm's **conditional input demand**.

At this point the similitudes between the consumer theory should be apparent. Let's compare both the expenditure function and the cost functions:

$$e(\boldsymbol{p}, u) = \min_{\boldsymbol{x} \in \mathbb{R}^n_{++}} \boldsymbol{p} \cdot \boldsymbol{x}, \quad \text{s.t.} \quad u(\boldsymbol{x}) \geq u \qquad \text{(Expenditure Function)} \tag{18}$$

$$c(\boldsymbol{w}, y) = \min_{\boldsymbol{x} \in \mathbb{R}^n_{++}} \boldsymbol{w} \cdot \boldsymbol{x}, \quad \text{s.t.} \quad f(\boldsymbol{x}) \geq y \qquad \text{(Cost Function)} \tag{19}$$

At least mathematically, these two functions are identical. This equivalence implies that the properties of the expenditure function also hold for the cost function:

Theorem 4.1. ***Properties of the Cost Function***. *If f is continuous and strictly increasing, then $c(\boldsymbol{w}, y)$ is:*

1. *Zero when $y = 0$,*

2. *Continuous on its domain,*
3. *For all $\boldsymbol{w} >> \mathbf{0}$, strictly increasing and unbounded above in y,*
4. *Increasing in $\boldsymbol{w}$,*
5. *Homogenous of degree one in $\boldsymbol{w}$,*
6. *Concave in $\boldsymbol{w}$.*

   Moreover, if f is strictly quasiconcave we have:
7. *Shephard's lemma: $c(\boldsymbol{w}, y)$ is differentiable in $\boldsymbol{w}$ at $(\boldsymbol{w}^0, y^0)$ whenever $\boldsymbol{w} >> \mathbf{0}$, and*

$$\frac{\partial c(\boldsymbol{w}^0, y^0)}{\partial w_i} = x_i(\boldsymbol{w}^0, y^0), \quad i = 1, ...n \tag{20}$$

If we continue our comparison, it is natural to think that the input demand functions are analogous to the Hicksian demands of the consumer, and therefore they share the same properties as well:

Theorem 4.2. ***Properties of the Conditional Input Demands***. *Suppose that the production function satisfies Assumption (2.1) and that the associated cost function is twice continuously differentiable. Then:*

1. *$\boldsymbol{x}(\boldsymbol{w}, y)$ is homogenous of degree zero in $\boldsymbol{w}$,*
2. *The substitution matrix, defined and denoted*

$$\sigma(\boldsymbol{w}, y)^* = \begin{bmatrix} \frac{\partial x_1(\boldsymbol{w},y)}{\partial p_1} & \cdots & \frac{\partial x_1(\boldsymbol{w},y)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n(\boldsymbol{w},y)}{\partial p_1} & \cdots & \frac{\partial x_n(\boldsymbol{w},y)}{\partial p_n} \end{bmatrix} \tag{21}$$

is symmetric and negative semidefinite. In particular, the negative semidefiniteness implies that $\partial x_i(\boldsymbol{w}, y)/\partial w_i \leq 0 \; \forall \; i$.

Homogenous, or more generally homothetic[1] production technologies are common in theoretical and applied fields. The cost and conditional demands associated with these technologies have some special properties worth mentioning:

Theorem 4.3. ***Cost and Conditional Input Demands when Production is Homothetic***.

1. *When the production function satisfies Assumption (2.1) and is homothetic,*
   - (a) *The cost function is multiplicatively separable in input prices and output, and can be written as $c(\boldsymbol{w}, y) = h(y)c(\boldsymbol{w}, 1)$, where $h(y)$ is strictly increasing and $c(\boldsymbol{w}, 1)$ is the unit cost function, or the cost of 1 unit of output;*
   - (b) *The conditional input demands are multiplicatively separable in input prices and the output, and can be written $\boldsymbol{x}(\boldsymbol{w}, y) = h(y)\boldsymbol{x}(\boldsymbol{w}, 1)$, where $h'(y) > 0$ and $\boldsymbol{x}(\boldsymbol{w}, 1)$ is the conditional input demand for 1 unit of output.*

[1] A real-valued function h on $D \subset \mathbb{R}_n$ is called homothetic if it can be written in the form $g(f(x))$, where $g : \mathbb{R} \to \mathbb{R}$ is strictly increasing and $f : D \to \mathbb{R}$ is homogeneous of degree 1.

2. *when the production function is homogenous of degree* $\alpha > 0$*,*

   (a) $c(\boldsymbol{w}, y) = y^{1/\alpha} c(\boldsymbol{w}, 1)$

   (b) $\boldsymbol{x}(\boldsymbol{w}, y) = y^{1/\alpha} \boldsymbol{x}(\boldsymbol{w}, 1)$

Now, we can properly examined how the firm behaves in the short run when one or more input levels are fixed. Let us define the **short-run or restricted cost function** as:

Definition 4.2. The Short-Run, or Restricted, Cost Function. Let the production function be $f(\boldsymbol{z})$, where $\boldsymbol{z} \equiv (\boldsymbol{x}, \bar{\boldsymbol{x}})$. Suppose that $\boldsymbol{x}$ is a subvector of variable inputs, and $\bar{\boldsymbol{x}}$ is a subvector of fixed inputs. Let $\boldsymbol{w}$ and $\bar{\boldsymbol{w}}$ be the associated input prices for the variable and fixed inputs, respectively. The short-run, or restricted, total cost function is defined as:

$$sc(\boldsymbol{w}, \bar{\boldsymbol{w}}, y; \bar{\boldsymbol{x}}) \equiv \min_{\boldsymbol{x}} \boldsymbol{w} \cdot \boldsymbol{x} + \bar{\boldsymbol{w}} \cdot \bar{\boldsymbol{x}}, \quad s.t. \quad f(\boldsymbol{x}, \bar{\boldsymbol{w}}) \geq y \tag{22}$$

If $\boldsymbol{x}(\boldsymbol{w}, \bar{\boldsymbol{w}}, y; \bar{\boldsymbol{x}})$ solves the minimization problem, then

$$sc(\boldsymbol{w}, \bar{\boldsymbol{w}}, y; \bar{\boldsymbol{x}}) = \boldsymbol{w} \cdot \boldsymbol{x}(\boldsymbol{w}, \bar{\boldsymbol{w}}, y; \bar{\boldsymbol{x}}) + \bar{\boldsymbol{w}} \cdot \bar{\boldsymbol{x}} \tag{23}$$

The optimized cost of the variable inputs, $\boldsymbol{w} \cdot \boldsymbol{x}(\boldsymbol{w}, \bar{\boldsymbol{w}}, y; \bar{\boldsymbol{x}})$ is called **total variable cost**, while the cost of the fixed inputs, $\bar{\boldsymbol{w}} \cdot \bar{\boldsymbol{x}}$ is called **total fixed cost**.

Paying attention to the short-run cost, it becomes obvious that as the firm is not able to choose freely the combination of inputs, the long-run cost would never be greater than the short-run cost, as they can allocate optimally across inputs. That doesn't mean the firm will never choose to have the same level of inputs as in fixed scenario. Let $\bar{\boldsymbol{x}}(y)$ denote the optimal choice of the fixed inputs to minimize short-run cost of output, y, at the given input prices. Now it follows that:

$$c(\boldsymbol{w}, \bar{\boldsymbol{w}}, y) \equiv sc(\boldsymbol{w}, \bar{\boldsymbol{x}}, y; \bar{\boldsymbol{x}}(y)) \tag{24}$$

must hold for any y. Furthermore, because we have chosen the fixed inputs to minimize short-run costs, the optimal amounts of $\bar{\boldsymbol{x}}(y)$ must satisfy (identically) the first order conditions for a minimum:

$$\frac{\partial sc(\boldsymbol{w}, \bar{\boldsymbol{x}}, y; \bar{\boldsymbol{x}}(y))}{\partial \bar{\boldsymbol{x}}_i} \equiv 0 \tag{25}$$

For all inputs i. Now, differentiating (24), and using (25), we have:

$$\frac{dc(\boldsymbol{w}, \bar{\boldsymbol{w}}, y)}{dy} = \frac{\partial sc(\boldsymbol{w}, \bar{\boldsymbol{x}}, y; \bar{\boldsymbol{x}}(y))}{\partial y} \tag{26}$$

To clarify this identity we have to recall the process we followed until now. First, the short-run cost-minimization problem involves more constraints than the long-run, so it must be that $sc(\boldsymbol{w}, \bar{\boldsymbol{x}}, y; \bar{\boldsymbol{x}}(y)) \geq c(\boldsymbol{w}, \bar{\boldsymbol{w}}, y)$ for all levels of output and levels of the fixed inputs. Second, for every level of output, equation (24) says that the short-run and long-run costs will coincide for some short-run cost function associated with some level of the fixed inputs. Finally, equation (26) is telling us that the slope of this short-run function will be equal to the slope of the long-run function in the cost-output plane. So, if we have two functions taking the same value and with equal slopes, they are tangent, which leads to the proposition:

Proposition 4.1. *The long-run total cost curve is the lower envelope of the entire family of short-run total curves.*

5 Duality in Production

So far, the similitude between the consumer expenditure-minimization problem and the firm cost-minimization problem arise naturally. In the same way, it is intuitive to think that as there is a duality between utility and expenditure, there is also a duality between production and costs.

The principles under which this relationship is regulated are identical to the consumer case. If the original production function is quasiconcave, the derived production function would be identical. If the original production function is not quasiconcave, the derived cost function is a concavication of it. Moreover, any function with all the properties of a production function generates some production cost for which it is the production function of itself.

The previous statement is a powerful one as it allows to backup and estimate a firm's cost function using its output and the inputs market prices. Formally:

Theorem 5.1. ***Recovering a Production Function from a Cost Function.*** *Let $c : \mathbb{R}^n_{++} \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfy properties 1 to 7 of a cost function given in Theorem (4.1). Then the function $f : \mathbb{R}^n_+ \to \mathbb{R}_+$ defined by*

$$f(\boldsymbol{w}) \equiv \max\{y \geq 0 \| \boldsymbol{w} \cdot \boldsymbol{x} \geq c(\boldsymbol{w}, y),\ \forall\ \boldsymbol{w} >> \mathbf{0}\} \tag{27}$$

is increasing, unbounded above, and quasiconcave. Moreover, the cost function generated by f is c.

We can also state an integrability-type theorem for input demand that allows to establish if quantities were optimally chosen by the firm. That is, those demands will be consistent with cost minimization at each output level, if and only if there is a cost function c satisfying:

$$\frac{\partial c(\boldsymbol{x}, y)}{\partial w_i} = x_i(\boldsymbol{w}, y), \quad i = 1, ..., n \tag{28}$$

To evaluate such question we can make use of the following theorem:

Theorem 5.2. *A continuously differentiable function $\boldsymbol{x}(\boldsymbol{w}, y)$ mapping $\mathbb{R}^n_{++} \times \mathbb{R}_+$ into $\mathbb{R}^n_+$ is the conditional input demand function generated by some strictly increasing, quasiconcave production function if and only if it is homogeneous of degree zero in $\boldsymbol{w}$, its substitution matrix, whose ijth entry is $\partial x_i(\boldsymbol{w}, y)/\partial w_j$ is symmetric and negative semidefinite, and $\boldsymbol{w} \cdot \boldsymbol{x}(\boldsymbol{w}, y)$ is strictly increasing in y.*

6 The Competitive Firm

So far we have examined the behavior when the firm is a perfect competitor in the input market. Now we will turn to the case when the firm is a perfect competitor in both the input and output markets. The competitive firm sees the market price for its product, assumes it will remain the same regardless of how much or how little it sells, and makes its plans accordingly. Such a firm is thus a price taker on both output and input markets.

6.1 Profit Maximization

Profit is the difference between revenue from seeing output and total cost of acquiring all the factors to produce it. The firm can sell each unit of output at price p, and its revenue is given by $R(y) = py$. If the firm aims for output level y^0, if $\boldsymbol{x}^0$ is a feasible vector of inputs to produce y^0, and $\boldsymbol{w}$ is the vector of factor prices, the cost of using $\boldsymbol{x}^0$ to produce y is imply $\boldsymbol{w} \cdot \boldsymbol{x}^0$. Profits in this case, are given by $py^0 - \boldsymbol{w} \cdot \boldsymbol{x}^0$. It follows that the firm must decide which level of output, and what combination of inputs are the best way to produce.

With the ultimate objective of maximizing profits, we let the firm solve the problem:

$$\max_{(\boldsymbol{x},y)\geq \boldsymbol{0}} py - \boldsymbol{w} \cdot \boldsymbol{x}, \quad s.t. \quad f(\boldsymbol{x}) \geq y \tag{29}$$

with f as a function satisfying Assumption (2.1). Similarly, we can replace the inequality in the constraint by an equality as the production function is strictly increasing. Consequently, since $y = f(\boldsymbol{x})$, we may rewrite the objective function as:

$$\max_{\boldsymbol{x}\in\mathbb{R}^n_+} pf(\boldsymbol{x}) - \boldsymbol{w} \cdot \boldsymbol{x}. \tag{30}$$

If the problem has an interior solution $\boldsymbol{x}^* >> 0$, then the first order conditions require:

$$p\frac{\partial f(\boldsymbol{w}^*)}{\partial x_i} = w_i, \quad \text{for every } i = 1, ..., n. \tag{31}$$

The term on the left hand side, the product of the output price with the marginal product of input i, is often referred to as **marginal revenue product** of input i. It gives the rate at which revenue increases per additional unit of input i. At the optimum, the marginal revenue product must equal the cost per unit of input i, namely , w_i. Assuming that all w_i are positive, we can use the first order condition to yield some information about the ratios:

$$\frac{\partial f(\boldsymbol{w}^*)/\partial x_i}{\partial f(\boldsymbol{w}^*)/\partial x_j} = \frac{w_i}{w_j}, \quad \forall\, i, j, \tag{32}$$

In other words, the MRTS between any two inputs has to equal the ratio of their prices. Recall, however, that this condition is the same condition we derived in Equation (17). The condition for maximizing profit is the same as the one to minimize cost; hence the golden rule for profit maximization is to equal marginal revenue to marginal cost.

6.2 The Profit Function

When f satisfies Assumption (2.1) and is also strictly concave, solutions to the maximization problem, if they exist, will be unique for each price vector $(p, \boldsymbol{w})$. The optimal choice of output $y^* \equiv y(p, \boldsymbol{w})$ is called the firm's **output supply function**, and the optimal choice of inputs, $\boldsymbol{x}^* \equiv \boldsymbol{x}(p, \boldsymbol{w})$, gives the vector of firm **input demand function**. The latter are full-fledged demand functions because, unlike the conditional input demands that depend partly on output, these input demands achieve the ultimate goal of the firm; they maximize profit. Formally:

Definition 6.1. The Profit Function. the firm's profit function depends only on input and output prices, and is defined as the maximum-value function,

$$\pi(p, \boldsymbol{w}) \equiv \max_{(\boldsymbol{w},y)\geq \boldsymbol{0}} py - \boldsymbol{w} \cdot \boldsymbol{x}, \quad s.t. \quad f(\boldsymbol{x}) \geq y. \tag{33}$$

When the profit function is well defined, it has several useful properties, that will seem analog to the consumer utility function:

Theorem 6.1. *If f satisfies Assumption (2.1), then for $p \geq 0$ and $\boldsymbol{w} \geq \boldsymbol{0}$, the profit function $\pi(p, \boldsymbol{w})$, where well-defined, is continuous and*

1. *Increasing in p,*
2. *Decreasing in $\boldsymbol{w}$,*
3. *Homogenous of degree one in $(p, \boldsymbol{w})$,*
4. *Convex in $(p, \boldsymbol{w})$,*
5. *Differentiable in $(p, \boldsymbol{w}) >> \mathbf{0}$. Moreover, under the additional assumption that f is strictly concave (Hotelling's lemma),*

$$\frac{\partial \pi(p, \boldsymbol{w})}{\partial p} = y(p, \boldsymbol{w}), \quad \text{and} \quad \frac{-\partial \pi(p, \boldsymbol{w})}{\partial w_i} = x_i(p, \boldsymbol{w}), \quad i = 1, ..., n. \tag{34}$$

Looking closely at Hotelling's lemma, we see that output supply and input demand can be obtained directly from the profit function by means of differentiating. From this property we can deduce some restrictions on firm behavior that derive from the profit maximization assumption:

Theorem 6.2. *Suppose that f is a strictly concave production function satisfying Assumption (2.1) and that its associated profit function, $\pi(\boldsymbol{p}, y)$, is twice continuously differentiable. Then, for all $p > 0$ and $\boldsymbol{w} >> 0$ where it is well defined:*

1. *Homogeneity of degree zero:*

$$y(tp, t\boldsymbol{w}) = y(p, \boldsymbol{w}) \; \forall \; t > 0, \tag{35}$$
$$x_i(tp, t\boldsymbol{w}) = x_i(p, \boldsymbol{w}) \; \forall \; t > 0 \; \wedge \; i = 1, ..., n \tag{36}$$

2. *Own-price effects:*

$$\frac{\partial y(p, \boldsymbol{w})}{\partial p} \geq 0 \tag{37}$$
$$\frac{\partial x_i(p, \boldsymbol{w})}{\partial w_i} \leq 0, \; \forall \; i = 1, ..., n \tag{38}$$

3. *The substitution matrix*

$$\begin{pmatrix} \frac{\partial y(p,\boldsymbol{w})}{\partial p} & \frac{\partial y(p,\boldsymbol{w})}{\partial w_1} & \cdots & \frac{\partial y(p,\boldsymbol{w})}{\partial w_n} \\ \frac{-\partial x_1(p,\boldsymbol{w})}{\partial p} & \frac{-\partial x_1(p,\boldsymbol{w})}{\partial w_1} & \cdots & \frac{-\partial x_1(p,\boldsymbol{w})}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-\partial x_n(p,\boldsymbol{w})}{\partial p} & \frac{-\partial x_n(p,\boldsymbol{w})}{\partial w_1} & \cdots & \frac{-\partial x_n(p,\boldsymbol{w})}{\partial w_n} \end{pmatrix} \tag{39}$$

   is symmetric and positive semidefinite.

Therefore, analog to the case of consumer demand and conditional input demand, there is an integrability theorem for input demand and output supply.

Because we have supposed the firm is free to choose its output and all input levels as it sees fit, the profit function we have defined so far is really best thought of as the long-run profit function. As we did for the cost function then, we can construct a short-run or restricted profit function to describe firm behavior when some of its inputs are variable and some are fixed.

Theorem 6.3. ***The Short-Run, or Restricted, Profit Function***. *Suppose that $f : \mathbb{R}^n_+ \to \mathbb{R}_+$ is strictly concave and satisfies Assumption (2.1). For $k < n$, let $\bar{\boldsymbol{x}} \in \mathbb{R}^k_+$ be a subvector of fixed inputs and consider $f(\boldsymbol{x}, \bar{\boldsymbol{x}})$ as a function of the subvector of variable inputs $\boldsymbol{x} \in \mathbb{R}^{n-k}_+$. Let $\boldsymbol{w}$ and $\bar{\boldsymbol{w}}$ be the associated input prices for variable and fixed inputs, respectively. The* ***short-run, or restricted, profit function*** *is defined as:*

$$\pi(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}) \equiv \max_{y,\boldsymbol{x}} py - \boldsymbol{w} \cdot \boldsymbol{x} - \bar{\boldsymbol{w}} \cdot \bar{\boldsymbol{x}} \tag{40}$$

The solutions to $y(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}})$ and $\boldsymbol{x}(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}})$ are called the short-run, or restricted , output supply and variable input demand functions, respectively.

For all $p > 0$ and $\boldsymbol{w} >> \boldsymbol{0}$, $\pi(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}})$ is well-defined if it is continuous in p and $\boldsymbol{w}$, increasing in p, decreasing in $\boldsymbol{w}$, and convex in $(p, \boldsymbol{w})$. If $\pi(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}})$ is twice continuously differentiable, $y(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}})$ and $\boldsymbol{x}(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}})$ posses al three properties listed in the ***Second Welfare Theorem****.*

To finish our treatment of the firm's behavior, let's focus on supply. We can subsume the input choice problem into the short-run cost function and express short-run profits as:

$$\pi(p, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}) = \max_{y} py - sc(y, \boldsymbol{w}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}) \tag{41}$$

The first order conditions imply the for optimal output $y^* > 0$,

$$p = \frac{dsc(y^*)}{dy} \tag{42}$$

or that price equals (short-run) marginal cost.

The last question that remains is: is marginal revenue equal to marginal cost a foolproof rule? Let the price equal to marginal cost at some $y^1 > 0$. The short-run total cost can be expressed as the sum of the total variable cost, $tvc(y)$, and the total fixed cost, tfc. The former is the optimized cost of variable inputs and the latter is the cost of fixed inputs. Short-run profits can then be written as:

$$\pi^1 = py^1 - tvc(y) - tfc \tag{43}$$

What if π^1 is negative? Is it still the best decision to produce y^1 even though it entails a loss? From the previous condition, we know that the profit maximizing (or loss minimizing) output is y^1, where price equals marginal cost. However, the firm can always shut down and produce nothing. If it produces $y = 0$, it will have no revenues, and no variable costs, although there will be fixed costs. The associated losses of shutting down are given by:

$$\pi^0 = p \cdot 0 - tvc(0) - tfc = -tfc < 0 \tag{44}$$

So a profit maximizer will choose between producing $y^1 > 0$ at a loss, or producing $y = 0$ at a loss according to which gives greater profit (smaller loss). The firm produces y^1 if and only if $\pi^1 - \pi^0 \geq 0$ or:

$$py^1 - tvc(y^1) \geq 0 \tag{45}$$

which is equivalent to:

$$p \geq \frac{tvc(y^1)}{y^1} \equiv avc(y^1) \tag{46}$$

We now have everything we need to describe short-run profits maximizing behavior. If the firm produces a positive amount of output, then it will produce the amount where price equals marginal cost (and marginal cost is non-decreasing), and price is not below the average variable cost at that level of output. If price is less than the average variable cost where price equals marginal cost, the firm will shut down and produce no output.