Notes for Consumer and Producer Theory

Francis J. DiTraglia
The University of California, San Diego

June 19, 2008

Abstract

I wrote the following notes in preparation for the Microeconomics Qualifying Exam at the University of California, San Diego. The content is based principally on selected chapters of Hal Varian's *Microeconomic Analysis, Third Edition*, with additional material from Varian's *Intermediate Microeconomics*, Sundaram's *A First Course in Optimization Theory*, and my lecture notes from Econ 200A at UCSD. Any errors and omissions are my own.

Contents

Chapter 1

A Little Math

This chapter reviews a few useful definitions and techniques that will be used later in the notes. Much of this material is drawn from Sundaram's *A First Course in Optimization Theory*

1.1 Convexity and Concavity

Definition 1.1.1 (Convex Set). *A set $\mathcal{D} \subset \mathbb{R}^n$ is called convex if for any two points $x, y \in \mathcal{D}$ and all $\lambda \in (0,1)$, we have $\lambda x + (1-\lambda)y \in \mathcal{D}$*

Remark 1.1.1. *There is no such thing as a concave set.*

Definition 1.1.2 (Concave and Convex Functions). *Let $\mathcal{D}$ be a convex subset of $\mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called concave on the set $\mathcal{D}$ if for any $x, y \in \mathcal{D}$ and any $\lambda \in (0,1)$,*

$$f[\lambda x + (1-\lambda)y] \leq \lambda f(x) + (1\lambda)f(y).$$

A function is called convex on $\mathcal{D}$ if for any $x, y \in \mathcal{D}$ and any $\lambda \in (0,1)$,

$$f[\lambda x + (1-\lambda)y] \geq \lambda f(x) + (1\lambda)f(y).$$

Definition 1.1.3 (Alternative Definitions). *A function is concave if and only if the area lying below its graph (the subgraph) is convex, e.g. $f(x) = -x^2$. A function is convex if and only if the area lying above its graph (the epigraph) is convex, e.g. $f(x) = x^2$.*

Note that convexity and concavity are neither exhaustive nor mutually exclusive. It is possible for a function to be both concave and convex (e.g. any affine function) or neither concave nor convex (e.g. $f(x) = x^3$).

Definition 1.1.4 (Strict Convexity and Concavity). *A function is strictly concave if it satisfies the definition of concavity with a strict rather than* *weak inequality. Similarly, a function is strictly convex if it satisfies the definition of convexity with a strict rather than weak inequality.*

Note that any affine function is both convex and concave, but neither strictly convex nor strictly concave. Convexity and concavity as defined here are sometimes called *weak* convexity and concavity to distinguish them from *strict* concavity.

Theorem 1.1.1. *Let $\mathcal{D}$ be a convex subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be a concave function. Then both of the following hold:*

(a) Any local maximum of f is a global maximum.

(b) The set of maximizers of f on $\mathcal{D}$ is either empty or convex.

Theorem 1.1.2 (Convexity, Concavity and the Second Derivative)**.** *Let $f : \mathcal{D} \to \mathbb{R}$ be twice continuously differentiable, where $\mathcal{D} \subseteq \mathbb{R}^n$ is open and convex. Then:*

(a) A function f is concave on $\mathcal{D}$ if and only if the matrix $D^2 f(x)$ of second derivatives is negative semidefinite for all $x \in \mathcal{D}$.

(b) A function f is convex on $\mathcal{D}$ if and only if the matrix $D^2 f(x)$ of second derivatives is positive semidefinite for all $x \in \mathcal{D}$.

(c)

(d)

1.2 Quasi-concavity and Quasi-convexity

Definition 1.2.1 (Upper and Lower Contour Sets)**.** *Let $\mathcal{D}$ be a convex subset of $\mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, and $a \in \mathcal{D}$. Then the set $U_f(a) = \{x \in \mathcal{D} | f(x) \geq a\}$ is called the upper contour set of f at a, while $L_f(a) = \{x \in \mathcal{D} | f(x) \leq a\}$ is called the lower contour set of f at a.*

Definition 1.2.2 (Quasi-concave and Quasi-convex Functions)**.** *A function f is quasi-concave if its upper contour set $U_f(a)$ is convex for all a. It is called quasi-convex if its lower contour set $L_f(a)$ is convex for all a.*

Definition 1.2.3 (Alternative Definitions)**.** *A function f is quasi-concave on $\mathcal{D}$ if and only if for any $x, y \in \mathcal{D}$ and any $\lambda \in (0, 1)$,*

$$f[\lambda x + (1 - \lambda) y] \geq \min\{f(x), f(y)\}.$$

A function is called quasi-convex on $\mathcal{D}$ if and only if for any $x, y \in \mathcal{D}$ and any $\lambda \in (0, 1)$,

$$f[\lambda x + (1 - \lambda) y] \leq \max\{f(x), f(y)\}.$$

Definition 1.2.4 (Strict Quasi-concavity and Quasi-convexity). *A function is strictly quasi-concave if it satisfies either of the definitions of quasi-concavity given above where a strict inequality replaces the weak inequality. A function is strictly quasi-convex if it satisfies either definition of quasi convexity given above with a strict rather than weak inequality.*

Note that quasi-convex and quasi-concave functions as defined here are sometimes called *weakly* quasi-concave or *weakly* quasi-convex.

Theorem 1.2.1. *Concavity implies quasi-concavity, and convexity implies quasi-convexity.*

Theorem 1.2.2. *Any monotonic transformation of a quasi-concave function results in another quasi-concave function. In particular, any monotonic transformation of a concave function results in a quasi-convex function.*

Note that concavity is *not* preserved under monotonic transformations, but quasi-concavity is preserved. Further, there exist quasi-concave functions that *cannot* be obtained by taking a monotonic transformation of some concave function. **Note: unlike concave functions, quasi-concave functions can have a local maximum that is *not* a global maximum.**

1.3 Homogeneous Functions

Definition 1.3.1 (Homogeneous Function). *A function $f : \mathbb{R}^n_+ \to \mathbb{R}$ is homogeneous of degree k if $f(\lambda \mathbf{x}) = \lambda^k f(\mathbf{x})$ for all $\lambda > 0$.*

Theorem 1.3.1 (Euler's Theorem). *If f is differentiable and homogeneous of degree k, then:*

$$\nabla f(\mathbf{x}) \cdot \mathbf{x} = k f(\mathbf{x})$$

In particular, if f is homogeneous of degree 1:

$$f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i} x_i.$$

Theorem 1.3.2 (Derivative of Homogeneous Function). *If a function $f(\mathbf{x})$ is homogeneous of degree $k \geq 1$, then $\partial f / \partial x_i$ is homogeneous of degree $k-1$.*

Definition 1.3.2 (Homothetic Function). *A function f is called homothetic if it is a positive, strictly increasing transformation of another function that is homogeneous of degree 1, that is, if $f(\mathbf{x}) = g(h(\mathbf{x}))$ where $h(\mathbf{x})$ is homogeneous of degree 1, and $g(\cdot)$ is positive and monotonic.*

Note: perfect complements, perfect substitutes, and Cobb-Douglas preferences are represented by homothetic utility functions.

Theorem 1.3.3 (Homogeneous and Homothetic Functions)**.** *The slopes of the level surfaces of any homogeneous or homothetic function are constant along rays through the origin. Formally, homogeneous and homothetic functions satisfy:*

$$\frac{\partial f(\lambda \mathbf{x})}{\partial x_i} \Big/ \frac{\partial f(\lambda \mathbf{x})}{\partial x_j} = \frac{\partial f(\mathbf{x})}{\partial x_i} \Big/ \frac{\partial f(\mathbf{x})}{\partial x_j}$$

for any $\lambda > 0$.

Chapter 2

Technology

In this chapter we study the production possibilities of the firm, the **technologically feasible** input and output combinations. Unless otherwise specified, there are n inputs and outputs.

2.1 Measurement of Inputs and Outputs

We usually think of inputs and outputs as **flows**.

2.2 Specification of Technology

Definition 2.2.1 (Net Output). *Let y_j^i denote the firm's input of good j, y_j^o its output. Then $y_j = y_j^o - y_j^i$ is its net output.*

Definition 2.2.2 (Production Plan). *A production plan is a list of net outputs of various goods, represented as a vector $\mathbf{y} \in \mathbb{R}^n$ with negative elements representing net inputs, positive elements net outputs.*

Definition 2.2.3 (Production Possibilities Set). *The set Y of all technologically feasible production plans is the Production Possibilities Set.*

Definition 2.2.4 (Short-Run Production Possibilities). *In the short run, some inputs may be fixed. Let $\mathbf{z} \in \mathbb{R}^n$ be a list of maximum inputs and outputs in the short run. Then $Y(\mathbf{z}) = \{\mathbf{y} \in Y : y_n \leq z_n, \forall n\}$ is the restricted, or short-run production possibilities set.*

Definition 2.2.5 (Input Requirement Set). *Consider a firm with one output y and input vector $\mathbf{x}$ so that $(y, -\mathbf{x})$ is the net output vector. The set $V(y) = \{\mathbf{x} \in \mathbb{R}^n_+ : (y, -\mathbf{x}) \in Y\}$ is called the input requirement set. It is the set of all input bundles that produce at least y units of output.*

Definition 2.2.6 (Isoquant). *An isoquant is the set of all input bundles that produce exactly y units of output: $Q(y) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in V(y), \mathbf{x} \notin V(y'), \forall y' > y\}$.*

Definition 2.2.7 (Production Function)**.** *Suppose a firm has only one output. Then the production function* $f(\mathbf{x})$ *is the set of all* $y \in \mathbb{R}$ *such that* y *is the maximum output associated with* $-\mathbf{x}$ *in* Y*.*

2.3 Some Properties of Technology

Definition 2.3.1 (Monotonicity)**.** *If a certain input bundle* $\mathbf{x}$ *is sufficient to produce output level* y*, and some other input bundle* $\mathbf{x}'$ *has more of each input than* $\mathbf{x}$*, then* $\mathbf{x}'$ *is also sufficient to produce* y*. That is, if* $\mathbf{x} \in V(y)$ *and* $x'_k \geq x_k$ *for all* k*, then* $\mathbf{x}' \in V(y)$*.*

Definition 2.3.2 (Convexity of Input Requirement Sets)**.** *The input requirement set* $V(y)$ *is said to be convex if, for any* $\mathbf{x}, \mathbf{x}' \in V(y)$*, and for any* $\lambda \in [0,1]$*,* $\lambda\mathbf{x} + (1-\lambda)\mathbf{x}' \in V(y)$*.*

Definition 2.3.3 (Convextity of Production Set)**.** *The production set* Y *is said to be convex if for any* $\mathbf{y}, \mathbf{y}' \in Y$ *and any* $\lambda \in [0,1]$*,* $\lambda\mathbf{y} + (1-\lambda)\mathbf{y}' \in Y$

NOTE: Convexity of the production set is a *much* stronger assumption than convexity of the input requirement set. A convex production set **rules out scale economies** (including start-up costs).

Proposition 2.3.1. *If the production set* Y *is convex, then so is the associated input requirement set* $V(y)$*. The converse is not true.*

Proposition 2.3.2. *The input requirement set* $V(y)$ *is convex if and only if the production function* $f(\mathbf{x})$ *is a quasiconcave function.*

Definition 2.3.4 (Regular Technologies)**.** *Technology is said to be regular if* $V(y)$ *is a closed, nonempty set for all* $y \geq 0$*.*

As a consequence of the definition, if technology is regular, there it is possible to produce *any* non-negative level of output.

2.4 The Technical Rate of Substitution

The technical rate of substitution $TRS_{1,2}$ gives the amount of input two required to keep output constant in the face of a one unit decrease in input one. Thus, is is the slope of the isoquant evaluated at a particular point:

$$TRS_{i,j} = \frac{\partial f(\mathbf{x})}{\partial x_i} / \frac{\partial f(\mathbf{x})}{\partial x_j}$$

The derivation of this result is the same as the derivation of the MRS given in Chapter 7 of these notes.

2.5 Returns to Scale

Definition 2.5.1 (Constant Returns to Scale)**.** *A technology is said to exhibit constant returns to scale if any of the following properties are satisfied:*

1. *If* $\mathbf{y} \in Y$, *then* $\lambda \mathbf{y} \in Y$ *for any* $\lambda \geq 0$.
2. *If* $\mathbf{x} \in V(y)$, *then* $\lambda \mathbf{x} \in V(y)$ *for any* $\lambda \geq 0$.
3. *The production function is homogeneous of degree one, that is* $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ *for any* $\lambda \geq 0$.

For constant returns to scale technology, doubling all of the inputs exactly doubles output.

Definition 2.5.2 (Increasing Returns to Scale)**.** *A technology is said to exhibit increasing returns to scale if* $f(\lambda \mathbf{x}) > \lambda f(\mathbf{x})$ *for any* $\lambda \geq 0$.

For increasing returns to scale technology, doubling all of the inputs more than doubles output.

Definition 2.5.3 (Decreasing Returns to Scale)**.** *A technology is said to exhibit decreasing returns to scale if* $f(\lambda \mathbf{x}) < \lambda f(\mathbf{x})$ *for any* $\lambda \geq 0$.

For decreasing returns to scale technology, doubling all of the inputs less than doubles output.

Chapter 3

Profit Maximization

3.1 Profit Maximization

Let **p** be the vector or prices for the inputs and outputs of a price-taking firm. Then the profit maximization problem is given by:

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

subject to $\mathbf{y} \in Y$. Since outputs are measured as positive numbers, and inputs as negative numbers, **the objective function is revenues minus costs**. The optimal value function for this maximization problem, $\pi(\mathbf{p})$, which gives profits as a function of input and output prices, is called the **profit function**.

The **short-run**, or **restricted** profit function is give by:

$$\pi(\mathbf{p}, \mathbf{z}) = \max_{\mathbf{y}} \mathbf{p} \cdot \mathbf{y}$$

subject to $\mathbf{y} \in Y(\mathbf{z})$.

If the firm produces a single output, the profit function is:

$$\pi(p, \mathbf{w}) = \max \ p \cdot f(\mathbf{x}) - \mathbf{w} \cdot \mathbf{x}$$

where p is the price of the firm's output, $\mathbf{x}$ is its vector of inputs, $f(\cdot)$ is the production function, and $\mathbf{w}$ the vector of input prices. Note that this is an *unconstrained* problem since the constraint is already contained in the production function. The first order conditions for this problem are:

$$p \cdot \frac{\partial f(\mathbf{x})}{\partial x_i} = w_i$$

for all $i = 1, 2, \ldots$. That is, the firm chooses the level of each input so that the *marginal value product* of that input equals the price of the input. The second order condition requires that the Hessian matrix of the production function be negative semi-definite at the optimal point (global concavity is sufficient although not necessary).

The solution function to the profit maximization problem, the function that gives the optimal choice of inputs given prices, is called the **factor demand curve** and is denoted $\mathbf{x}(p, \mathbf{w})$. The production function evaluated at the optimal input choice $y(p, \mathbf{w}) = f(\mathbf{x}(p, \mathbf{w}))$ is called the **supply function** of the firm.

The first order condition can be interpreted graphically provided the firm uses only one input x and produces only one output y. In this case, profits are given by $\Pi = p \cdot y - w \cdot x$. Rearranging, we can write the *isoprofit line*, the line of all input-output combinations that give the same profit level $\bar{\Pi}$, as

$$y = \frac{\bar{\Pi}}{p} + \frac{w}{p}x$$

This is a line with slope w/p in x-y space. The first order condition is satisfied at the isoprofit that is *tangent* to the production function.

3.2 Difficulties

We usually assume that the factor demand function and supply function are well-behaved, but in certain cases they aren't. The following are common problems than can arise.

1. The production function may not be differentiable, e.g. Leontief technology. In this case we cannot take first order conditions.

2. The first order conditions given above assume an interior solution, i.e. that the firm uses a *positive* amount of each input. This will not be the case, for example, if technology is linear. In general, the Kuhn-Tucker conditions are:
$$p \cdot \frac{\partial f(\mathbf{x})}{\partial x_i} \leq w_i$$
for every $i = 1, 2, \ldots$, with equality if $x_i > 0$.

3. A profit maximizing plan *might not exist*. Whenever technology exhibits constant returns to scale, if there is a point at which profits are positive, the firm can increase its profits be increasing its scale. Hence, the only non-trivial profit-maximizing position for constant returns to scale is zero profits for a positive level of output. In this case, the firm is indifferent about its scale.

4. The profit maximizing production plan might not be unique, as in the case of constant returns to scale just mentioned.

3.3 Properties of Firm Demand and Supply

Proposition 3.3.1. *Factor demand is homogeneous of degree zero in input and output prices, that is* $\mathbf{x}(\lambda p, \lambda \mathbf{w}) = \mathbf{x}(p, \mathbf{w})$ *for any* $\lambda > 0$*.*

Proof. By definition $\mathbf{x}(p, \mathbf{x})$ is the profit-maximizing input bundle at output price p and input prices $\mathbf{w}$. But if $\mathbf{x}(p, \mathbf{x})$ maximizes $p \cdot f(\mathbf{x}) - \mathbf{w} \cdot \mathbf{x}$, it also maximizes $\lambda(p \cdot f(\mathbf{x}) - \mathbf{w} \cdot \mathbf{x})$. □

Theorem 3.3.1. *If the firm is profit maximizing,* $\mathbf{\Delta p} \cdot \mathbf{\Delta y} \geq 0$*.*

Proof. Follows from the weak axiom of cost minimization. □

By the above theorem, if the price of a firm's output good increases with all other things held constant, the firm's supply of this good *cannot* decrease. Similarly, if the price of a firm's input good increases with all other prices held constant, the firm's *demand* for this input cannot decrease.

Chapter 4

Profit Function

Recall that the profit function gives the maximum profits a firm can earn as a function of the vector of prices of net outputs:

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

such that $\mathbf{y}$ is a production plan in the production set Y. The important feature of the profit function is that it is the optimal value function for a *linear* maximization problem.

4.1 Properties of the Profit Function

The following important facts follow directly from the assumption that firms maximize profits. No regularity conditions, such as convexity or monotonicity, are needed.

Theorem 4.1.1 (Properties of Profit Function)**.** *The profit function is:*

(a) Non-decreasing in output prices, non-increasing in input prices

(b) Homogeneous of degree 1 in $\mathbf{p}$*, i.e.* $\pi(t\mathbf{p}) = t\pi(\mathbf{p})$

(c) Convex in $\mathbf{p}$

(d) Continuous in $\mathbf{p}$ *whenever* π *is well-defined and all prices are non-zero.*

Proof. Note that this proof uses only the assumption of profit maximization, and makes no assumptions about the properties of technology:

(a) Let $\mathbf{y}$ be the profit-maximizing production plan given prices $\mathbf{p}$, and $\mathbf{y}'$ be the profit-maximizing production plan given prices $\mathbf{p}'$. For a production plan to be profit maximizing given a set of prices, there cannot exist any alternative production plan that gives higher profits at the given prices. Hence, $\mathbf{p}' \cdot \mathbf{y}' \geq \mathbf{p}' \cdot \mathbf{y}$ and $\mathbf{p}' \cdot \mathbf{y}' \geq \mathbf{p}' \cdot \mathbf{y}$

(b)

(c)

(d) The proof of continuity follows from the theorem of the maximum.

□

Example: Price Stabilization

Suppose a competitive firm faces random prices: p_1 with probability q, p_2 with probability $1 - q$. Consider the stable price $\bar{p} = qp_1 + (1 - q)p_2$. Does the firm prefer the stable price or the random prices? We need to compare expected profits under random prices to profits at price $\bar{p}$. Since the profit function is convex in prices:

$$\mathbb{E}[\pi] = q\pi(p_1) + (1 - q)\pi(p_2) \geq \pi(qp_1 + (1 - q)p_2) = \pi(\bar{p})$$

Hence, the firm prefers random prices. It produces more when output prices are high, less when they are low.

4.2 Firm Supply and Demand Functions

Given a firm's net supply function $\mathbf{y}(\mathbf{p})$, its profit function is simply

$$\pi(\mathbf{p}) = \mathbf{p} \cdot \mathbf{y}(\mathbf{p}).$$

We can also go in the opposite direction using Hotelling's Lemma.

Theorem 4.2.1. *(Hotelling's Lemma) Let $y_i(\mathbf{p})$ be the firm's net supply function for good i. Then, for $p_i > 0$ and assuming the derivative esists:*

$$y_i(\mathbf{p}) = \frac{\partial \pi(\mathbf{p})}{\partial p_i}.$$

Proof. Use the envelope theorem. □

The intuition is as follows: there are two effects from an increase in the price of an output. First, the direct effect means that even if the firm continued making the same quantity as before, profits will be higher. Second, the indirect effect means that due to the price change the firm will re-optimize, adjusting its quantity. For an infinitesimal change in price, the indirect effect is zero by the envelope theorem.

4.3 Comparative Statics from Profit

Hotelling's Lemma allows us to derive properties of the net supply functions from the profit function. For example, since profit is homogeneous of degree 1 in prices, its partial derivative with respect to p_i (the net supply function of good i) is homogeneous of degree zero in prices.

Similarly, since profit is convex in prices, the matrix of its second derivatives is P.S.D. Since this matrix is the same as the matrix of *first* derivatives of the net supply functions, the substitution matrix, it too is P.S.D. This means that $\partial y_i/\partial p_i \geq 0$ for all i – when the price of a net input goes up, you employ no more of it and when the price of a net output goes up you produce no less of it.

Chapter 5

Cost Minimization

Cost minimization gives a way to study the supply behavior of firms in competitive *and* non-competitive output markets.

5.1 Calculus of Cost Minimization

In the cost minimization problem, the firm chooses the cheapest way to produce a specified amount of output:

$$\min_{\mathbf{x}} \mathbf{w} \cdot \mathbf{x}$$

subject to $f(\mathbf{x}) = y$, where $\mathbf{w}$ is the price vector corresponding to productive inputs $\mathbf{x}$, f is the production function for a certain output good, and y is a specified amount of that output good.

Differentiating the objective function and the constrain, we find that the first order conditions for minimization are:

$$\mathbf{w} = \lambda Df(\mathbf{x}^*)$$

where $\mathbf{x}^*$ is the optimum. Dividing the i^{th} condition by the j^{th}, we find

$$\frac{w_i}{w_j} = \frac{\partial f(\mathbf{x}^*)}{\partial x_i} / \frac{\partial f(\mathbf{x}^*)}{\partial x_j}.$$

This says that the economic rate of substitution between inputs i and j – the rate at which one input can be substituted for the other while maintaining a constant cost – equals the technical rate of substitution – the rate at which one input can be substituted for another while maintaining constant output. If this equality did not hold, that would mean that there was a way to produce the same output at a lower price.

This situation can be represented graphically. The MRTS is the slope of the isoquant $f(x_1, x_2) = \bar{y}$, a curve in $\mathbb{R}^2$ that shows all combinations of inputs x_1 and x_2 that produce a given level of output $\bar{y}$, i.e. a level curve of the production function. The economic rate of substitution, w_1/w_2 is the negative of the slope of the isocost curve $w_1x_1 + w_2x_2 = C$, a line that shows all of the combinations of inputs that cost C given input prices. Cost minimization occurs when the at the value of C that makes the isocost curve tangent to the *given* isoquant. We vary C until tangency is achieved, holding $\bar{y}$ fixed.

To verify that we have a cost minimum and not a maximum, we need to look at the second order conditions, which involve a bordered Hessian (see Varian pg. 23).

Problems With First Order Conditions

There are several ways that the first order conditions can break down:

1. If the production function is not differentiable (e.g. Leontieff) we cannot apply calculus.

2. The first order conditions assume an interior optimum. If $x_i^* = 0$, the appropriate first order condition becomes $\lambda \partial f(\mathbf{x}^*)/\partial x_i \leq w_i$.

3. We assumed that a cost-minimizing input combination exists; it might not. This isn't really a problem. The objective function $\mathbf{w} \cdot \mathbf{x}$ is continuous, so all we need to do is restrict attention to a compact set to guarantee the existence of a minimum.

4. The first order conditions are necessary, but not sufficient.

5.2 Conditional Factor Demand

Definition 5.2.1 (Conditional Factor Demand)**.** *The function that gives the cost minimizing choice of inputs* $\mathbf{x}^*$ *given input prices* $\mathbf{w}$ *and output requirement* y *is the conditional factor demand function* $\mathbf{x}(\mathbf{w}, y)$*.*

1. Conditional factor demands depend on factor prices *and* output level.

2. The conditional factor demand is sometimes called the *output constrained* factor demand.

3. **Note that cross-price effects for conditional factor demand are equal:** $\partial x_i/\partial w_j = \partial x_j/\partial w_i$.

4. When there are only two factors x_1, x_2 they must be **substitutes**: $\partial x_1/\partial w_2 = \partial x_2/\partial w_2 > 0$. When there are more factors, cross-price effects can go either way.

5. **The substitution matrix is negative semidefinite:** $\mathbf{Dx}(\mathbf{w}) \leq 0$.

5.3 Algebra of Cost Minimization

Consider observed output levels of a firm y_t, factor prices $\mathbf{w}^t$, and factor levels $\mathbf{x}^t$ for $t = 1, \ldots, T$.

Definition 5.3.1 (Weak Axiom of Cost Minimization)**.** *If a firm is cost minimizing, then the cost of its observed choice of inputs must be no greater than the cost of any other level of inputs that would produce at least as much output:* $\mathbf{w}^t \cdot \mathbf{x}^t \leq \mathbf{w}^t \cdot \mathbf{x}^s$ *for all* s *and* t *such that* $y^s \geq y^t$*.*

Proposition 5.3.1 (Change in Factor Prices and Conditional Demands)**.** *The vector of factor demands must move opposite the direction of factor prices:* $\boldsymbol{\Delta}\mathbf{w}\boldsymbol{\Delta}\mathbf{x} \leq 0$*.*

Chapter 6

Cost Function

6.1 Average and Marginal Costs

Definition 6.1.1 (Cost Function)**.** *The cost function is expressed as the value of the conditional factor demands:* $c(\mathbf{w}, y) = \mathbf{w} \cdot \mathbf{x}(\mathbf{w}, y)$.

In the short run, some factors are fixed. Let $\mathbf{x}_f$ be the fixed factors, and $\mathbf{x}_v$ be the variable factors. Accordingly, write $\mathbf{w} = (\mathbf{w}_v, \mathbf{w}_f)$. We write the short-run conditional factor demand functions as $\mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f)$. These usually depend on the level of fixed costs.

Definition 6.1.2 (Short Run Costs)**.** *The various short run cost functions are defined as follows:*

- *Short-Run Total Cost:* $STC = c(\mathbf{w}, y, \mathbf{x}_f) = \mathbf{w}_v \cdot \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f) + \mathbf{w}_f \cdot \mathbf{x}_f$
- *Short-Run Variable Cost:* $SVC = \mathbf{w}_v \cdot \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f)$
- *Fixed Cost:* $FC = \mathbf{w}_f \cdot \mathbf{x}_f$
- *Short-Run Average Total Cost:* $SATC = STC/y$
- *Short-Run Average Variable Cost:* $SAVC = SVC/y$
- *Short-Run Average Fixed Cost:* $SAFC = FC/y$
- *Short-Run Marginal Cost:* $SMC = \partial STC/\partial y$

In the long-run, all factors are variable. We can write the long-run cost function in terms of the short-run cost function by considering a two-stage optimization. Let $\mathbf{x}_f^*(\mathbf{w}, y)$ be the optimal choice of fixed factors and $\mathbf{x}_v^*(\mathbf{w}, y) = \mathbf{x}_v^*(\mathbf{w}, y, \mathbf{x}_f^*(\mathbf{w}, y))$ be the optimal long-run choice of variable factors. Then the long-run cost function can be written:

$$c(\mathbf{w}, y) = \mathbf{w}_v \cdot \mathbf{x}_v^*(\mathbf{w}, y) + \mathbf{w}_f \cdot \mathbf{x}_f^*(\mathbf{w}, y) = c(\mathbf{w}, y, \mathbf{x}_f^*(\mathbf{w}, y))$$

Definition 6.1.3 (Short Run Costs)**.** *The various short run cost functions are defined as follows:*

- *Long-Run Total Cost:* $LTC = c(\mathbf{w}, y) = c(\mathbf{w}, y, \mathbf{x}_f^*(\mathbf{w}, y))$
- *Long-Run Average Cost:* $LAC = LTC/y = c(\mathbf{w}, y)/y$
- *Long-Run Marginal Cost:* $LMC = \partial LTC/\partial y = \partial c(\mathbf{w}, y)/\partial y$

NOTE: Long-run Average Variable Cost (LAVC) equals Long-run Average Cost (LAC) because all factors are variable in the long-rong. For the same reason, Long-run Fixed Costs are zero.

Proposition 6.1.1 (Total Costs Constant Returns to Scale)**.** *If the production function exhibits constant returns to scale, the cost function is linear in output:*

$$c(\mathbf{w}, y) = y \cdot c(\mathbf{w}, 1)$$

Proof. The intuition is that once the firm has chosen the cheapest combinations of inputs to produce a single unit of output (i.e. $c(\mathbf{w}, 1)$), by constant returns to scale it can replicate this process for as many units as it likes. □

Proposition 6.1.2 (Average and Marginal Costs under CRS)**.** *For constant returns to scale technology, average cost, average variable cost, and marginal cost curves are equivalent.*

Proof. By CRS, $c(\mathbf{w}, y) = y \cdot c(\mathbf{w}, 1)$. Differentiating with respect to y gives $LMC = c(\mathbf{w}, 1)$. Dividing by y gives $LAC = c(\mathbf{w}, 1)$. Hence $LMC = LAC$. Further, since fixed costs are zero in the long-run, $LAC = LAVC$. □

NOTE: for constant returns to scale technology, total and average cost grow *linearly* with desired output level, for decreasing returns they grow *faster* than linearly, and for increasing returns they grow *slower* than linearly.

6.2 Shape of Average Cost Curves

In the short-run, costs have a variable and fixed component. Thus:

$$SAC = \frac{c(\mathbf{w}, y, \mathbf{x}_f)}{y} = \frac{\mathbf{w}_f \cdot \mathbf{x}_f}{y} + \frac{\mathbf{w}_v \cdot \mathbf{x}_v(\mathbf{w}, y, \mathbf{x_f})}{y} = SAFC + SAVC$$

SAFC is decreasing, and SAVC (although it may initially decrease because of a scale economy) is *eventually* increasing as output nears capacity. Hence, SAC is U-shaped. In the long-run, all factors are fixed, so AFC is zero. Long-run average costs could be constant or decreasing by a replication argument. However if there are long-run fixed factors, LATC will be U-shaped by the same argument as given in the short-run case.

6.3 Average and Marginal Costs

From here onwards, we suppress the dependence of cost on factor prices and write $TC = c(y)$, etc. Everything in this section applies to *both* the long and short run.

Assuming that average costs are U-shaped, the following first order condition is necessary and sufficient for minimum ATC:

$$\frac{y \cdot c'(y) - c(y)}{y^2} = 0$$

Rearranging and letting $y^* = \arg\min_y ATC$:

$$c'(y*) = \frac{c(y^*)}{y^*}$$

That is, **marginal and average costs are equal at minimum average cost**. For $y < y^*$, $MC < ATC$ and for $y > y^*$, $MC > ATC$.

Let $c_v(y)$ be variable costs. Then $AVC(y) = c_v(y)/y$. By L'Hôpital's rule:

$$\lim_{y \to 0} \frac{c_v(y)}{y} = \frac{c_v'(0)}{1} = c_v'(0)$$

Hence, **marginal cost and average variable cost are equal at** $y = 0$.

6.4 Long and Short-run Costs

The long-run cost curve must *never* lie above any short-run cost curve because the short-run cost minimization problem is a constrained version of the unconstrained, long-run cost minimization problem.

Let $c(y) = c(y, z(y))$ denote the long run cost function, holding factor prices fixed, and letting $z(y)$ be the cost minimizing demand for a single fixed factor. Let y^* be some given level of output, and $z^* = z(y^*)$ be the long-run demand for the fixed factor. The short-run cost $c(y, z^*)$ must be at least as great as the long-run cost $c(y, z(y))$ for all levels of output. The two are equal precisely when the desired level of output is y^*. **Hence, the long and short-run cost curves must be tangent at** y^*.

This is just the envelope theorem. The slope of the long-run cost curve at y^* is

$$\begin{aligned} \frac{dc(y^*, z(y*))}{dy} &= \begin{bmatrix} \frac{\partial c(y^*,z^*)}{\partial y} & \frac{\partial c(y^*,z^*)}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{dy}{dy} \\ \frac{dx}{dy} \end{bmatrix} \\ &= \frac{\partial c(y^*, z^*)}{\partial y} + \frac{\partial c(y^*, z^*)}{\partial z} \cdot \frac{dz(y^*)}{dy} \end{aligned}$$

But since z^* is the optimal choice of fixed factors at output level y^*, it must be that $\frac{\partial c(y^*,z^*)}{\partial z} = 0$. The result follows.

NOTE: Since long and short run *total* costs are tangent at y^*, so are the corresponding *average* cost curves.

6.5 Properties of the Cost Function

Theorem 6.5.1 (Properties of the Cost Function)**.** *The cost function satisfies the following:*

1. *Costs are non-decreasing in factor prices. That is, if* $\mathbf{w}' \geq \mathbf{w}$*, then* $c(\mathbf{w}', y) \geq c(\mathbf{w}, y)$*.*

2. *Costs are homogeneous of degree 1 in factor prices. In other words,* $c(\lambda \cdot \mathbf{w}, y) = \lambda \cdot c(\mathbf{x}, y)$*, for* $\lambda > 0$*.*

3. *Costs are concave in factor prices, that is:* $c(\lambda \cdot \mathbf{w} + (1 - \lambda) \cdot \mathbf{w}', y) \geq \lambda \cdot c(\mathbf{w}, y) + (1 - \lambda) \cdot c(\mathbf{w}', y)$*, for* $\lambda \in [0, 1]$*.*

4. *Costs are continuous in factor prices (for positive factor prices).*

Proof. We consider each in turn:

1. Let $\mathbf{w}' \geq \mathbf{w}$ (i.e. $w'_k \geq w_k$ for all k), and let $\mathbf{x}'$ and $\mathbf{x}$ be the associated cost-minimizing input bundles. By cost minimization, $\mathbf{w} \cdot \mathbf{x} \leq \mathbf{w} \cdot \mathbf{x}' \leq$ by cost minimization. Further, since $\mathbf{w} \leq \mathbf{w}'$ we have $\mathbf{w} \cdot \mathbf{x}' \leq \mathbf{w}' \cdot \mathbf{x}'$. Combining these, $\mathbf{w}\mathbf{x}' \leq \mathbf{w}' \cdot \mathbf{x}'$.

2. If $\mathbf{x}$ minimizes $\mathbf{w} \cdot \mathbf{x}$ subject to $f(\mathbf{x}) = y$, then it also minimizes $\lambda \cdot \mathbf{w} \cdot \mathbf{x}$ subject to $f(\mathbf{x}) = y$ for $\lambda > 0$.

3. Omitted. See intuitive argument below.

4. Follows from the theorem of the maximum.

□

NOTE: There is some nice intuition for the concavity of the cost function. If one factor price increases while all others remain the same, costs cannot go down (by property 1 from the above theorem). However, costs increase at a *decreasing* rate because the firm shifts away from the factor whose price has risen.

For example, suppose x^* is the cost-minimizing input bundle at factor prices w^*. And suppose the price of factor 1 changes from w_1^* to w_1. If the firm doesn't re-optimize its inputs, continuing to x^*, its costs will be $w_1 \cdot x_1^* + \sum_{i=2}^{k} w_i^* \cdot x_i^*$. The minimal cost of production $c(\mathbf{w}, y)$ *must be less* than this "passive" cost function. Thus, the graph of $c(\mathbf{w}, y)$ lies *below* the passive cost curve, with the two curves coinciding (tangency) at $w_1 = w_1^*$. Thus, the tangent is above the function, so $c(\mathbf{w}, y)$ is concave with respect to w_1.

Theorem 6.5.2 (Shephard's Lemma). *Let $x_i(\mathbf{w}, y)$ be the firm's conditional factor demand for input i. Then if the cost function is differentiable at $(\mathbf{x}, y)$, and all factor prices are positive,*

$$x_i(\mathbf{w}, y) = \frac{\partial c(\mathbf{w}, y)}{\partial w_i}$$

Proof. Use the envelope theorem for constrained optimization. □

Proposition 6.5.1. *The Lagrange multiplier in the cost minimization is simply marginal cost:*

$$\frac{\partial c(\mathbf{w}, y)}{\partial y} = \lambda$$

6.6 Comparative Statics from the Cost Function

A number of properties follow immediately from Shephard's Lemma:

1. Because costs are nondecreasing in factor prices, $\partial c(\mathbf{w}, y)/\partial w_i \geq 0$. Hence, by Shepard's Lemma, conditional factor demands are non-negative: $x_i(\mathbf{w}, y) \geq 0$.

2. The cost function is homogeneous of degree one in factor prices. Since the derivative of a function that is homogeneous of degree k is homogeneous of degree $k - 1$ (see Chapter 1), conditional factor demands are homogeneous of degree zero in factor prices.

3. By the concavity of costs in factor prices, cross-price effects on conditional factor demands are symmetric:

$$\frac{\partial x_i(\mathbf{w}, y)}{\partial w_j} = \frac{\partial^2 c(\mathbf{w}, y)}{\partial w_i \partial w_j} = \frac{\partial^2 c(\mathbf{w}, y)}{\partial w_j \partial w_i} = \frac{\partial x_j(\mathbf{w}, y)}{\partial w_i}$$

4. Because costs are concave in factor prices and the diagonal elements of a negative semi-definite matrix are nonpositive, conditional factor demand curves are weakly downward sloping:

$$\frac{\partial x_i(\mathbf{x}, y)}{\partial w_i} = \frac{\partial^2 c(\mathbf{w}, y)}{\partial w_i^2} \leq 0$$

5. The vector of changes in factor demands moves "opposite" the vector of changes in factor prices. That is, $\mathbf{dw} \cdot \mathbf{dx} \leq 0$.

Chapter 7

Consumer Preferences

Let X denote a consumer's *consumption set*, where $X = \mathbb{R}^k_+$. Given bundles $\mathbf{x}, \mathbf{y} \in X$ the preference relation $\mathbf{x} \succsim \mathbf{y}$ indicates that the consumer considers $\mathbf{x}$ at least as good as $\mathbf{y}$.

7.1 Properties of Preferences

We make a number of assumptions about the preference relation $\succsim$.

Definition 7.1.1 (Completeness)**.** *For all* $\mathbf{x}, \mathbf{y} \in X$*, either* $\mathbf{x} \succsim \mathbf{y}$*,* $\mathbf{y} \succsim \mathbf{x}$*, or both.*

Completeness means that the consumer can meaningfully rank any two bundles in the consumption set.

Definition 7.1.2 (Reflexivity)**.** *For all* $\mathbf{x} \in X$*,* $\mathbf{x} \succsim \mathbf{x}$*.*

Reflexivity says that any bundle is weakly preferred to itself.

Definition 7.1.3 (Transitivity)**.** *For all* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$*, if* $\mathbf{x} \succsim \mathbf{y}$ *and* $\mathbf{y} \succsim \mathbf{z}$*, then* $\mathbf{x} \succsim \mathbf{z}$*.*

Transitivity is needed to consider preference maximization. It ensures that any collection of bundles contains a set of best elements.

Definition 7.1.4 (Continuity)**.** *For all* $\mathbf{x} \in X$*, the weakly better than set* $\{\mathbf{z} \in X : \mathbf{z} \succsim \mathbf{x}\}$ *and the weakly worse than set* $\{\mathbf{z} \in X : \mathbf{x} \succsim \mathbf{z}\}$ *are closed (or equivalently the strictly better and worse than sets are open).*

By continuity, the weakly better than and worse than sets contain all of their limit points. Thus if $\lim_{k\to\infty} \mathbf{x}^k = \mathbf{x}^*$ and $\mathbf{x}^k \succsim \mathbf{y}$ for all k, we have that $\mathbf{x}^* \succsim z$. We can also think of continuity in terms of the strictly better and worse than sets. As these are assumed to be open sets, each of their elements is an interior point. Hence, if $\mathbf{y} \succ \mathbf{z}$ and $\mathbf{x}$ is sufficiently close to $\mathbf{y}$, we have $\mathbf{x} \succ \mathbf{z}$.

Continuity is a technical requirement that allows the preference relation to be represented by a continuous utility function, provided that the relation is also complete, reflexive, and transitive.

Definition 7.1.5 (Utility Function). *A function $u : X \to \mathbb{R}$ is called a utility function representing preferences $\succsim$ provided that $\mathbf{x} \succsim \mathbf{y}$ if and only if $u(\mathbf{x}) \geq u(\mathbf{y})$.*

Proposition 7.1.1 (Ordinal Character of Utility). *Let $f : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function, and $(\mathbf{x})$ be a utility function representing preferences $\succsim$. Then $f(u(\mathbf{x}))$ is a utility function representing the same preferences.*

Proof. By definition, $u(\mathbf{x}) \geq u(\mathbf{y})$ for any $\mathbf{x} \succsim \mathbf{y}$. Since f is a strictly increasing function, it preserves the direction of inequalities. Hence $f(u(\mathbf{x})) \geq f(u(\mathbf{y}))$ for any $\mathbf{x} \succsim \mathbf{y}$. □

We have several alternative assumptions that capture the notion that "more is better."

Definition 7.1.6 (Weak Monotonicity). *If $x_i \geq y_i$ for all i, then $\mathbf{x} \succsim \mathbf{y}$*

Weak monotonicity says that **at least as much as everything is at least as good**. Assuming free disposal of unwanted goods, this assumption is trivial.

Definition 7.1.7 (Strong Monotonicity). *If $x_i \geq y_i$ for all i and $x_i > y_i$ for some i, then $\mathbf{x} \succsim \mathbf{y}$*

Strong monotonicity says that **no less of anything, and strictly more of something is strictly better**. This assumption requires that none of the goods is a "bad."

Definition 7.1.8 (Local Nonsatiation). *For any $\mathbf{x} \in X$ and any $\epsilon > 0$ there exists some bundle $\mathbf{y} \in N_\epsilon(\mathbf{x})$ such that $\mathbf{y} \succ \mathbf{x}$.*

Local nonsatiation says that in any neighborhood of a given bundle, no matter how small, there is something strictly better. **This rules out "thick" indifference curves**.

Definition 7.1.9 (Convexity of Preferences). *For any $\mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathbf{x} \succsim \mathbf{z}$ and $\mathbf{y} \succsim \mathbf{z}$, it follows that $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \succsim \mathbf{z}$ for all $\lambda \in [0, 1]$.*

Definition 7.1.10 (Strict Convexity of Preferences). *For any $\mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathbf{x} \succsim \mathbf{z}$ and $\mathbf{y} \succsim \mathbf{z}$, it follows that $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \succ \mathbf{z}$ for all $\lambda \in (0, 1)$.*

Convexity says that agents prefer mixtures to extremes. Weak convexity allows indifference curves with flat sections; strict convexity does not. convexity is a generalization of "diminishing MRS."

7.2 Marginal Rate of Substitution

Let $u(\mathbf{x})$ be a consumer's utility function. The **Marginal Rate of Substitution** between goods i and j ($MRS_{i,j}$), gives the amount of good j required to compensate the consumer for a unit loss of good i. Thus, it is the slope of the indifference curve in i-j space.

We want the slope of a *particular* indifference curve between goods i and j, so we treat u as a function of these goods alone, holding the levels of all other goods constant, and evaluate at a fixed utility level $\bar{u}$. Then $u(x_i, x_j) = \bar{u}$. Now we totally differentiate both sides of the expression with respect to x_i, treating x_j as a function of x_i:

$$\begin{aligned}
\frac{d}{dx_i}u(x_i, x_j) &= \frac{d}{dx_i}\bar{u} \\
\left[\begin{array}{cc} \partial u/\partial x_i & \partial u/\partial x_i \end{array}\right]\left[\begin{array}{c} dx_i/dx_i \\ (dx_j/dx_i)_{\bar{u}} \end{array}\right] &= 0 \\
\left[\begin{array}{cc} \partial u/\partial x_i & \partial u/\partial x_i \end{array}\right]\left[\begin{array}{c} 1 \\ (dx_j/dx_i)_{\bar{u}} \end{array}\right] &= 0 \\
\frac{\partial u}{\partial x_i} + \frac{\partial u}{\partial x_j}\left(\frac{dx_j}{dx_i}\right)_{\bar{u}} &= 0
\end{aligned}$$

The notation $(dx_j/dx_i)_{\bar{u}}$ indicates that although x_j is not an explicit function of x_i, it can be treated as an *implicit* function of x_i when $u(\cdot)$ is held fixed at $\bar{u}$. Rearranging:

$$MRS_{i,j} = \left(\frac{dx_j}{dx_i}\right)_{\bar{u}} = -\frac{\partial u}{\partial x_i}/\frac{\partial u}{\partial x_j}$$

Proposition 7.2.1 (MRS Invariant to Increasing Transformations)**.** *Let $v(u)$ be an increasing function. Then the MRS between two goods computed from $v(u(\mathbf{x}))$ is the same as that computed from $u(\mathbf{x})$. That is, the* ***MRS is a property of underlying preferences that does not depend on the utility function used to represent them****.*

Proof. Differentiate $v(u(\mathbf{x}) = \bar{v}$ totally with respect to x_i as above:

$$\begin{aligned}
\left[\begin{array}{cc} v'(u)\partial u/\partial x_i & v'(u)\partial u/\partial x_i \end{array}\right]\left[\begin{array}{c} 1 \\ (dx_j/dx_i)_{\bar{v}} \end{array}\right] &= 0 \\
v'(u)\frac{\partial u}{\partial x_i} + v'(u)\frac{\partial u}{\partial x_j}\left(\frac{dx_j}{dx_i}\right)_{\bar{v}} &= 0
\end{aligned}$$

Rearranging:

$$MRS^v_{i,j} = \left(\frac{dx_j}{dx_i}\right)_{\bar{v}} = -v'(u)\frac{\partial u}{\partial x_i}/v'(u)\frac{\partial u}{\partial x_j} = -\frac{\partial u}{\partial x_i}/\frac{\partial u}{\partial x_j} = MRS^u_{i,j}$$

□

7.3 The Utility Maximization Problem

Let m be the consumer's income, $\mathbf{p}$ the vector of prices, and k be the number of goods available. Then the consumer's budget set (the set of all affordable consumption bundles) is given by $B = \{\mathbf{x} \in X : \mathbf{p} \cdot \mathbf{x} \leq m\}$. The consumer chooses $\mathbf{x}$ to maximize his utility over the budget set B:

$$\max_{\mathbf{x} \in X} u(\mathbf{x})$$

subject to:

$$\mathbf{p} \cdot \mathbf{x} \leq m$$

Assuming an interior solution, the first order conditions are:

$$\frac{\partial u(\mathbf{x}^*)}{\partial x_i} = \lambda p_i$$

for all i. In other words: $MU_i/p_i = \lambda$ for all i, so **the marginal utility *per dollar* of each good is equal to λ, the marginal utility of income**. Dividing the first order condition for good i by that for good j, we have:

$$\frac{\partial u(\mathbf{x}^*)}{\partial x_i}/\frac{\partial u(\mathbf{x}^*)}{\partial x_j} = \frac{p_i}{p_j}$$

for all i,j. That is, $-MRS_{i,j} = p_i/p_j$ for all i,j. **The marginal rate of substitution between any two goods equals the economic rate of substitution**.

Some Considerations

1. In order for a solution of the above to exist, B must be compact, and u continuous. By the assumption of continuous preferences, u can be assumed continuous. In order for B to be compact (since it is a subset of $\mathbb{R}^k$) is must be closed and bounded. This requires $p_i > 0$ for all i.

2. The specific choice of u has no effect on the solution. Any utility function that represents the underlying preferences will find the same constrained maximizer.

3. If we multiply m and $\mathbf{p}$ be a constant, the constraint set B is unchanged, so the optimal $\mathbf{x}^*$ is also unchanged.

4. Everything we did above assumed an interior optimum. If this is not the case, i.e. $x_i^* = 0$ for some good i, then $\partial u(\mathbf{x}^*)/\partial x_i \leq \lambda p_i$, so $MU_i/p_i \leq \lambda$. **The marginal utility per dollar of good i is less than or equal to λ.** The inequality becomes an equality in the case of a corner tangency; otherwise it is strict.

5. Technically, we need to check the second order conditions. These depend on some notion of local quasi-concavity, but are complicated.

7.4 Indirect Utility

Definition 7.4.1. *Let $\mathbf{x}^*$ denote the optimal consumption bundle from the consumer's optimization problem above. The indirect utility function $v(\mathbf{x}, m)$ is given by $v(\mathbf{x}, m) = u(\mathbf{x}^*)$. Thus, it gives maximum utility as a function of prices and income.*

Theorem 7.4.1 (Properties of Indirect Utility)**.** *Indirect utility has the following properties:*

1. *$v(\mathbf{x}, m)$ is non-increasing in $\mathbf{p}$ (it remains unchanged if the price of a good you don't buy rises).*

2. *$v(\mathbf{x}, m)$ is non-decreasing in m, strictly increasing in m under local non-satiation.*

3. *$v(\mathbf{x}, m)$ is homogeneous of degree 0 in prices and income.*

4. *$v(\mathbf{x}, m)$ is quasi-convex in prices. That is, the worse-than sets of v holding m constant are convex.*

5. *$v(\mathbf{x}, m)$ is continuous for positive prices and income.*

Proof. We consider each in turn:

1. let B be the budget set at prices $\mathbf{p}$ and B' be the budget set at prices $\mathbf{p}'$ where $\mathbf{p} \leq \mathbf{p}'$. Then $B' \subseteq B$, so the maximum of u over B' cannot exceed the maximum of u over B.

2. Same idea as part 1.

3. As mentioned above, if income and all prices are multiplied by a positive constant, the budget set remains unchanged.

4. Proof by contradiction (see Varian).

5. Follows from the Theorem of the Maximum.

□

The level sets of the indirect utility function are sometimes called "price indifference curves." By the preceding theorem, the direction of improvement is southwest, and the lower contour sets (which lie to the northeast) are convex.

7.5 Expenditure Function

Under local non-satiation, $v(\mathbf{p}, m)$ is strictly increasing in m. Then, holding prices constant, its inverse function exists: we can solve for m as a function of utility holding prices fixed. This function is known as the **expenditure function**. It gives the minimum amount of income necessary to attain utility level u at prices $\mathbf{p}$, denoted $e(\mathbf{p}, u)$.

An equivalent way to derive the expenditure function is by considering the "dual" of the utility maximization problem–the expenditure minimization problem:

$$e(\mathbf{p}, u) = \min_{\mathbf{x} \in X} \mathbf{p} \cdot \mathbf{x}$$

subject to

$$u(\mathbf{x}) \geq \bar{u}$$

Note that the expenditure function is exactly analogous to the firm's cost function, and hence has the same properties.

Theorem 7.5.1 (Properties of Expenditure Function)**.** *The expenditure function has the following properties:*

1. *$e(\mathbf{p}, u)$ is nondecreasing in prices.*
2. *$e(\mathbf{p}, u)$ is homogeneous of degree 1 in prices.*
3. *$e(\mathbf{p}, u)$ is concave in prices.*
4. *$e(\mathbf{p}, u)$ is continuous in $\mathbf{p}$ for positive prices.*
5. *Let $\mathbf{x}^* = \mathbf{h}(\mathbf{p}, u)$ be the expenditure minimizing bundle that achieves utility level u at prices $\mathbf{p}$. Then, whenever the derivative exists and $p_i > 0$, $h_i(\mathbf{p}, u) = \partial e(\mathbf{p}, u)/\partial p_i$.*

Proof. Note that the following arguments are identical those for the firm's cost minimization problem.

1.
2. If $\mathbf{x}^*$ minimizes $\mathbf{p} \cdot \mathbf{x}$ subject to $u(\mathbf{x}) \geq \bar{u}$, it also minimizes $\mathbf{p}' \cdot \mathbf{x}$ subject to the same constraint for any $\mathbf{p}'$. Suppose $\mathbf{p}' = \lambda \mathbf{p}$. Then $e(\mathbf{p}', m) = \mathbf{p}' \cdot \mathbf{x}^* = \lambda \mathbf{p} \cdot \mathbf{x}^* = \lambda e(\mathbf{p}, m)$.

3.

4.

5. Envelope theorem.

□

7.6 Hicksian vs. Marshallian Demand

The solution function to the utility maximization problem, $\mathbf{x}(\mathbf{p}, m)$, is called the **Marshallian Demand Function** (when it is single-valued). Marshallian Demand allows utility to vary in the face of changing prices and income. Own price effects are *can* be positive for Marshallian Demand.

The solution function to the expenditure minimization problem $\mathbf{h}(\mathbf{p}, u)$ is called **Hicksian Demand Function**, (when it is single-valued). It is sometimes called the **Compensated Demand Curve** since it allows income to vary in order to maintain a fixed utility level.

Theorem 7.6.1 (Properties of Hicksian Demand)**.** *The Hicksian Demand function $\mathbf{h}(\mathbf{p}, u)$ has the following properties:*

1. *$\mathbf{h}(\mathbf{p}, u)$ is homogeneous of degree zero in prices.*
2. *$h_i(\mathbf{p}, u) = \partial e(\mathbf{p}, u)/\partial p_i$ whenever the derivative exists and $p_i > 0$.*
3. *Own price effects are non-positive for Hicksian Demand.*
4. *Cross price effects are symmetric for Hicksian Demand.*

Proposition 7.6.1 (Some Important Properties)**.** *The expenditure minimization and Utility Maximization Problems are related in the following important ways:*

1. *The minimum expenditure necessary to reach utility level $v(\mathbf{p}, m)$ is m. That is, $e(\mathbf{p}, v(\mathbf{p}, m)) = m$.*
2. *The maximum utility from income $e(\mathbf{p}, u)$ is u. That is, $v(\mathbf{p}, e(\mathbf{p}, u))$.*
3. *The Marshallian Demand at income m is the same as the Hicksian Demand at utility $v(\mathbf{p}, m)$. That is, $x_i(\mathbf{p}, m) = h_i(\mathbf{p}, v(\mathbf{p}, m))$.*
4. *The Hicksian Demand at utility u is the same as the Marshallian Demand at income $e(\mathbf{p}, u)$. That is, $h_i(\mathbf{p}, u) = x_i(\mathbf{p}, e(\mathbf{p}, u))$.*

Theorem 7.6.2 (Roy's Identity)**.** *Let* $\mathbf{x}(\mathbf{p}, m)$ *be the Marshallian Demand Function. Then for all* $i = 1, \ldots, k$

$$x_i(\mathbf{p}, m) = -\frac{\partial v(\mathbf{p}, m)}{\partial p_i} / \frac{\partial v(\mathbf{p}, m)}{\partial m}$$

provided that the derivatives are well-defined, and $m > 0$*.*

Proof. What follows is essentially a proof of the envelope theorem. First, note that the indirect utility function is defined by $v(\mathbf{p}, m) \equiv u(\mathbf{x}(\mathbf{p}, m))$. Differentiating both sides with respect to p_j,

$$\frac{\partial v(\mathbf{p}, m)}{\partial p_j} = \sum_{i=1}^{k} \frac{\partial u(\mathbf{x})}{\partial x_i} \frac{\partial x_i}{\partial p_j}$$

Since $\mathbf{x}(\mathbf{p}, m)$ is a Marshallian Demand Function, it satisfies the first order condition for utility maximization: $\partial u(\mathbf{x})/\partial x_i = \lambda p_i$. Substituting into the above,

$$\frac{\partial v(\mathbf{p}, m)}{\partial p_j} = \lambda \sum_{i=1}^{k} p_i \frac{\partial x_i}{\partial p_j} \tag{7.1}$$

The budget constraint gives $\mathbf{p} \cdot \mathbf{x}(\mathbf{p}, m) = m$. Differentiating this expression with respect to p_j and rearranging,

$$-x_j(\mathbf{p}, m) = \sum_{i=1}^{k} p_i \frac{\partial x_i}{\partial p_j} \tag{7.2}$$

Combining 7.1 and 7.2,

$$\frac{\partial v(\mathbf{p}, m)}{\partial p_j} = -\lambda x_j(\mathbf{p}, m) \tag{7.3}$$

Now follow the same steps for m. First differentiate $v(\mathbf{p}, m) \equiv u(\mathbf{x}(\mathbf{p}, m))$ with respect to m, yielding:

$$\frac{\partial v(\mathbf{p}, m)}{\partial m} = \sum_{i=1}^{k} \frac{\partial u(\mathbf{x})}{\partial x_i} \frac{\partial x_i}{\partial m}$$

Substituting the first order condition for utility maximization: $\partial u(\mathbf{x})/\partial x_i = \lambda p_i$ gives

$$\frac{\partial v(\mathbf{p}, m)}{\partial m} = \lambda \sum_{i=1}^{k} p_i \frac{\partial x_i}{\partial m} \tag{7.4}$$

Differentiating $\mathbf{p} \cdot \mathbf{x}(\mathbf{p}, m) = m$ with respect to m,

$$\sum_{i=1}^{k} p_i \frac{\partial x_i}{\partial m} = 1 \tag{7.5}$$

Combining 7.4 and 7.5, we see that λ is the marginal utility of income:

$$\frac{\partial v(\mathbf{p}, m)}{\partial m} = \lambda \tag{7.6}$$

Finally, combining 7.6 with 7.3

$$\frac{\partial v(\mathbf{p}, m)}{p_j} = -\frac{\partial v(\mathbf{p}, m)}{m} \cdot x_j(\mathbf{p}, m)$$

Rearranging this expression gives the desired result. □

7.7 Money Metric Utility Functions

Definition 7.7.1 (Direct Money Metric Utility). *The direct money metric utility function, $m(\mathbf{p}, \mathbf{x})$, gives the minimum expenditure at prices $\mathbf{p}$ necessary to purchase a bundle at least as good as $\mathbf{x}$. That is, $m(\mathbf{p}, \mathbf{x}) \equiv e(\mathbf{p}, u(\mathbf{p}))$.*

For a fixed $\mathbf{x}$, $u(\mathbf{x})$ is fixed. Hence $m(\mathbf{p}, \mathbf{x})$ has all the properties of the expenditure function. When $\mathbf{p}$ is fixed, $m(\mathbf{p}, \mathbf{x})$ is a utility function. Note that the expenditure function is strictly increasing in u for fixed prices provided utility is locally non-satiated: to get a higher utility you need to spend more. Hence, for fixed prices $m(\mathbf{p}, \mathbf{x})$ is simply **an increasing transformation of the utility function**. Hence it is a utility function itself.

Definition 7.7.2 (Indirect Money Metric Utility Function). *The indirect money metric utility function $\mu(\mathbf{p}; \mathbf{q}, m)$ measures how much money one would need at prices $\mathbf{p}$ to be as well off as one would be facing prices $\mathbf{q}$ and income m. That is, $\mu(\mathbf{p}; \mathbf{q}, m) \equiv e(\mathbf{p}, v(\mathbf{q}, m))$.*

Chapter 8

Choice

8.1 Comparative Statics

Holding prices fixed but allowing income to vary, the resulting locus of utility-maximizing bundles is called the **income expansion path**, or sometimes the *income offer curve*. The function that plots the demand for a given good against income (with income on the y-axis) is called the **Engel Curve**. There are several possibilities:

1. The income expansion path, and hence the Engel curve for each good, is a straight line through the origin. This means that the consumer's demand functions have **unit income elasticity**: he consumes the same *proportion* of each commodity at any income.

2. The income expansion path bends *away* from one good and *towards* the other. This means that the when the consumer has more income he buys more of both goods, but proportionally more of one than the other. The good whose budget share increases with income is called a **luxury good**; the other good is called a **necessary good**.

3. The income expansion path bends *backwards*. This means that at some point an increase in income leads to a *decrease* in the consumption of one of the goods. A good whose demand decreases with income is called an **inferior good**. A good whose demand increases with income is called a **normal good**.

Holding income fixed but allowing prices to vary gives the **price offer curve**. For most goods, demand decreases as price rises. A good whose demand *increases* when price rises is called a **Giffen good**.

Note: Demand can slope upwards, as it does in the case of a Giffen good. However, there is always an upper limit on demand: **a consumer cannot demand more than m/p_i of good i or he will spend more than his income!**

8.2 The Slutsky Equation

Hicksian demand is not directly observable because one of its arguments is utility. In contrast, Marshallian demand is always obervable: its arguments are income and prices. However, the two share an intimate connection via the **Slutsky Equation**:

$$\frac{\partial x_j(\mathbf{p}, m)}{\partial p_i} = \frac{\partial h_j(\mathbf{p}, v(\mathbf{p}, m))}{\partial p_i} - \frac{\partial x_j(\mathbf{p}, m)}{\partial m} \cdot x_i(\mathbf{p}, m)$$

This equation states that the change in the Marshallian demand for good j in response to a change in the price of good i can be decomposed into two effects: the income and substitution effects. This is because a change in prices affects overall purchasing power *and* relative prices.

The **substitution effect** is given by:

$$\frac{\partial h_j(\mathbf{p}, v(\mathbf{p}, m))}{\partial p_i}$$

This is the change in demand for good j that would result from a change in the price of good i if we were to keep utility stable at its initial level $v(\mathbf{p}, m)$ by adjusting income. It measures the effect of *relative prices* on demand. Remember that compensated *own-price* effects are non-positive, so if $i = j$, the substitution effect is non-positive.

The **income effect** is given by:

$$-\frac{\partial x_j(\mathbf{p}, m)}{\partial m} \cdot x_i(\mathbf{p}, m)$$

When the price of good i increases, if the consumer wants to keep buying as much of this good as before, he will have less money left over to spend on other goods, such as good j. The magnitude of this effect depends on how much of good i the consumer was buying before the price change: $x_i(\mathbf{p}, m)$. If the consumer wasn't buying any of goof i before the price change, there is no income effect.

8.2.1 Deriving the Slutsky Equation

Let x^* be the utility-maximizing bundle at prices $\mathbf{p}$ and income m, and let $u^* = u(\mathbf{x}^*)$. We know from above that $h_j(\mathbf{p}, u^*) = x_j(\mathbf{p}, e(\mathbf{p}, u^*))$. Differentiating both sides with respect to p_i:

$$\frac{\partial h_j(\mathbf{p},u^*)}{\partial p_i} = \begin{bmatrix} \frac{\partial x_j(\cdot)}{\partial p_1} & \dots & \frac{\partial x_j(\cdot)}{\partial p_k} & \frac{\partial x_j(\cdot)}{\partial m} \end{bmatrix} \begin{bmatrix} \frac{\partial p_1}{\partial p_i} \\ \vdots \\ \frac{\partial p_i}{\partial p_i} \\ \vdots \\ \frac{\partial p_k}{\partial p_i} \\ \frac{\partial e(\mathbf{p},u^*)}{\partial p_i} \end{bmatrix}$$

Hence:

$$\frac{\partial h_j(\mathbf{p},u^*)}{\partial p_i} = \frac{\partial x_j(\mathbf{p},m)}{\partial p_i} + \frac{\partial x_j(\mathbf{p},m)}{\partial p_i} \cdot \frac{\partial e(\mathbf{p},u^*)}{\partial p_i}$$

But $\partial e(\mathbf{p},u^*)/\partial p_i = h_j(\mathbf{p},u^*)$ and we know that $h_j(\mathbf{p},u^*) = x_j(\mathbf{p},m)$. Thus:

$$\frac{\partial h_j(\mathbf{p},u^*)}{\partial p_i} = \frac{\partial x_j(\mathbf{p},m)}{\partial p_i} + \frac{\partial x_j(\mathbf{p},m)}{\partial p_i} \cdot x_j(\mathbf{p},m)$$

Rearranging gives the Slutsky equation.

8.3 Properties of Demand

A number of properties follow from the fact that the expenditure function is concave in prices, and the fact that $h_j(\mathbf{p},u) = \partial e(\mathbf{p}.u)/\partial p_j$, which gives us that $\partial h_j(\mathbf{p},u)/\partial p_j = \partial^2 e(\mathbf{p},u)/\partial p_j \partial p_i$. Becuase $e(\mathbf{p},u)$ is concave in prices, the second derivative matrix (in prices) is negative semidefinite. Recall also that cross partials are equal so that $\partial^2 e(\mathbf{p},u)/\partial p_j \partial p_i = \partial^2 e(\mathbf{p},u)/\partial p_i \partial p_j$. These observations immediately imply the following:

Theorem 8.3.1. *Properties of Demand The following properties of Hicksian demand follow immediately from its relationship with the expenditure function*

1. *The matrix of substitution terms $\partial h_j(\cdot)/\partial p_j$ is negative semidefinite.*
2. *The matrix of substitution terms is symmetric.*
3. *Compensated, own-price effects are non-positive.*
4. *The substitution matrix*

$$\left(\frac{\partial x_j(\mathbf{p},m)}{\partial p_i} + \frac{\partial x_j(\mathbf{p},m)}{\partial m} \cdot x_i(\mathbf{p},m)\right)$$

   is negative semidefinite.

Proof. From the properties of the expenditure function:

1. Concavity of expenditure function.
2. Equality of cross-partials.
3. Corollary to part (1).
4. Rearrange the Slutsky equation and use part (1).

□

8.4 Elasticity

Definition 8.4.1 (Elasticity). *Suppose $y = f(x_1, x_2, \ldots, x_k)$. Then the Elasticity of y with respect to x_i, denoted E_{y,x_i}, is given by:*

$$E_{y,x_i} = \frac{\partial}{\partial x_i} f(x_1, x_2, \ldots, x_k) \cdot \left(\frac{x_i}{y}\right)$$

Proposition 8.4.1 (Constant Elasticity). *The elasticity of y with respect to x is constant if and only if $y = f(x)$ where $f(x) = \alpha x^\rho$ where α and ρ are arbitrary constants.*

Note: by this proposition any constant function and any line through the origin have constant elasticity.

Definition 8.4.2 (Elasticity of Substitution). *The elasticity of substitution is defined as the elasticity of $MRS_{1,2}$ with respect to x_2/x_1. That is:*

$$E_{sub} = \frac{dMRS_{1,2}}{d(x_2/x_1)} \cdot \frac{x_2/x_1}{MRS_{1,2}}$$

Proposition 8.4.2 (Constant Elasticity of Substitution). *The following utility function, commonly called the C.E.S. utility function, has constant elasticity of substitution $1 - \rho$:*

$$u(x_1, x_2) = \alpha x_1^\rho + \beta x_2^\rho$$

Proof. First, note that

$$MRS_{1,2} = \frac{\alpha}{\beta}\left(\frac{x_2}{x_1}\right)^{1-\rho}$$

Hence

$$\frac{dMRS_{1,2}}{d(x_2/x_1)} = (1-\rho)\frac{\alpha}{\beta}\left(\frac{x_2}{x_1}\right)^{-\rho}$$

Multiplying and simplifying the result, $E_{sub} = 1 - \rho$. □

NOTE: The CES utility (or production) function is more commonly written as $[\alpha x_1^\rho + \beta x_2^\rho]^{1/\rho}$. When $\rho = -\infty$ we have perfect substitutes, when $\rho = 1$ we have perfect complements.

8.4.1 Some Identities

Proposition 8.4.3 (Own, Cross-price and Income Elasticities)**.** *The own, cross-price, and income elasticities of demand for good i satisfy*

$$E_{x_i,p_1} + E_{x_i,p_2} + \ldots + E_{x_i,p_k} + E_{x_i,m} = 0$$

Proof. Changing income and the prices of all goods by the same factor leaves the constraint set for the utility maximization problem unchanged. Thus, Marshallian demand is homogeneous of degree zero in prices and income. Applying Euler's Theorem with $k = 0$, $\nabla x_i(\mathbf{p}, m) \cdot (\mathbf{x}, m) = 0$, or equivalently

$$\frac{\partial x_i(\mathbf{p}, m)}{\partial p_1} \cdot p_1 + \ldots + \frac{\partial x_i(\mathbf{p}, m)}{\partial p_k} \cdot p_k + \frac{\partial x_i(\mathbf{p}, m)}{\partial m} \cdot m = 0$$

Dividing both sides by $x_i(\mathbf{p}, m)$ gives the desired result. □

Proposition 8.4.4 (Income Elasticities)**.** *The sum of each good's income elasticity multiplied by its budget share is one. That is,*

$$\sum_{i=1}^{k} E_{x_i,m} \cdot \frac{p_i \cdot x_i(\mathbf{p}, m)}{m} = 1$$

Note: This result implies that if one good is inferior, i.e. its income elasticity is negative, another good must be *superior*: its income elasticity must be greater than one. However the reverse does *not* hold. If one good is superior this does not imply that another must be inferior. Note that the sum of the budget shares equals one.

Proof. If preferences are locally non-satiated, the budget constraint is satisfied with equality. Hence,

$$\sum_{i=1}^{k} p_i \cdot x_i(\mathbf{p}, m) = m$$

Differentiating both sides with respect to m,

$$\sum_{i=1}^{k} p_i \cdot \frac{\partial x_i(\mathbf{p}, m)}{\partial m} = 1$$

Multiplying each term by $\frac{m}{x_i(\mathbf{p},m)} \cdot \frac{x_i(\mathbf{p},m)}{m} = 1$ gives the desired result. □

Proposition 8.4.5 (Price Elasticities)**.** *The price elasticities of demand and relative budget shares for each good satisfy the following relationship for any $j \in 1, 2, \ldots, k$:*

$$\sum_{i=1}^{k} E_{x_i,p_j} \cdot \frac{p_i \cdot x_i(\mathbf{p}, m)}{p_j \cdot x_j(\mathbf{p}.m)}$$

Proof. If preferences are locally non-satiated, the budget constraint is satisfied with equality. Hence,

$$\sum_{i=1}^{k} p_i \cdot x_i(\mathbf{p}, m) = m$$

Differentiating both sides with respect to p_j, by the product rule:

$$x_j(\mathbf{p}, m) + p_j \cdot \frac{\partial x_j(\mathbf{p}, m)}{\partial p_j} + \sum_{i \neq j} p_i \cdot \frac{\partial x_i(\mathbf{p}, m)}{\partial p_j} = 0$$

Rearranging and dividing through by $x_j(\mathbf{p}, m)$,

$$\sum_{i=1}^{k} \frac{\partial x_i(\mathbf{p}, m)}{\partial p_j} \cdot \frac{p_i}{x_j(\mathbf{p}, m)} = -1$$

Multiplying each term by $\frac{p_j}{x_i(\mathbf{p},m)} \cdot \frac{x_i(\mathbf{p},m)}{p_j}$ gives the desired result. □