CONSUMER CHOICE

1. THE CONSUMER CHOICE PROBLEM

1.1. Unit of analysis and preferences. The fundamental unit of analysis in economics is the economic agent. Typically this agent is an individual consumer or a firm. The agent might also be the manager of a public utility, the stockholders of a corporation, a government policymaker and so on.

The underlying assumption in economic analysis is that all economic agents possess a preference ordering which allows them to rank alternative states of the world.

The behavioral assumption in economics is that all agents make choices consistent with these underlying preferences.

1.2. Definition of a competitive agent. A buyer or seller (agent) is said to be competitive if the agent assumes or believes that the market price of a product is given and that the agent’s actions do not influence the market price or opportunities for exchange.

1.3. Commodities. Commodities are the objects of choice available to an individual in the economic system. Assume that these are the various products and services available for purchase in the market. Assume that the number of products is finite and equal to L (ℓ = 1, ..., L). A product vector is a list of the amounts of the various products:

\[
x = \begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_L
\end{bmatrix}
\]

The product bundle x can be viewed as a point in \( R^L \).

1.4. Consumption sets. The consumption set is a subset of the product space \( R^L \), denoted by \( X^L \subset R^L \), whose elements are the consumption bundles that the individual can conceivably consume given the physical constraints imposed by the environment. We typically assume that the consumption set is \( X = R^L_+ = \{ x \in R^L : x_\ell \geq 0 \text{ for } \ell = 1, ..., L \} \).

1.5. Prices. We will assume that all L products are traded in the market at dollar prices that are publicly quoted. How they are determined will be discussed later. The prices are represented by a price vector

\[
p = \begin{bmatrix}
p_1 \\
p_2 \\
\vdots \\
p_L
\end{bmatrix} \in R^L
\]

For now assume that all prices are strictly positive, i.e. \( p_\ell >> 0 \). We will also assume that all consumers are price takers in the sense that they cannot influence the price at which they buy or sell a product.
1.6. Income or wealth. Assume that each consumer has wealth equal to \( m_i \) or the representative consumer has wealth \( m \).

1.7. Affordable consumption bundles. We say that a consumption bundle \( x \) is affordable for the representative consumer if

\[
p x = p_1 x_1 + p_2 x_2 + \cdots + p_L x_L \leq m
\]

If \( x \) is also an element of \( R^L_+ \), then the set of feasible consumption bundles is \( x \in R^L_+ : p x \leq m \). This is called a Walrasian budget set and is denoted \( B_{p,m} \).

1.8. Preferences. We assume a preference relation over products \( \succeq \) with the following properties

1: complete in that for all \( x_1, x_2 \in X \) we have \( x_1 \succeq x_2 \) or \( x_2 \succeq x_1 \) (or both)
2: transitive in that \( \forall x_1, x_2, x_3 \in X \), if \( x_1 \succeq x_2 \) and \( x_2 \succeq x_3 \) then \( x_1 \succeq x_3 \).
3: locally nonsatiated in that for every \( x_1 \in X \) and every \( \epsilon > 0 \), there is \( x_2 \in X \) such that \( ||x_2 - x_1 \leq \epsilon || \) and \( x_2 \succeq x_1 \).
4: continuous in that for any sequence of pairs

\[
\{ (x_1^n, x_2^n) \}_{n=1}^{\infty} \text{ with } x_1^n \succeq x_2^n \forall n,
\]

\[
x_1 = \lim_{n \to \infty} x_1^n, \text{ and } x_2 = \lim_{n \to \infty} x_2^n,
\]

we have \( x_1 \succeq x_2 \).

1.9. Existence of a utility function. Based on the preferences defined in 1.8, there exists a continuous utility function \( v(x) \) that represents \( \succeq \) in the sense that \( x_1 \succeq x_2 \) iff \( v(x_1) \geq v(x_2) \).

1.10. Convexity. We often assume that preferences are convex in the sense that if \( x_1 \succeq x_2 \), then for \( 0 \leq \lambda \leq 1 \), \( \lambda x_1 + (1-\lambda)x_2 \succeq x_1 \). This implies that indifference curves are convex to the origin. If the utility function is quasi-concave, then the indifference curves will be convex and vice versa.

2. The Utility Maximization Problem

2.1. Formal Problem. The utility maximization problem for the consumer is then as follows

\[
\max_{x \geq 0} u = v(x)
\]

s.t. \( p x \leq m \)

where we assume that \( p >> 0, m > 0 \) and \( X = R^L_+ \).

This is called the primal preference problem. If we have smooth convex indifference curves and an interior solution, then the problem can be solved using standard Lagrangian techniques. Alternatively, Kuhn-Tucker methods can be used. The Lagrangian function is given by

\[
\mathcal{L} = v(x) - \lambda (\Sigma_{i=1}^n p_i x_i - m)
\]

The first order conditions are

\[
\frac{\partial v}{\partial x_i} - \lambda p_i = 0, \quad i = 1, 2, \ldots, n
\]

\[
- \Sigma_{i=1}^n p_i x_i + m = 0
\]

The value of \( \lambda \) is the amount by which \( \mathcal{L} \) would increase given a unit relaxation in the constraint (an increase in income). It can be interpreted as the marginal utility of expenditure. This units of this are of
course arbitrary. The solution to 2 is given by \( x(p,m) = g(p,m) \). These functions are called Marshallian demand equations. Note that they depend on the prices of all good and income. Based on the structure of preferences and the nature of the optimization problem, they will have certain properties which we will discuss shortly.

2.2. **Cobb-Douglas Example.** Consider a utility function given by

\[
u = v(x) = x_1^{\alpha_1} x_2^{\alpha_2}
\]

We usually assume that \( \alpha_i > 0 \). To maximize utility subject to a budget constraint we obtain we set up a Lagrangian function.

\[
\mathcal{L} = x_1^{\alpha_1} x_2^{\alpha_2} - \lambda \left[ p_1 x_1 + p_2 x_2 - m \right]
\]

Differentiating equation 6 we obtain

\[
\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\alpha_1 x_1^{\alpha_1} x_2^{\alpha_2}}{x_1} - \lambda p_1 = 0 \quad (7a)
\]

\[
\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2}}{x_2} - \lambda p_2 = 0 \quad (7b)
\]

\[
\frac{\partial \mathcal{L}}{\partial \lambda} = - \left[ p_1 x_1 + p_2 x_2 \right] + m = 0 \quad (7c)
\]

Take the ratio of the 7a and 7b to obtain

\[
\frac{\frac{\alpha_1 x_1^{\alpha_1} x_2^{\alpha_2}}{x_1}}{\frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2}}{x_2}} = \frac{p_1}{p_2}
\]

\[
\Rightarrow \frac{\alpha_1 x_2}{\alpha_2 x_1} = \frac{p_1}{p_2} \quad (8)
\]

We can now solve the equation for the quantity of good 2 as a function of the quantity of good 1 and the prices of both goods. Doing so we obtain

\[
x_2 = \frac{\alpha_2 x_1 p_1}{\alpha_1 p_2} \quad (9)
\]

Now substitute 9 in 7c to obtain

\[
\frac{\partial \mathcal{L}}{\partial \lambda} = - \left[ p_1 x_1 + p_2 x_2 \right] + m = 0
\]

\[
\Rightarrow p_1 x_1 + p_2 \frac{\alpha_2 x_1 p_1}{\alpha_1 p_2} = m
\]

\[
\Rightarrow \frac{p_1 \alpha_1 x_1}{\alpha_1} + \frac{\alpha_2 x_1 p_1}{\alpha_1} = m
\]

\[
\Rightarrow \frac{p_1 x_1}{\alpha_1} (\alpha_1 + \alpha_2) = m
\]

\[
\Rightarrow x_1 = \frac{\alpha_1 m}{(\alpha_1 + \alpha_2) \frac{p_1}{p_1}}
\]

Similarly for \( x_2 \) so that we have
CONSUMER CHOICE

\[ x_2 = \frac{\alpha_2 m}{(\alpha_1 + \alpha_2) p_2} \]  (11)

Note that demand for the kth good only depends on the kth price and is homogeneous of degree zero in prices and income. Also note that it is linear in income. This implies that the expenditure elasticity is equal to 1. This can be seen as follows.

\[ x_1 = \frac{\alpha_1 m}{(\alpha_1 + \alpha_2) p_1} \]

\[ \Rightarrow \frac{\partial}{\partial m} \frac{m}{x_1} = \left[ \frac{\alpha_1}{(\alpha_1 + \alpha_2) p_1} \right] \left[ \frac{m}{(\alpha_1 + \alpha_2) p_1} \right] = 1 \]  (12)

We can find the value of the optimal \( u \) by substitution

\[ u = x_1^{\alpha_1} x_2^{\alpha_2} \]

\[ = \left( \frac{m}{p_1 \left( \frac{\alpha_1}{(\alpha_1 + \alpha_2)} \right)} \right)^{\alpha_1} \left( \frac{m}{p_2 \left( \frac{\alpha_2}{(\alpha_1 + \alpha_2)} \right)} \right)^{\alpha_2} \]  (13)

This can also be written

\[ u = x_1^{\alpha_1} x_2^{\alpha_2} \]

\[ = \left[ \frac{m}{p_1 \left( \frac{\alpha_1}{(\alpha_1 + \alpha_2)} \right)} \right]^{\alpha_1} \left[ \frac{m}{p_2 \left( \frac{\alpha_2}{(\alpha_1 + \alpha_2)} \right)} \right]^{\alpha_2} \]  (14)

For future reference note that the derivative of the optimal \( u \) with respect to \( m \) is given by

\[ \frac{\partial u}{\partial m} = (\alpha_1 + \alpha_2) m^{\alpha_1 + \alpha_2 - 1} p_1^{-\alpha_1} p_2^{-\alpha_2} \alpha_1^{\alpha_1} \alpha_2^{\alpha_2} (\alpha_2 + \alpha_2)^{-\alpha_1 - \alpha_2} \]  (15)

We obtain \( \lambda \) by substituting in either the first or second equation as follows

\[ \alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} - \lambda p_1 = 0 \]

\[ \Rightarrow \lambda = \frac{\alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2}}{p_1} \]  (16)

\[ \alpha_2 x_1^{\alpha_1} x_2^{\alpha_2 - 1} - \lambda p_2 = 0 \]

\[ \Rightarrow \lambda = \frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2 - 1}}{p_2} \]
If we now substitute for $x_1$ and $x_2$, we obtain

$$\lambda = \frac{\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2}}{p_1}$$

$$x_1 = \frac{m}{p_1} \left[ \frac{\alpha_1}{\alpha_1 + \alpha_2} \right]$$

$$x_2 = \frac{m}{p_2} \left[ \frac{\alpha_2}{\alpha_1 + \alpha_2} \right]$$

$$\Rightarrow \lambda = \frac{\alpha_1 \left( \frac{m}{p_1} \left[ \frac{\alpha_1}{\alpha_1 + \alpha_2} \right] \right)^{\alpha_1-1} \left( \frac{m}{p_2} \left[ \frac{\alpha_2}{\alpha_1 + \alpha_2} \right] \right)^{\alpha_2}}{p_1}$$

$$= \frac{\alpha_1 m^{\alpha_1 + \alpha_2 - 1} p_1^{1 - \alpha_1} p_2^{\alpha_2} \alpha_1^{\alpha_1 - 1} \alpha_2^{\alpha_2} (\alpha_1 + \alpha_2)^{1 - \alpha_1 - \alpha_2}}{p_1}$$

$$= m^{\alpha_1 + \alpha_2 - 1} p_1^{-\alpha_1} p_2^{-\alpha_2} \alpha_1^{\alpha_1} \alpha_2^{\alpha_2} (\alpha_1 + \alpha_2)^{1 - \alpha_1 - \alpha_2}$$

Thus $\lambda$ is equal to the derivative of the optimal $u$ with respect to $m$.

To check for a maximum or minimum we set up the bordered Hessian. The bordered Hessian in this case is

$$H_B = \begin{bmatrix}
\frac{\partial^2 L(x^*, \lambda^*)}{\partial x_1 \partial x_1} & \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_1 \partial x_2} & \frac{\partial g(x^*)}{\partial x_1} \\
\frac{\partial^2 L(x^*, \lambda^*)}{\partial x_2 \partial x_1} & \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_2 \partial x_2} & \frac{\partial g(x^*)}{\partial x_2} \\
\frac{\partial g(x^*)}{\partial x_1} & \frac{\partial g(x^*)}{\partial x_2} & 0
\end{bmatrix}$$

We compute the various elements of the bordered Hessian as follows
The derivatives of the constraints are constants. The bordered Hessian is given by

\[
H_B = \begin{bmatrix}
(\alpha_1)(\alpha_1 - 1)x_1^{\alpha_1 - 2}x_2^{\alpha_2} & \alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} & p_1 \\
\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} & (\alpha_2)(\alpha_2 - 1)x_1^{\alpha_1}x_2^{\alpha_2 - 2} & p_2 \\
p_1 & p_2 & 0
\end{bmatrix}
\]

(19)

To find the determinant of the bordered Hessian expand by the third row as follows

\[
|H_B| = (-1)^4 p_1 \left| \begin{array}{ccc}
\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} & p_1 \\
(\alpha_2)(\alpha_2 - 1)x_1^{\alpha_1}x_2^{\alpha_2 - 2} & p_2
\end{array} \right| + (-1)^5 p_2 \left| \begin{array}{ccc}
(\alpha_1)(\alpha_1 - 1)x_1^{\alpha_1 - 2}x_2^{\alpha_2} & p_1 \\
\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} & p_2
\end{array} \right| + 0
\]

\[
= p_1 \left| \begin{array}{ccc}
\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} & p_1 \\
(\alpha_2)(\alpha_2 - 1)x_1^{\alpha_1}x_2^{\alpha_2 - 2} & p_2
\end{array} \right| - p_2 \left| \begin{array}{ccc}
(\alpha_1)(\alpha_1 - 1)x_1^{\alpha_1 - 2}x_2^{\alpha_2} & p_1 \\
\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} & p_2
\end{array} \right|
\]

(20)

\[
= p_1p_2\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} - p_1^2(\alpha_2)(\alpha_2 - 1)x_1^{\alpha_1}x_2^{\alpha_2 - 2} - p_2^2(\alpha_1)(\alpha_1 - 1)x_1^{\alpha_1 - 2}x_2^{\alpha_2} + p_1p_2\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1}
\]

\[
= 2p_1p_2\alpha_1\alpha_2x_1^{\alpha_1 - 1}x_2^{\alpha_2 - 1} - p_1^2(\alpha_2)(\alpha_2 - 1)x_1^{\alpha_1}x_2^{\alpha_2 - 2} - p_2^2(\alpha_1)(\alpha_1 - 1)x_1^{\alpha_1 - 2}x_2^{\alpha_2}
\]

For a maximum we want equation 20 to be positive. Rewriting it we obtain
\[ 2p_1p_2\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2-1} - p_1^2(\alpha_2)(\alpha_2 - 1)x_1^{\alpha_1} x_2^{\alpha_2-2} - p_2^2(\alpha_1)(\alpha_1 - 1)x_1^{\alpha_1-2} x_2^{\alpha_2} > 0 \]  

(21)

We can also write it in the following convenient way

\[ 2p_1p_2\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2-1} + \alpha_2 p_1^2 x_1^{\alpha_1} x_2^{\alpha_2-2} - \alpha_2^2 p_1^2 x_1^{\alpha_1} x_2^{\alpha_2-2} + \alpha_1 p_2^2 x_1^{\alpha_1-2} x_2^{\alpha_2} > 0 \]

(22)

To eliminate the prices we can substitute from the first-order conditions.

\[ p_1 = \frac{\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2}}{\lambda} \]

\[ p_2 = \frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2-1}}{\lambda} \]

This then gives

\[ 2 \left( \frac{\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2}}{\lambda} \right) \left( \frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2-1}}{\lambda} \right) \alpha_1 \alpha_2 x_1^{\alpha_1-1} x_2^{\alpha_2-1} \]

\[ + \alpha_2 \left( \frac{\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2}}{\lambda} \right)^2 x_1^{\alpha_1} x_2^{\alpha_2-2} - \alpha_2 \left( \frac{\alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2}}{\lambda} \right)^2 x_1^{\alpha_1-2} x_2^{\alpha_2-2} \]

\[ + \alpha_1 \left( \frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2-1}}{\lambda} \right)^2 x_1^{\alpha_1-2} x_2^{\alpha_2} - \alpha_1 \left( \frac{\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2-1}}{\lambda} \right)^2 x_1^{\alpha_1-2} x_2^{\alpha_2} > 0 \]

(23)

Multiply both sides by \( \lambda^2 \) and combine terms to obtain

\[ 2\alpha_1^2 \alpha_2^2 x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} \]

\[ + \alpha_2^2 \alpha_1 x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} - \alpha_2^2 \alpha_1 x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} \]

\[ + \alpha_1^2 \alpha_2 x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} - \alpha_1^2 \alpha_2 x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} > 0 \]

(24)

Now factor out \( x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} \) to obtain

\[ x_1^{3\alpha_1-2} x_2^{3\alpha_2-2} (2\alpha_1^2 \alpha_2^2 + \alpha_2^2 \alpha_1 - \alpha_2^2 \alpha_1 + \alpha_1^2 \alpha_2 - \alpha_1^2 \alpha_2) > 0 \]

(25)

With positive values for \( x_1 \) and \( x_2 \) the whole expression will be positive if the last term in parentheses is positive. Then rewrite this expression as

\[ (\alpha_1^2 \alpha_2 + \alpha_1 \alpha_2^2) > 0 \]

(26)

Now divide both sides by \( \alpha_1^2 \alpha_2 \) (which is positive) to obtain

\[ \left( \frac{1}{\alpha_2} + \frac{1}{\alpha_1} \right) > 0 \]

(27)
3. THE EXPENDITURE (COST) MINIMIZATION PROBLEM

3.1. Basic duality formulation. The fundamental (primal) utility maximization problem is given by

$$\max_{x \geq 0} u = v(x)$$

s.t. $px \leq m$  \hspace{1cm} (28)

Dual to the utility maximization problem is the cost minimization problem

$$\min_{x \geq 0} m = px$$

s.t. $v(x) = u$  \hspace{1cm} (29)

3.2. Marshallian and Hicksian demand functions. The solution to equation 29 gives the Hicksian demand functions $x = h(u, p)$. The Hicksian demand equations are sometimes called "compensated" demand equations because they hold $u$ constant. The solutions to the primal and dual problems coincide in the sense that

$$x = g(p, m) = h(u, p)$$  \hspace{1cm} (30)

3.3. Indirect objective functions. We can substitute the optimal levels of the decision variables as functions of the parameters back into the objective functions to obtain the indirect objective functions. For the primal problem this gives

$$u = v(x_1, x_2, \ldots, x_n) = v[g_1(m, p), g_2(m, p), \ldots, g_n(m, p)] = \psi(m, p)$$  \hspace{1cm} (31)

This is called the indirect utility function and specifies utility as a function of prices and income. We can also write it as follows

$$\psi(m, p) = \max_x [v(x) : px = m]$$  \hspace{1cm} (32)

The indirect utility function for the Cobb-Douglas utility function is given by

$$u = x_1^{\alpha_1} x_2^{\alpha_2}$$

$$= \left( \frac{m}{p_1} \left( \frac{\alpha_1}{\alpha_1 + \alpha_2} \right) \right)^{\alpha_1} \left( \frac{m}{p_2} \left( \frac{\alpha_2}{\alpha_1 + \alpha_2} \right) \right)^{\alpha_2}$$

$$= \left( \frac{\alpha_1}{\alpha_1 + \alpha_2} \right)^{\alpha_1} \left( \frac{\alpha_2}{\alpha_1 + \alpha_2} \right)^{\alpha_2} \left( \frac{m}{p_1} \right)^{\alpha_1} \left( \frac{m}{p_2} \right)^{\alpha_2}$$  \hspace{1cm} (33)

For the dual problem the indirect objective function is

$$m = \Sigma_{j=1}^{n} p_j h_j(u, p) = c(u, p)$$  \hspace{1cm} (34)

This is called the cost (expenditure) function and specifies cost or expenditure as a function of prices and utility. We can also write it as follows

$$c(u, p) = \min_x [px : v(x) = u]$$  \hspace{1cm} (35)
3.4. Inversion of $\psi(m, p)$ and $c(u, p)$. Because $c(u, p) = m$, we can rearrange or invert it to obtain $u$ as a function of $m$ and $p$. This will give $\psi(m, p)$. Similarly inversion of $\psi(m, p)$ will give $c(u, p)$. These relationships between the utility maximization cost minimization problems are summarized in figure 1.

**Figure 1. Utility Maximization and Cost Minimization**

![Diagram of utility maximization and cost minimization](image)

3.5. Properties of the cost function. The cost function in the consumer problem has a number of analogous to its properties in the production problem.

3.5.1. C.1. The cost function is nondecreasing in $p$, increasing in $u$, and increasing in at least one $p$.

Let $p^1 \geq p^2$. Let $x^1$ be the cost minimizing input bundle with $p^1$ and $x^2$ be the cost minimizing input bundle with $p^2$. Then $p^2 x^2 \leq p^2 x^1$ because $x^1$ is not cost minimizing with prices $p^2$. Now $p^1 x^1 \geq p^2 x^1$ because $p^1 \geq p^2$ by assumption so that

$$C(p^1, y) = p^1 x^1 \geq p^2 x^1 \geq p^2 x^2 = C(p^2, y)$$

Nonsatiation guarantees that the function will be increasing in $u$. Let $V(u_0)$ be the set of all bundles that are equivalent to or preferred to bundles that provide utility level $u_0$. Now let $u_1 \leq u_2$. Because $V(u_1)$ is a subset of $V(u_2)$ if $u_1 \geq u_2$ then
CONSUMER CHOICE

\[ C(u_1, p) = \min_x \{ px : x \in V(u_1) \} \geq \min_x \{ px : x \in V(u_2) \} = c(u_2, p) \]

The point is that if we have a smaller set of possible x’s to choose from then cost must increase.

3.5.2. C.2. Positively linearly homogenous in p

\[ C(u, \theta p) = \lambda C(u, p), \quad p > 0. \]

Let the cost minimization problem with prices p be given by

\[ C(u, p) = \min_x \{ px : x \in V(u) \}, \quad u \in \text{Dom} V, p > 0, \tag{36} \]

where

\[ \text{Dom} V = \{ u \in R^1_+ : V(u) \neq \emptyset \} \]

The x vector that solves this problem will be a function of u and p, and is usually denoted h(u,p). This is the Hicksian demand function. The cost function is then given by

\[ C(u, p) = ph(u, p) \tag{37} \]

Now consider the problem with prices tp (p > 0)

\[ \hat{C}(y, tp) = \min_x \{ tp x : x \in V(y) \}, \quad u \in \text{Dom} V, p > 0 \]

\[ = t \min_x \{ px : x \in V(y) \}, \quad y \in \text{Dom} V, p > 0 \tag{38} \]

The x vector that solves this problem will be the same as the vector which solves the problem in equation 36, i.e., h(u,p). The cost function for the revised problem is then given by

\[ \hat{C}(p, tp) = tp h(u, p) = tC(u, p) \tag{39} \]

3.5.3. C.3. C is concave and continuous in w

To demonstrate concavity let (p, x) and (p', x') be two cost-minimizing price-consumption combinations and let p''= tp + (1-t)p' for any 0 \leq t \leq 1. Concavity implies that C(u, p'') \geq tC(u, p) + (1-t) C(u, p'). We can prove this as follows.

We have that C(u, p'') = p''. x'' = tp' \cdot x'' + (1-t)p' \cdot x'' \tag{40}

where x'' is the optimal choice of x at prices p''. Because x'' is not necessarily the cheapest way to obtain utility level u at prices p' or p, we have p \cdot x'' \geq C(u, p) and p' \cdot x'' \geq C(u, p') so that by substitution C(u, p'') \geq tC(u, p) + (1-t) C(u, p').

The point is that if p \cdot x'' and p' \cdot x'' are each larger than the corresponding term is the linear combination then C(u, p'') is larger than the linear combination.

Rockafellar [10, p. 82] shows that a concave function defined on an open set (p > 0) is continuous.


3.6.1. Definition. If indifference curves are convex, the cost minimizing point is unique. Then we have

\[ \frac{\partial C(u, p)}{\partial p_i} = h_i(u, p) \tag{40} \]

which is a Hicksian Demand Curve.
3.6.2. Constructive proof using the envelope theorem. The cost minimization problem is given by

\[ C(y, w) = \min_x px : v(x) - u = 0 \]  (41)

The associated Lagrangian is given by

\[ L = px - \lambda(v(x) - u) \]  (42)

The first order conditions are as follows

\[ \frac{\partial L}{\partial x_i} = p_i - \lambda \frac{\partial v}{\partial x_i} = 0, \quad i = 1, \ldots, n \]  (43a)

\[ \frac{\partial L}{\partial \lambda} = -(vx - u) = 0 \]  (43b)

Solving for the optimal \( x \)'s yields

\[ x_i(u, p) = h_i(u, p) \]  (44)

with \( C(u, p) \) given by

\[ C(u, p) = px(u, p) = ph(u, p) \]  (45)

If we now differentiate 45 with respect to \( p_i \), we obtain

\[ \frac{\partial C}{\partial p_i} = \sum_{j=1}^{n} p_j \frac{\partial x_j(u, p)}{\partial p_i} + x_i(u, p) \]  (46)

From the first order conditions in equation 43a (assuming that the constraint is satisfied as an equality) we have

\[ p_j = \lambda \frac{\partial v}{\partial x_j} \]  (47)

Substitute the expression for \( p_j \) from equation 47 into equation 46 to obtain

\[ \frac{\partial C}{\partial p_i} = \sum_{j=1}^{n} p_j \frac{\partial v(x)}{\partial x_j} \frac{\partial x_j(u, p)}{\partial p_i} + x_i(u, p) \]  (48)

If \( \lambda > 0 \) then equation 43b implies \( [v(x)-u] = 0 \). Now differentiate equation 43b with respect to \( p_i \) to obtain

\[ \sum_{j=1}^{n} \frac{\partial v(x(u, p))}{\partial x_j} \frac{\partial x_j(u, p)}{\partial p_i} = 0 \]  (49)

which implies that the first term in equation 48 is equal to zero and that

\[ \frac{\partial C(u, p)}{\partial p_i} = x_i(u, p) \]  (50)

3.7. The cost function and Marshallian demand functions. If we substitute the indirect utility function in the Hicksian demand functions obtained via Shephard’s lemma in equation 40, we get \( x \) in terms of \( m \) and \( p \). Specifically

\[ x_i = x_i(u, p) = h_i(u, p) = h_i[\psi(m, p), p] = g_i(m, p) = x_i(m, p) \]  (51)
3.8. **Cobb-Douglas Example.** The utility function is given by

\[ v(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} x_i^{\alpha_i} = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \ldots \]  

(52)

First set up the Lagrangian problem

\[ \mathcal{L} = \sum_{i=1}^{n} p_i x_i - \lambda \left( \prod_{i=1}^{n} x_i^{\alpha_i} - u \right) \]  

(53)

The first order conditions are as follows

\[
\frac{\partial \mathcal{L}}{\partial x_i} = p_i - \lambda \left[ \alpha_i x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_{i-1}^{\alpha_{i-1}} x_{i+1}^{\alpha_{i+1}} \ldots \right] = 0, \quad i = 1, \ldots, n
\]  

(54a)

\[
\frac{\alpha_i y}{x_i} \lambda = 0, \quad i = 1, \ldots, n
\]  

(54b)

\[
\frac{\partial \mathcal{L}}{\partial \lambda} = - \prod_{i=1}^{n} x_i^{\alpha_i} + u = 0
\]  

(54c)

Taking the ratio of the ith and jth equations we obtain

\[
\frac{p_i}{p_j} = \frac{\alpha_i x_j}{\alpha_j x_i} \]  

(55)

We can now solve the equation for the jth quantity as a function of the ith quantity and the ith and jth prices. Doing so we obtain

\[
x_j = \frac{\alpha_j x_i p_i}{\alpha_i p_j} \]  

(56)

where we treat the first good asymmetrically and solve for each demand for a good as a function of the first. Now substituting in the utility function we obtain

\[
v = \prod_{j=1}^{n} x_j^{\alpha_j} = \prod_{j=1}^{n} \left( \frac{\alpha_j x_1 p_1}{\alpha_1 p_j} \right)^{\alpha_j} \]  

(57)

Because \( x_1, p_1 \) and \( \alpha_1 \) do not change with \( j \), they can be factored out of the product to obtain

\[
u = \left( \frac{x_1 p_1}{\alpha_1} \right)^{\alpha_1} \prod_{j=1}^{n} \left( \frac{\alpha_j}{p_j} \right)^{\alpha_j} \]  

(58)

We then solve this expression for \( x_1 \) as a function of \( u \) and the other \( x \)'s. To do so we divide both sides by the product term to obtain
\[ x_1 = \left( \frac{\alpha_1}{p_1} \right) \left( \frac{u}{\prod_{j=1}^{n} \left( \frac{\alpha_j}{p_j} \right)^{\alpha_j}} \right)^{\frac{1}{\sum_{j=1}^{n} \alpha_j}} \]  

(61)

Similarly for the other \( x_k \) so that we have

\[ x_k = \left( \frac{\alpha_k}{p_k} \right) \left( \frac{u}{\prod_{j=1}^{n} \left( \frac{\alpha_j}{p_j} \right)^{\alpha_j}} \right)^{\frac{1}{\sum_{j=1}^{n} \alpha_j}} \]  

(62)

Now if we substitute for the \( i \)th \( x \) in the cost expression we obtain

\[ C = \sum_{i=1}^{n} p_i \left( \frac{\alpha_i}{p_i} \right) \left( \frac{u}{\prod_{j=1}^{n} \left( \frac{\alpha_j}{p_j} \right)^{\alpha_j}} \right)^{\frac{1}{\sum_{j=1}^{n} \alpha_j}} \]  

(63)

3.9. The indirect utility function and Hicksian demands. If we substitute \( C(u,p) \) in the Marshallian demands, we get the Hicksian demand functions

\[ x_i = x_i(m, p) = g_i(m, p) = g_i[C(u, p), p] = h_i(u, p) = x_i(u, p) \]  

(64)

3.10. Roy’s identity. We can also rewrite Shephard’s lemma in a different way. First write the identity

\[ \psi(C(u, p), p) = u \]  

(65)

Then totally differentiate both sides of equation 65 with respect to \( p_i \), holding \( u \) constant as follows

\[ \frac{\partial \psi[C(u, p), p]}{\partial p_i} \frac{\partial C(u, p)}{\partial p_i} + \frac{\partial \psi[C(u, p), p]}{\partial p_i} \frac{\partial p_i}{\partial p_i} = 0 \]  

(66)

Rearranging we obtain
\[
\frac{\partial C(u,p)}{\partial p_i} = \frac{\frac{\partial \psi(C(u,p),p)}{\partial p_i}}{\frac{\partial \psi(C(u,p),p)}{\partial m}} = g_i(m,p)
\]

where the last equality follows because we are evaluating the indirect utility function at income level \( m \). Figure 2 makes these relationships clear.

**FIGURE 2. Demand, Cost and Indirect Utility Functions**

4. **Properties of Indirect Utility Function**

The indirect utility function has the following properties.

4.1. \( \psi(m,p) \) is nonincreasing in \( p \), that is if \( p' \geq p \), \( \psi(m,p') \leq \psi(m,p) \).

4.2. \( \psi(m,p) \) is nondecreasing in \( m \), that is if \( m' \geq m \), \( \psi(m',p) \geq \psi(m,p) \).

4.3. \( \psi(m,p) \) is homogeneous of degree 0 in \( (p, m) \) so that \( \psi(tm, tp) = \psi(m, p) \) for \( t > 0 \).

4.4. \( \psi(m,p) \) is quasiconvex in \( p \); that is \( \{ p : \psi(m,p) < \alpha \} \) is a convex set for all \( \alpha \).
4.5. \( \psi \). \( \psi(m,p) \) is continuous for all \( p > 0, m > 0 \).

5. DISCUSSION AND PROOFS OF PROPERTIES OF INDIRECT UTILITY FUNCTION

5.1. \( \psi \). \( \psi(m,p) \) is nonincreasing in \( p \), that is if \( p' \geq p \), \( \psi(m,p') \leq \psi(m,p) \).

If prices go up, indirect utility cannot increase. Let \( B = \{ x : px < m \} \) and \( B' = \{ x : p'x < m \} \) for \( p' > p \). Then \( B' \) is contained in \( B \). Therefore the maximum of \( v(x) \) over \( B \) is at least as great as the maximum of \( v(x) \) over \( B' \).

5.2. \( \psi \). \( \psi(m,p) \) is nondecreasing in \( m \), that is if \( m' \geq m \), \( \psi(m',p) \geq \psi(m,p) \).

If income goes up, indirect utility cannot decrease. Let \( B = \{ x : px < m \} \) and \( B' = \{ x : px < m' \} \) for \( m' > m \). Then \( B \) is contained in \( B' \). Therefore the maximum of \( v(x) \) over \( B' \) is at least as great as the maximum of \( v(x) \) over \( B \).

5.3. \( \psi \). \( \psi(m,p) \) is homogeneous of degree 0 in \( (p, m) \) so that \( \psi(tm,tp) = \psi(m,p) \) for \( t > 0 \).

This is called the absence of money illusion. If prices and income are multiplied by the same positive number, the budget set does not change. Specifically,

\[
\sum_{i=1}^{n} (tp_i)x_i = tm
\]

\[
\Rightarrow \sum_{i=1}^{n} p_i x_i = m
\]

With the same budget set, the utility maximization problem has the same solution.

5.4. \( \psi \). \( \psi(m,p) \) is quasiconvex in \( p \); that is \( \{ p : \psi(m,p) < \alpha \} \) is a convex set for all \( \alpha \).

Suppose \( p \) and \( p' \) are such that \( \psi(m,p) \leq \alpha, \psi(m,p') \leq \alpha \). Let \( p'' = tp + (1-t)p' \). We want to show that \( \psi(m,p'') \leq \alpha \). Define the budget sets:

\[
B = \{ x : px < m \}
\]

\[
B' = \{ x : p'x < m \}
\]

\[
B'' = \{ x : p''x < m \}
\]

We can show that any \( x \) in \( B' \) must be in either \( B \) or \( B' \); that is that \( B \cup B' \supset B'' \). Assume not; then \( x \) is such that \( tp + (1-t)p'x \leq m \), but \( px > m \) and \( p'x > m \). These two inequalities can be written as

\[
(1-t)p'x > (1-t)m
\]

Summing the two expressions in equation 68 we obtain

\[
 tp + (1-t)p'x > m
\]

But this contradicts the original assumption that \( x \) is in neither \( B \) or \( B' \).

Now note that
ψ(m, p′′) = \max_x v(x), \text{ such that } x \in B′′ \\
\leq \max_x v(x), \text{ such that } x \in (B \cup B'); \text{ because } B \cup B' \supset B'' \\
\leq \alpha \text{ because } \psi(m, p) \leq \alpha \text{ and } \psi(m, p') \leq \alpha

5.5. \psi(m, p) is continuous for all p > 0, m > 0.

By the theorem of the maximum (given below) \psi(m, p) is continuous for p > 0, m > 0. In the utility maximization problem, f(x, \lambda) in the theorem of the maximum is the utility function. It does not depend on \lambda. The constraint set is those values of x that are in the budget set as parameterized by p and m. So \lambda in this case is (p, m). The indirect utility function \psi(m, p) is M(\lambda) while the ordinary demand functions x(m, p) are m(\lambda). The utility function is continuous by assumption. The constraint set is closed. If p > 0 and m > 0, the constraint set will be bounded. If some price were zero, the consumer might want to consume infinite amounts of this good. We rule that out.

**Theorem 1** (Theorem of the Maximum). Let f(x, \lambda) be a continuous function with a compact range and suppose that the constraint set \gamma(\lambda) is a non-empty, compact-valued, continuous correspondence of \lambda. Then

(i) The function M(\lambda) = \max_x \{ f(x, \lambda) : x \in \gamma(\lambda) \} is continuous \\
(ii) The correspondence m(\lambda) = \{ x \in \gamma(\lambda) : f(x, \lambda) = M(\lambda) \} is nonempty, compact valued and upper semi-continuous.

**Proof:** See Berge [1, p. 116].

6. Money Metric Utility Functions

6.1. Definition of m(p, x). Assume that the consumption set X is closed, convex, and bounded from below. The common assumption that the consumption set is X = \mathbb{R}^L_+ = \{ x \in \mathbb{R}^L : x_\ell \geq 0 \text{ for } \ell = 1, \ldots, L \} is more than sufficient for this purpose. Assume that the preference ordering satisfies the properties given in section 1.8. Then for all x \in X, let BT(x) = \{ y \in BT | y \succeq x \}. For the price vector p, the money metric m(p, x) is defined by

\[ m(p, x) = \min_{y \geq 0} py \text{ s.t. } y \in BT(x) \]  \hspace{1cm} (69)

If p is strictly greater than zero and if x is a unique element of the least cost commodity bundles at prices p, then m(p, x) can be viewed as a utility function for a fixed set of prices. It can also be defined as follows.

\[ m(p, x) = C(u(x), p) \] \hspace{1cm} (70)

The money metric defines the minimum cost of buying bundles as least as good as x. Consider figure 3. All points on the indifference curve passing through x will be assigned the same level of m(p, x), and all points on higher indifference curves will be assigned a higher level.

6.2. Example of a money metric utility function. Consider the Cobb-Douglas utility function

\[ v(x_1, x_2, \ldots, x_n) = A \prod_{i=1}^n x_i^{\alpha_i} = A x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \cdots \] \hspace{1cm} (71)

The cost function associated with this utility function is given in equation 63, which we repeat here.
To obtain the money metric we replace $v$ in equation 72 with $v(x)$ from equation 71.
\[ C = (\Sigma_{i=1}^{n} \alpha_i) v \prod_{j=1}^{n} \left( \frac{p_j}{\alpha_j} \right)^{\alpha_j} \]

\[ \Rightarrow m(p, x) = (\Sigma_{i=1}^{n} \alpha_i) \left( \prod_{i=1}^{n} x_i^{\alpha_i} \right) \left( \prod_{j=1}^{n} \left( \frac{p_j}{\alpha_j} \right)^{\alpha_j} \right)^{\frac{1}{\sum_{j=1}^{n} \alpha_j}} \]  

\[ (73) \]

**7. Properties of Demand Functions**

Demand functions have the following properties

7.1. **Adding up or Walras law.**

\[ \sum_{i=1}^{n} p_i h(u, p) = \sum_{i=1}^{n} p_i g_i(m, p) = m \]  

(74)

7.2. **Homogeneity.**

\[ h_i(u, \theta p) = h_i(u, p) = g_i(\theta m, \theta p) = g_i(m, p) \]  

(75)

The Hicksian demands are derivatives of a function that is homogeneous of degree one, so they are homogeneous of degree zero. Euler’s theorem then implies that

\[ \sum_{j=1}^{n} \frac{\partial h_j(u, p)}{\partial p_j} p_j = 0 \]  

(76)

If all prices and income are multiplied by a constant \( t > 0 \), the budget set does not change and so the optimal levels of \( x(m, p) \) do not change. We can also write this in differential form using the Euler equation.

\[ \sum_{j=1}^{n} \frac{\partial g_j(m, p)}{\partial p_j} p_j + \frac{\partial g_j(m, p)}{\partial m} m = 0 \]

\[ \Rightarrow \sum_{j=1}^{n} \frac{\partial g_j(m, p)}{\partial p_j} p_j = -\frac{\partial g_j(m, p)}{\partial m} m \]  

(77)

7.3. **Symmetry.** The cross price derivatives of the Hicksian demands are symmetric, that is, for all \( i \neq j \)

\[ \frac{\partial h_j(u, p)}{\partial p_i} = \frac{\partial h_i(u, p)}{\partial p_j} \]  

(78)

This is clear from the definition of the Hicksian demands as derivatives of the cost function. Specifically,

\[ \frac{\partial C(u, p)}{\partial p_i} = h_i(u, p) \]  

(79)

so that

\[ \frac{\partial^2 C(u, p)}{\partial p_j \partial p_i} = \frac{\partial h_i(u, p)}{\partial p_j} \quad \text{and} \quad \frac{\partial^2 C(u, p)}{\partial p_i \partial p_j} = \frac{\partial h_j(u, p)}{\partial p_i} \]  

(80)
by Young’s theorem on the equality of cross-partials.

7.4. **Negativity.** The nxn matrix formed by the elements $\frac{\partial h_i(u, p)}{\partial p_j}$ is negative semi-definite, that is, for any vector $z$, the quadratic form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} z_i z_j \frac{\partial h_i(u, p)}{\partial p_j} \leq 0 \quad (81)$$

If the vector $z$ is proportional to $p$, then the inequality becomes an equality and the quadratic form is zero. This means the matrix is negative semi-definite. This follows from the concavity of the cost function. If we denote $\frac{\partial h_i(u, p)}{\partial p_j}$ by $s_{ij}$, then we can write the entire matrix of cross partial derivatives as $S = s_{ij}$. This then implies that

$$z' Sz \leq 0 \quad (82)$$

By the properties of a negative semi-definite matrix, this means that $s_{ii} \leq 0$, or that the Hicksian demand functions have a slope which is non-positive. This follows from concavity of cost, and does not require convex indifference curves.

7.5. **The Slutsky equation.** If we differentiate equation 64 with respect to $p_j$ and then substitute from Shephard’s lemma for $\frac{\partial C(u, p)}{\partial p_j}$, we obtain

$$x_i = x_i(m, p) = g_i[C(u, p), p] = h_i(u, p) = x_i(u, p)$$

$$\Rightarrow x_i(u, p) = h_i(u, p) = g_i[C(u, p), p]$$

$$\Rightarrow \frac{\partial x_i(u, p)}{\partial p_j} = s_{ij} = \frac{\partial h_i(u, p)}{\partial p_j} = \frac{\partial g_i(m, p)}{\partial m} \frac{\partial C(u, p)}{\partial p_j} + \frac{\partial g_i(m, p)}{\partial p_j} \frac{\partial C(u, p)}{\partial m} \frac{\partial p_j}{\partial p_j} \quad (83)$$

The last term in equation 83 is the uncompensated derivative of $x_i$ with respect to $p_j$. To compensate for this, an amount, $x_i$, times $\frac{\partial g_i}{\partial m}$ must be added on. We can also write equation 83 as follows

$$\frac{\partial x_i(m, p)}{\partial p_j} = \frac{\partial x_i(u, p)}{\partial p_j} - \frac{\partial x_i(m, p)}{\partial m} x_j(m, p) \quad (84)$$

The first term is called the substitution effect, the second term the income effect.

8. **Integrability**

In section 7 we determined that a system of demand equations satisfies

(i) Walras Law

(ii) Homogeneity of degree zero in prices and income

(iii) Symmetry

(iv) Negativity

The integrability question has to do with whether a system of equations satisfy these four properties whether there exists a utility function from which this system can be derived. The answer to this question is yes as indicated in the following theorem.
Theorem 2. A continuously differentiable function \( x \) which maps \( \mathbb{R}^n_+ \) into the real line is the demand function generated by some increasing, quasiconcave utility function if (and only if, when utility continuous, strictly increasing, and strictly quasiconcave) it satisfies Walras law, symmetry and negativity.

This question is typically posed in terms of the cost function. One shows that a system of Hicksian demand functions satisfying Walras law, symmetry and negativity has an associated cost function from which it can be derived. Once one has the cost function, it can be used to obtain the utility function. Shephard’s lemma (equation 40) states that
\[
\frac{\partial C(u, p)}{\partial p_i} = h_i(u, p)
\]

Ordinary (Marshallian demand equations can be written in terms of Hicksian demand equations (see equation 64) as follows
\[
x_i(m, p) = x_i(C(u, p), p) = h_i(u, p) = x_i(u, p)
\]

Combining equations 85 and 86 we can then write
\[
\frac{\partial C(u, p)}{\partial p_i} = h_i(u, p) = x_i(u, p) = x_i(C(u, p), p), \quad i = 1, 2, \ldots, k
\]

Now suppose we have a system of demand functions \( x_i(m, p) \) for \( i = 1, 2, \ldots, k \). Now pick some point \( x^0(m, p^0) \) and assign it an arbitrary level of utility \( u^0 \). Also assume that
\[
C(u^0, p^0) = p^0 x(m^0, p^0)
\]

If a cost function which generated this system of demand functions exists, then it must satisfy the system of partial differential equations given by
\[
\frac{\partial C(u^0, p)}{\partial p_i} = h_i(u^0, p) = x_i(C(u^0, p), p), \quad i = 1, 2, \ldots, k
\]

If this system of partial differential equations has a solution, then \( x(m, p) \) is the demand system generated the cost function \( C(u, p) \). In order to understand the conditions for such a system to have a solution, consider the derivative of equation 89 with respect to \( p_j \).
\[
\frac{\partial^2 C(u^0, p)}{\partial p_j \partial p_i} = \frac{\partial}{\partial p_j} \left( \frac{\partial C(u^0, p)}{\partial p_i} \right) = \frac{\partial x_i(C(u^0, p), p)}{\partial p_j} + \frac{\partial x_i(C(u^0, p), p)}{\partial p_j} + \frac{\partial}{\partial p_j} \left( \frac{\partial x_i(C(u^0, p), p)}{\partial p_j} \right)
\]

Substituting for \( \frac{\partial C(u^0, p)}{\partial p_j} \) using Shephard’s lemma and writing \( m \) in place of \( C \) we obtain
\[
\frac{\partial^2 C(u^0, p)}{\partial p_j \partial p_i} = \frac{\partial x_i(m, p)}{\partial m} x_j(m, p) + \frac{\partial x_i(m, p)}{\partial p_j}
\]

Notice that the left hand side of equation 91 is symmetric in \( i \) and \( j \) by Young’s theorem. But if the left hand side of 91 is symmetric the right hand side of 91 must also be symmetric. This implies that the right hand side of 91 being symmetric is a necessary condition for the system of partial differential equations in 89 to have a solution. Frobenius’ theorem states that symmetry of the right hand side of 91 is also sufficient for 89 to have a solution. Then notice that the right hand side of 91 is the Slutsky matrix associated with the demand system \( x(m, p) \). So symmetry of the Slutsky matrix is necessary and sufficient for a function \( C(u, p) \) to exist, from which \( x(p, m) \) can be derived. The question remaining is whether the function \( C(u, p) \) which solves 89 is a proper cost function.
We therefore need to verify that the properties of the cost function stated in subsection 3.5 hold for the function which solves the system in equation 89. These properties are

1. Nondecreasing in $p$, increasing in $u$, and increasing in at least one $p$.
2. Positively linearly homogenous in $p$
3. Concave and continuous in $w$

$C(u, p)$ as a solution to 89 will be nondecreasing in $p$ because Shephard’s lemma shows that the derivative of $C(u, p)$ with respect to any price is Hicksian demand which is nonnegative. The other properties in items 1 and 3 can be similarly shown. $C(u, p)$ will be concave if it has a Hessian matrix which is negative semi-definite. But the Hessian of $C(u, p)$ is just the Slutsky matrix. So if a system of demand equations has a negative semi-definite Slutsky matrix, then the solution to the system partial differential equations in 89 will be concave.

A continuously differentiable function $x$ which maps $\mathbb{R}^n_+$ into the real line is the demand function. The bottom line is that a system of demand functions that satisfies Walras law, symmetry and negativity is consistent with some increasing, quasiconcave utility function.
9. Some Notes on Functions, Correspondences, and Functional Structure

9.1. Functions. By a function we mean a rule that assigns to each element in a set $X$, a unique element \( f(x) \) in another set $Y$. Consider the set $X = \mathbb{R}^1$ and $Y = \mathbb{R}^1$ and the rule $f(x) = 3x$. For any real $x$, the function assigns a unique real number in the set $Y$. We often use the following notation for a function $f: X \to Y$

\[
f : X \to Y
\]

The set $X$ is called the domain of the function $f$. The set of values taken by $f$, that is, the set $y \in Y: (\exists x)$ \([y = f(x)]\) is called the range of $f$. The range of $f$ will generally be smaller than $Y$. Consider the case where $X$ is the rational numbers and $Y$ is the real numbers. The function $f(x) = 3x$ will not cover all members of $Y$. A function whose range is all of $Y$ is said to be onto $Y$. If $A$ is a subset of $X$, the image under $f$ of $A$ is defined to be the set of elements in $Y$ such that $y = f(x)$ for some $x$ in $A$. This is denoted as $f[A]$ and formally given by

\[
f [A] = \{ y \in Y : (\exists x) [x \in A \text{ and } y = f(x)] \}
\]

The function $f$ is onto $Y$ iff $Y = f[X]$. If $B$ is a subset of $Y$, the inverse image $f^{-1}[B]$ is the set of $x$ in $X$ for which $f(x)$ is in $B$. Formally

\[
f^{-1}[B] = \{ x \in X : f(x) \in B \}
\]

The function $f$ is onto $Y$ iff the inverse image of each nonempty set of $Y$ is nonempty. Consider again the example where $X$ is the rational numbers. There are elements of $Y$ such that the is no element of $X$ that could generate them under the function $f(x) = 3x$. A function $f: X \to Y$ is called one-to-one if $f(x_1) = f(x_2)$ only when $x_1 = x_2$. Consider the function mapping the real line into the real line where $f(x) = x^2$. The function is not one-to-one since $f(-3) = f(3)$. The function is also not onto since there is no real $x$ for which $f(x) = -25$.

9.2. Correspondences. Let $X$ and $Y$ be two sets. If with each element of $X$ we associate a subset $\Gamma(X)$ of $Y$, we say that the correspondence $x \to \Gamma(X)$ is a mapping of $X$ into $Y$. The set $\Gamma(x)$ is called the image of $x$ under the mapping $\Gamma$. If the set $\Gamma(x)$ always consists of a single element, we say that $\Gamma$ is a function. Consider as an example the case where $X$ is $\mathbb{R}^1_+$ and $Y$ is $\mathbb{R}^1$. Now consider the mapping $\Gamma(x) = y \in \mathbb{R}^1: y \leq -\sqrt{x}$. This is a correspondence from $X$ to $Y$.

9.3. Functional Structure. Functional structure has to do with the amount of information we have about a given mapping. Considered in a different fashion, it is about the number of constraints that we know are imposed on a mapping. For example if we know that $y = 3x$ for all elements $x \in X$, then we know everything there is to know about the mapping. Alternatively, the mapping is very constrained in that no other rule satisfies this mapping. Consider the case where $y = ax$, but $a$ is not known. In this case we know that the mapping is linear, but we cannot say much more than that about it. Consider a function which maps $\mathbb{R}^2$ into $\mathbb{R}^1$. Specifically consider the function $y = ax_1^2 + bx_2^2$. We know that the function is quadratic and that it has no linear or constant terms. It is also clear that it is homogeneous of degree 2 in the vector $x$ because $f(3x) = 9f(x)$. Thus the function is not completely general.

Functional structure relates to the way in which a mapping is constructed and the way in which the elements of the domain enter the mapping. Consider for example a general mapping from $\mathbb{R}^3$ to $\mathbb{R}^1$ that has the following form: $y = f(x_1, x_2, x_3) = g(x_1, x_2) + h(x_3)$. While this is not much information about the mapping, it rules out functions such as $y = 3x_1x_2 + 5x_1x_3 + 8x_2x_3$.

Functional structure is typically represented in two ways. The earliest and perhaps most common is through the use of differential constraints on the function. These constraints then imply something about
the parent function. The difficulty with this approach is that it requires differentiability. The other approach is to consider structure on the function directly as for example \( f(x) = F(h(x)) \) where \( h \) is required to be homogeneous in all the \( x \)'s.

9.4. **Functional Structure and Functional Equations.** Functional structure is often related to the solution of functional equations. Functional equations are equations in which the unknown (or unknowns) are functions. Such functions can be multi-place in the sense of having more than one argument and can also deal with several variables. The number of places in the equations need not equal the number of variables. For example, the famous Cauchy equation \( f(x+y) = f(x) + f(y) \) is a function with one place but two variables. A linear function such as \( f(z) = cz \) would satisfy the equation. As another example consider the cost function for a multiple output firm. Consider the restriction that the cost of producing the vector of outputs \((x+y)\) is the sum of the cost of producing either vector individually, i.e., \( C(w,x+y) = C(w,x) + C(w,y) \). The determination of the form that all cost functions satisfying this restriction must take involves solving the functional equation.
10. AGGREGATION ACROSS GOODS

10.1. Separability and Aggregation. Separability is related to the ability to aggregate variables in economic analysis. For example, is it reasonable to aggregate two or more types of cold cereal together in analyzing the demand for food. Or can the hours of male and female workers be added together for an analysis of productivity. Separability is specifically concerned with how the rate of substitution between two goods or factors is affected by levels of other goods or factors. For example the rate of substitution between beef and pork may or may not be affected by the amount of tofu consumed. If this rate is not affected by tofu consumption, then some types of aggregation of beef and pork may be possible. While beef and pork may not be separable from tofu in consumption, they may be separable from shirts. Separability is particularly important for aggregate analysis where inputs tend to come in generic bundles such as labor, capital and materials, and goods tend to come in bundles such as housing, food, transportation, entertainment and so on.

10.2. Differential definition of separability. Consider a function $f$ depending on $n$ variables with $f$ being twice differentiable and

$$\frac{\partial f}{\partial x_i} > 0$$  \hspace{1cm} (92)

Then variables $x_i$ and $x_j$ are separable from $x_k$ if and only if

$$\frac{\partial}{\partial x_k} \left( \frac{\partial f(x)}{\partial x_i} \right) = 0 \forall x \in \Omega^n$$ \hspace{1cm} (93)

$$\Omega^n = \{ (x_1, x_2, \ldots, x_n) = x \in \mathbb{R}^n \mid x \geq 0^n \text{ and } x \neq 0^n \}$$

This says that the marginal rate of substitution between $x_i$ and $x_j$ does not depend on the level of $x_k$. This definition is due to Leontief [7] and independently Sono [9].

10.3. Intuition of differential definition of separability. As $x_k$ changes, the indifference curve when projected into $x_i, x_j$ space will have the same slope at

$$(\bar{x}_i, \bar{x}_j)$$

In figure 4, changing $x_3$ to $\hat{x}_3$ does not change the set of $x_1$ and $x_2$ such that $u(x_1, x_2, x_3) \geq u(\bar{x}_1, \bar{x}_2, x_3)$. In figure 5, changing $x_3$ to $\hat{x}_3$ changes the set of $x_1$ and $x_2$ such that $u(x_1, x_2, x_3) \geq u(\bar{x}_1, \bar{x}_2, x_3)$. 
10.4. **Some simple implications of the definition of separability.** Carry out the differentiation implied in equation 93 to obtain

\[
\frac{\partial f}{\partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_k} - \frac{\partial f}{\partial x_i} \frac{\partial^2 f}{\partial x_j \partial x_k} = 0
\]

\[
\Rightarrow \frac{\partial^2 f}{\partial x_i \partial x_k} = \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j}
\]

\[
\Rightarrow \frac{f_{ik}}{f_{jk}} = \frac{f_i}{f_j}
\]

\[
\Rightarrow \frac{f_{ik}}{f_i} x_k = \frac{f_{jk}}{f_j} x_k
\]

\[
\Rightarrow \frac{\partial \ln f_i}{\partial \ln x_k} = \frac{\partial \ln f_j}{\partial \ln x_k}
\]

Thus the elasticity of marginal product of \(x_i\) with respect to \(x_k\) is equal to the elasticity of the marginal product of \(x_j\) with respect to \(x_k\).
10.5. **example.** Consider a Cobb-Douglas utility function which is separable.

\[
v = v(x_1, x_2, x_3) \\
= 10 x_1^{1/2} x_2^{1/4} x_3^{1/4}
\]  
(95)

Consider whether \( x_1 \) and \( x_2 \) are separable from \( x_3 \). Let the intimal reference vector be \( x = [1, 1, 1] \) and denote the utility of this bundle by \( v^* \).

\[
v^* = v(1, 1, 1) \\
= 10 \left( 1^{1/2} \right) \left( 1^{1/4} \right) \left( 1^{1/4} \right) \\
= 10
\]  
(96)

Now consider those bundles that are preferred to bundles with a utility level of 10 where \( x_3 = 1 \).

\[
B^{12}(x_1, x_2, 1) = [x_1, x_2 | 10 x_1^{1/2} x_2^{1/4} \geq 10] \\
= [x_1, x_2 | x_1^{1/2} x_2^{1/4} \geq 1]
\]  
(97)

Now consider a different reference vector with \( x_3 = 16 \), i.e., \( x = [1, 1, 16] \). The reference utility level in this case is given by
\[ v^* = v(1, 1, 16) \]
\[ = 10 \left( 1^{1/2} \right) \left( 1^{1/4} \right) \left( 16^{1/4} \right) \]
\[ = 20 \]  

Now consider those bundles that are preferred to bundles with a utility level of 20 where \( x_3 = 16 \).

\[ B^{12}(x_1, x_2, 16) = [x_1, x_2 | 10 x_1^{1/2} x_2^{1/4} 16^{1/4} \geq 20] \]
\[ = [x_1, x_2 | 20 x_1^{1/2} x_2^{1/4} \geq 20] \]
\[ = [x_1, x_2 | x_1^{1/2} x_2^{1/4} \geq 1] \]  

which is the same set as before.

10.6. **Functional structure.** Separability restrictions on the function \( f \) implies that the function cannot be completely general. Thus not all possible forms satisfy this restriction. For example if the restrictions in 93 were to hold for all possible choices of \( i \) and \( j \) the function would have to be of the form

\[ f(x) = f^* [\Sigma_{i=1}^n f^i(x_i)] \]  

where the functions \( f^i \) only depend on the level of \( x_i \) and \( f^* \) only depends on the sum of these functions. In this description \( f^i \) is called an aggregator and \( f^* \) is called a macro function. Many problems of this sort can be solved using the theory of functional equations. The specification of the form that \( f \) must take is based on what are called representation theorems.

10.7. **Representation Theorems.**

10.7.1. **Notation.** Consider a commodity space with \( n \) commodities. Let \( \Omega^n \) be the non-negative orthant of \( \mathbb{R}^n \) so that all commodity vectors \( x \in \Omega^n \). Then consider an index set \( I = \{1, 2, \ldots, n\} \) and a binary partition of this set denoted by \( \bar{I} \).

\[ \bar{I} = \{I^r, I^c\}, \text{ where} \]
\[ I^r \cup I^c = I \]
\[ I^r \cap I^c = \emptyset \]
\[ I^c \neq \emptyset \]  

\( x = (x^r, x^c) \) is a partitioned commodity vector

\( p = (p^r, p^c) \) is a partitioned commodity vector

10.7.2. **Theorem on Representation of Separability.**

**Theorem 3.** Assume \( v \) is continuous and nondecreasing. Then a partition \( I^r \) is strictly separable in \( v \) from its complement \( I^c \) iff there exist functions

\[ v^r : \Omega^r \rightarrow \mathbb{R}^1 \]
\[ \bar{v} : \Omega^c \times \rho(v^r) \rightarrow \mathbb{R}^1 \]  

such that
\[ v(x) = \bar{v}(v^r(x^r), x^c) \] (103)

where \( \rho(v^r) \) is the range of \( v^r \) and \( \bar{v} \) is nondecreasing in \( (v^r(x^r)) \).

### 10.7.3. Example of Theorem 3 with Differentiability

Let
\[ v(x) = \bar{v}^r(x^r), x^c \] (104)

and consider two elements in \( x^r \) denoted \( x_i \) and \( x_j \) and one element in \( x^c \) denoted in \( x_k \). Differentiating equation (104) with respect to \( x_i \) and \( x_j \) we obtain

\[ \frac{\partial v}{\partial x_i} = \frac{\partial \bar{v}}{\partial v^r} \frac{\partial v^r}{\partial x_i} \] (105a)
\[ \frac{\partial v}{\partial x_j} = \frac{\partial \bar{v}}{\partial v^r} \frac{\partial v^r}{\partial x_j} \] (105b)

Taking the ratio of (105a) and (105a) we obtain

\[ \frac{\partial v}{\partial x_i} \frac{\partial v}{\partial x_j} = \frac{\partial v^r}{\partial x_i} \frac{\partial v^r}{\partial x_j} = \gamma(x^r) \] (106)

The ratio only depends on commodities in the group designated by \( r \). Therefore

\[ \frac{\partial \gamma(x^r)}{\partial x_k} = 0 \] (107)

because there are no \( x_k \) in the expression \( \gamma(x^r) \). The aggregator \( v^r \) is typically constructed as

\[ v^r(x^r) = v(x^r, 0^c) \] (108)

where \( 0^c \) is an arbitrary reference vector.

Consider the example from subsection 10.5

\[ v = v(x_1, x_2, x_3) \]
\[ = 10 x_1^{1/2} x_2^{1/4} x_3^{1/4} \]

with \( I^r = \{1, 2\}, I^c = \{3\} \). Define

\[ v^r(x_1, x_2) = v(x_1, x_2, 16) \]
\[ = 10 x_1^{1/2} x_2^{1/4} \cdot 2 \]
\[ = 20 x_1^{1/2} x_2^{1/4} \]

Given that \( v(x_1, x_2, x_3) = \bar{v}(v^r(x_1, x_2), x_3) \) we must have

\[ 10 x_1^{1/2} x_2^{1/4} x_3^{1/4} = \bar{v} \left( (20 x_1^{1/2} x_2^{1/4}), x_3 \right) \]

One possible choice for \( \bar{v} \) would be

\[ \bar{v} = 1/2 \ v^r x_3^{1/4} \]
\[ = 10 x_1^{1/2} x_2^{1/4} x_3^{1/4} \]
10.7.4. Representation Theorems and Functional Structure.

**Theorem 4.** Suppose \( v \) is continuous and nondecreasing and satisfies the separability assumptions of Theorem 3. Then \( \bar{v} \) and \( v^r \) in the representation 104 can be so chosen that if \( v \) has any of the following properties, then \( \bar{v} \) and \( v^r \) possess the same property.

(i) strict quasi-concavity
(ii) strict positive monotonicity
(iii) positive homotheticity
(iv) positive linear homogeneity

If \( I^r \) is strictly separable from \( I^c \), then (directional) differentiability of \( f \) implies (partial) differentiability of \( \bar{v} \) and \( v^r \), and if \( v \) is (strictly) concave, \( v^r \) can be chosen to be (strictly) concave.

The implication of Theorem 4 is that if we impose structure on the function \( f \) then the macro function \( \bar{v} \) and the aggregator \( v^r \) have the same structure.

10.8. Importance of separability. The implication of separability is that we can create aggregates to represent separable groups. For instance in Theorem ?? if \( x^r \) is food and \( x^c \) is all goods then if separability conditions hold we can construct an index \( v^r(x^r) \) of food for in analyzing consumer behavior. The ability to test for separability before aggregating is then an important factor to consider in consumption analysis. Separability is also important in dividing decision making into multiple steps as in the two-stage budgeting procedure in consumer theory.

10.9. **Weak Separability.**

10.9.1. Definition of weak separability. Partition \( x \) into \( m \) mutually exclusive subsets with \( \tilde{I} = [I^1, I^2, \ldots, I^m] \) and \( x = [x^1, x^2, \ldots, x^m] \). Then weak separability implies that

\[
\frac{\partial}{\partial x_i^t} \left[ \frac{\partial v(x)}{\partial x_j^p} \right] = 0, \quad i \neq k, j, p = 1, 2, \ldots, n^i, t = 1, 2, \ldots, n^k
\]  

(109)

Here \( j \) and \( p \) index variables in the \( i \)th group while \( t \) indexes variables within the \( k \)th group. This simply implies that the MRS between pairs of elements in one group is unaffected by the level of variables in other groups. Note that rather than considering separability between groups and their complements we are considering separability with respect to a partition \( \tilde{I} \).

10.9.2. Implication of weak separability. If \( v(x) \) is weakly separable it can be represented as

\[
v(x) = \bar{v} \left( v^1(x^1), v^2(x^2), \ldots, v^m(x^m) \right)
\]  

(110)

\( \bar{v} \) is called the macro function and is increasing in its arguments \( v^i \) is the \( i \)th aggregator function or group utility function. \( v^i \) aggregates commodities in a group into an aggregate commodity. For a fuller discussion see Goldman and Uzawa [4, p. 390].

If the utility function can be written in the form 113, then the demand for the \( i \)th commodity in the \( j \)th group can be written as follows

\[
x_i^j = x_i^j(p^j, C^j)
\]  

(111)

where \( C^j = \sum_{k=1}^{n^j} p_k^j x_k^j \) is the expenditure committed to the \( j \)th group.

10.10. **Strong Separability.**
10.10.1. **Definition of strong separability.** A function is strongly separable if

\[
\frac{\partial}{\partial x_i^k} \left[ \frac{\partial f(x)}{\partial x_i^j} \frac{\partial f(x)}{\partial x_m^t} \right] = 0, \quad i \neq k, \ m \neq k, \ j = 1, 2, \ldots, n^i, \ p = 1, 2, \ldots, n^m, \ t = 1, 2, \ldots, n^k
\]  

(112)

Strong separability implies that MRS between any two variables (not necessarily in the same group) does not depend on the level of variables in a third group. Strong separability is only defined when the number of groups exceeds two.

10.10.2. **Implication of strong separability.** Strong separability implies weak separability and can be represented as

\[
v(x) = v^\ast \left[ \sum_{i=1}^{m^i} v^i(x^i) \right]
\]

For a fuller discussion see Goldman and Uzawa [4, p. 389] or Blackorby, Primont and Russell [2, p. 136].

10.11. **Homothetic Separability.**

10.11.1. **Definition of strong separability.** A function \( v \) is said to be homothetically separable if it is weakly separable in the partition \( I \) and each aggregator function \( v^r \) is homothetic.

10.11.2. **Implication of homothetic separability.** The overall utility function need not be homothetic but the aggregators for the groups must be homothetic or have straight-line expansion paths through the origin. Homothetic separability is a necessary and sufficient condition for two stage budgeting.

11. **Two-stage budgeting example**

Assume that a consumer is maximizing utility subject to a budget constraint. Further assume that the consumer makes consumption and expenditure decisions in two stages. In the first stage, the consumer minimizes the cost of consuming goods in the jth group (food, housing, transportation, clothing, entertainment, etc.) subject to a reference utility level for that group. These group cost functions can be used to obtain indirect utility functions for each group as a function of group expenditures and the prices of goods in the group. In the second stage the consumer maximizes utility (as a function of group indirect utility functions) subject to the constraint that the sum of expenditure on all groups \( \sum_{j=1}^{J} c^j(p^j, u^j) \) is equal to total income m.
11.1. Definitions.

\[ v(x) = v(x_1, x_2, \ldots, x_n) \]  
\[ = \tilde{v} \left( v^1(x^1), v^2(x^2), \ldots, v^J(x^J) \right) \]  
\[ = \prod_{r=1}^J v^r(x^r)^{\gamma_r} \]  

\[ c^r = c^r(p^r, u^r) = c^r(p_{1r}^r, p_{2r}^r, \ldots, p_{nr}^r, u_r) \]  
\[ = \sum_{k=1}^{n_r} p_{kr}^r x_{kr} \]  
\[ v^r = \prod_{i=1}^{n_r} x_i^{\alpha_i^r}, \sum_{k=1}^{n_r} \alpha_i^r = 1 \]  
\[ v(x) = \prod_{r=1}^J \left( \prod_{i=1}^{n_r} x_i^{\alpha_i^r} \right)^{\gamma_r} \]  
\[ = \prod_{r=1}^J \prod_{i=1}^{n_r} x_i^{\alpha_i^r \gamma_r} \]  

a: \( v(x) = v(x_1, x_2, \ldots, x_n) \) is a utility function defined over \( n \) goods.

b: \( v(x) = \tilde{v} \left( v^1(x^1), v^2(x^2), \ldots, v^J(x^J) \right) \) is a weakly separable representation of the utility function with aggregators \( v^i(x^i) \) and a macro function \( \tilde{v} \). The vector of goods in the \( i \)th group is denoted \( x_i \).

There are \( J \) groups.

c: \( v(x) = \prod_{r=1}^J v^r(x^r)^{\gamma_r} \) is a representation of the macro function.

d: \( c^r(p_{1r}^r, p_{2r}^r, \ldots, p_{nr}^r, u_r) \) is a cost function that denotes the minimum cost of obtaining utility level \( u_r \) with prices \( p_{kr}^r \) by choosing goods in the \( r \)th group, \( \{x_{1r}, x_{2r}, \ldots, x_{nr}\} \).

e: \( v^r = \prod_{i=1}^{n_r} x_i^{\alpha_i^r} \) is a sub-utility function that denotes the utility received from consuming goods in the \( r \)th group.

f: \( v(x) = \prod_{r=1}^J \prod_{i=1}^{n_r} x_i^{\alpha_i^r \gamma_r} \) is the overall utility function.

11.2. Finding the minimum cost of obtaining a group utility level. Set up the Lagrangian problem where the reference utility level is \( u^r \).

\[ L = \sum_{i=1}^{n_r} p_i x_i - \lambda \left( \prod_{i=1}^{n_r} x_i^{\alpha_i^r} - u^r \right) \]  

We will not superscript \( p_i \) and \( x_i \) with \( r \) given that we know the prices and quantities refer to the \( r \)th group. The first order conditions are as follows
\[
\frac{\partial L}{\partial x_i} = p_i - \lambda \left[ \alpha^r_1 x_1^{\alpha^r_1^+} x_2^{\alpha^r_2} \cdots x_{i-1}^{\alpha^r_{i-1}} x_i^{\alpha^r_i - 1} x_{i+1}^{\alpha^r_{i+1}} \cdots \right] = 0, \quad i = 1, \ldots, n_r
\]

\[
= p_i - \frac{\alpha^r_i u^r}{x_i} \lambda = 0, \quad i = 1, \ldots, n_r
\]

(115a)

\[
\frac{\partial L}{\partial \lambda} = -\prod_{i=1}^{n_r} x_i^{\alpha^r_i} + u^r = 0
\]

(115b)

Taking the ratio of the ith and jth equations we obtain

\[
\frac{p_i}{p_j} = \frac{\alpha^r_i x_j}{\alpha^r_j x_i}
\]

(116)

We can now solve the equation for the jth quantity as a function of the ith quantity and the ith and jth prices. Doing so we obtain

\[
x_j = \frac{\alpha^r_j x_1 p_i}{\alpha^r_i p_j}
\]

(117)

where we treat the first good in the rth group asymmetrically and solve for each demand for a good as a function of the first. Now substituting in equation 115b we obtain

\[
u^r = \prod_{j=1}^{n_r} \left( \frac{\alpha^r_j x_1 p_1}{\alpha^r_i p_j} \right)^{\alpha^r_j}
\]

(118)

Because \(x_1, p_1\) and \(\alpha^r_1\) do not change with \(j\), they can be factored out of the product to obtain

\[
u^r = \left( \frac{x_1 p_1}{\alpha^r_1} \right)^{\sum_{j=1}^{n_r} \alpha^r_j} \prod_{j=1}^{n_r} \left( \frac{\alpha^r_j}{p_j} \right)^{\alpha^r_j}
\]

(119)

We then solve this expression for \(x_1\) as a function of \(u^r\) and the other \(x\)’s. To do so we divide both sides by the product term to obtain

\[
\frac{\sum_{j=1}^{n_r} \alpha^r_j}{x_1} \left( \frac{p_1}{\alpha^r_1} \right) \prod_{j=1}^{n_r} \left( \frac{\alpha^r_j}{p_j} \right)^{\alpha^r_j} = \frac{\alpha^r_j}{\prod_{j=1}^{n_r} \left( \frac{\alpha^r_j}{p_j} \right)^{\alpha^r_j}}
\]

(120)

We now multiply both sides by \(\left( \frac{\alpha^r_j}{p_1} \right)^{\sum_{j=1}^{n_r} \alpha^r_j} \) to obtain

\[
\sum_{j=1}^{n_r} \alpha^r_j = \left( \frac{p_1}{\alpha^r_1} \right)^{\sum_{j=1}^{n_r} \alpha^r_j} u^r
\]

(121)

If we now raise both sides to the power \(\frac{1}{\sum_{j=1}^{n_r} \alpha^r_j}\) we find the value of \(x_1\)
\[ x_1 = \left( \frac{\alpha_r^1}{p_1} \right) \left( \frac{u^r}{\prod_{j=1}^{n_r} \left( \frac{\alpha_j^r}{p_j} \right)^{\alpha_j^r}} \right) \]  

Similarly for the other \( x_k \) so that we have
\[ x_k = \left( \frac{\alpha_r^k}{p_k} \right) \left( \frac{u^r}{\prod_{j=1}^{n_r} \left( \frac{\alpha_j^r}{p_j} \right)^{\alpha_j^r}} \right) \]  

Now if we substitute for the \( i^{th} x \) in equation 113d we obtain
\[ c^r = c^r(p^r, u^r) = c^r(p_{r1}, p_{r2}, \ldots, p_{rn}, u^r) \]
\[ = \sum_{k=1}^{n_r} p_k^r x_k^r \]
\[ = \sum_{k=1}^{n_r} p_k^r \left( \frac{\alpha_k^r}{p_k} \right) \left( \frac{u^r}{\prod_{j=1}^{n_r} \left( \frac{\alpha_j^r}{p_j} \right)^{\alpha_j^r}} \right) \]
\[ = \left( \sum_{k=1}^{n_r} \alpha_k^r \right) \left( \frac{u^r}{\prod_{j=1}^{n_r} \left( \frac{\alpha_j^r}{p_j} \right)^{\alpha_j^r}} \right) \]
\[ = \left( \sum_{k=1}^{n_r} \alpha_k^r \right) \left( u^r \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \]
\[ = \left( \sum_{k=1}^{n_r} \alpha_k^r \right) \left( u^r \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \]
\[ = \left( \sum_{k=1}^{n_r} \alpha_k^r \right) \left( u^r \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \]

Now we assumed that \( \sum_{k=1}^{n_r} \alpha_i^r = 1 \) so we obtain
\[ c^r = u^r \left( \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \]  

11.3. **Finding the indirect utility function for the rth group.** We can invert equation 124 to obtain the indirect utility function for the rth group.
\[ c^r = u^r \left( \prod_{j=1}^{n_r} \left( \frac{p^r_j}{\alpha^r_j} \right)^{\alpha^r_{ij}} \right) \]

\[ u^r = \frac{c^r}{\left( \prod_{j=1}^{n_r} \left( \frac{p^r_j}{\alpha^r_j} \right) \right)^{\alpha^r_{ij}}} \]

(125)

### 11.4. Maximizing utility by allocating expenditure across groups.

The weakly separable overall utility function is given by

\[ v(x) = v(x_1, x_2, \ldots, x_n) \]

\[ = \hat{v} \left( v^1(x^1), v^2(x^2), \ldots, v^J(x^J) \right) \]

\[ = \prod_{r=1}^{J} v^r(x^r)^{\beta_r}. \]  

(126)

Replace \( v^r(x^r) \) in equation 126 with \( u^r \) from equation 125 as follows

\[ v(x) = \prod_{r=1}^{J} u^r(x^r)^{\beta_r}; \]

\[ = \prod_{r=1}^{J} \left( c^r \left( \frac{n_r}{\prod_{j=1}^{n_r} \left( \frac{p^r_j}{\alpha^r_j} \right)^{\alpha^r_{ij}}} \right) \right)^{\beta_r}; \]

\[ = \prod_{r=1}^{J} (c^r)^{\beta_r} \left( \prod_{j=1}^{n_r} \left( \frac{\alpha^r_j}{p^r_j} \right)^{\alpha^r_{ij}} \right)^{\beta_r}; \]

\[ = \prod_{r=1}^{J} (c^r)^{\beta_r} \gamma_r(p^r), \quad \gamma_r(p^r) = \left( \prod_{j=1}^{n_r} \left( \frac{\alpha^r_j}{p^r_j} \right)^{\alpha^r_{ij}} \right) \]

(127)

Now consider maximizing \( v(x) \) in equation 127 by choosing \( c^r \) subject to the constraint that \( \sum_{r=1}^{J} c^r = m \).

The Lagrangian problem is as follows

\[ \mathcal{L} = \prod_{r=1}^{J} (c^r)^{\beta_r} \gamma_r(p^r) - \lambda \left[ \sum_{r=1}^{J} c^r - m \right] \]

(128)

Differentiating equation 128 we obtain
\[
\frac{\partial L}{\partial c^i} = \frac{\beta_i \prod_{r=1}^{J} (c^r)^{\beta_r \gamma_r(p^r)}}{c^i} - \lambda = 0 \quad (129a)
\]

\[
\frac{\partial L}{\partial \lambda} = - \sum_{r=1}^{J} c^r + m = 0 \quad (129b)
\]

If we take the ratio of any of the conditions in equation 129a we obtain

\[
\frac{\beta_i \prod_{r=1}^{J} (c^r)^{\beta_r \gamma_r(p^r)}}{\beta_j \prod_{r=1}^{J} (c^r)^{\beta_r \gamma_r(p^r)}} = 1
\]

\[
\Rightarrow \frac{\beta_i}{\beta_j} \frac{c^i}{c^j} = 1
\]

We can now solve equation 130 for the jth expenditure as a function of the ith expenditure. Doing so we obtain

\[
c^j = \frac{\beta_j}{\beta_i} c^i
\]

\[
= \frac{\beta_j}{\beta_1} c^1
\]

where we treat the first expenditure asymmetrically and solve for each expenditure as a function of the first expenditure. Now substituting in equation 129b we obtain

\[
\frac{\partial L}{\partial \lambda} = - \sum_{r=1}^{J} c^r + m = 0
\]

\[
\Rightarrow \sum_{r=1}^{J} \frac{\beta_r}{\beta_1} c^1 = m
\]

\[
\Rightarrow c^1 \sum_{r=1}^{J} \frac{\beta_r}{\beta_1} = m
\]

\[
\Rightarrow c^1 = \frac{\beta_1 m}{\sum_{r=1}^{J} \beta_r}
\]

Similarly for the other \(c^k\) so that we have

\[
c^k(p, m) = \frac{\beta_k m}{\sum_{r=1}^{J} \beta_r}
\]

To find the demand for the \(k^{th}\) variable in the \(r^{th}\) group rewrite equation 123 imposing the condition that \(\sum_{j=1}^{J} a^j_r = 1.\)
\[ x_k^r = (u^r)^{\sum_{j=1}^{n_r} \alpha_j} \left( \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \left[ \frac{\alpha_k^r}{p_k} \right] \]

(134)

Now replace \( u^r \) in equation 135 with its value in terms of \( c^r \) from equation 125

\[ x_k^r = u^r \left[ \frac{\alpha_k^r}{p_k^r} \right] \left( \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \]

\[ = c^r \left( \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \left[ \frac{\alpha_k^r}{p_k^r} \right] \left( \prod_{j=1}^{n_r} \left( \frac{p_j^r}{\alpha_j^r} \right)^{\alpha_j^r} \right) \]

\[ = c^r \left[ \frac{\alpha_k^r}{p_k^r} \right] \]

(135)

\[ = \frac{\beta_r}{\sum_{j=1}^{\beta_j} \frac{\alpha_k^r}{p_k^r}} \]

\[ = \frac{\beta_r \alpha_k^r}{\sum_{j=1}^{\beta_j} \frac{m}{p_k^r}} \]
12. AGGREGATION ACROSS CONSUMERS

12.1. Introduction to aggregation. Aggregation theory per se is concerned with aggregation across individuals or firms. For example, the economist may be interested in developing a cost function for the steel industry as opposed to the cost function for a single firm. She may also be interested in conditions under which an aggregate function defined over aggregate output and average prices gives the same results as the sum of cost functions defined on individual firm outputs and prices. Or the economist may be interested in aggregating demand functions or welfare across consumers. This issue is important for doing most types of analysis because aggregate data is much more common than individual or firm level data. Aggregation has to do the properties that functions must satisfy for the micro relations to add up properly to the industry wide relations.

12.2. Aggregate demand functions. Suppose that there exist I consumers with preference relations \( \succeq_i \) and ordinary demand functions \( x_i(p, m_i) \) for each of n commodities. The price vector \( p \) contains the prices of the various goods and is assumed to be the same for all consumers, and \( m_i \) represents the income of the \( i^{th} \) consumer. In general, given prices \( p \in \mathbb{R}^n \) and income levels \((m_1, m_2, \ldots, m_I)\) for the I consumers, aggregate demand can be written as

\[
x(p, m_1, m_2, \ldots, m_I) = \sum_{i=1}^{I} x_i(p, m_i)
\]

(136)

Thus, aggregate demand depends not only on prices but also on the specific income levels of the various consumers. A frequently asked question is when are we justified in writing aggregate demand in the simpler form \( x(p, m) \), where \( m = \sum_{i=1}^{I} m_i \). The question is whether aggregate demand depends only on aggregate income, not its distribution between agents.

In order for demand to depend only on aggregate income, aggregate demand must be identical for any two distributions of the same total amount of income across consumers. That is, for any \((m_1, m_2, \ldots, m_I)\) and \((m'_1, m'_2, \ldots, m'_I)\) such that \( \sum_{i=1}^{I} m_i = \sum_{i=1}^{I} m'_i \), we must have \( \sum_{i=1}^{I} x_i(p, m_i) = \sum_{i=1}^{I} x_i(p, m'_i) \).

To understand when this condition is satisfied, consider, starting from some initial distribution \( (m_1, m_2, \ldots, m_I) \), a differential change in income \((dm_1, dm_2, \ldots, dm_I) \in \mathbb{R}^I \), satisfying \( \sum_{i=1}^{I} dm_i = 0 \). Now rewrite equation 136 considering only the \( \ell^{th} \) commodity and imposing the fact that aggregate demand be written as a function of aggregate income.

\[
x^\ell\left(p, m = \sum_{i=1}^{I} m_i\right) = \sum_{i=1}^{I} x_i^\ell(p, m_i)
\]

(137)

where the superscript \( \ell \) refers to the \( \ell^{th} \) commodity and the subscript \( i \) refers to the \( i^{th} \) consumer. Now take the total differential of equation 137 to obtain
\[ \sum_{i=1}^{l} \frac{\partial x^\ell(p, \sum_{i=1}^{l} m_i)}{\partial \sum_{i=1}^{l} m_i} \sum_{i=1}^{l} m_i \, dm_i = \sum_{i=1}^{l} \frac{\partial x^\ell_i(p, m_i)}{\partial m_i} \, dm_i \]

\[ \Rightarrow \sum_{i=1}^{l} \frac{\partial x^\ell(p, \sum_{i=1}^{l} m_i)}{\partial \sum_{i=1}^{l} m_i} \sum_{i=1}^{l} m_i \, dm_i = \sum_{i=1}^{l} \frac{\partial x^\ell_i(p, m_i)}{\partial m_i} \, dm_i \]

\[ \Rightarrow \frac{\partial x^\ell(p, m)}{\partial m} \sum_{i=1}^{l} m_i = \sum_{i=1}^{l} \frac{\partial x^\ell_i(p, m_i)}{\partial m_i} \, dm_i \]

\[ \Rightarrow 0 = \sum_{i=1}^{l} \frac{\partial x^\ell_i(p, m_i)}{\partial m_i} \, dm_i \]

Equation 138 must hold for every \( \ell \). This can be true for all redistributions \((dm_1, dm_2, \ldots, dm_l)\) satisfying \( \sum_{i=1}^{l} dm_i = 0 \) and from any initial income distribution \((m_1, m_2, \ldots, m_l)\) if and only if the coefficients of the different \( dm_i \) are equal, that is

\[ \frac{\partial x^\ell_i(p, m_i)}{\partial m_i} = \frac{\partial x^\ell_j(p, m_j)}{\partial m_j} \] (139)

for every \( \ell \), any two individuals \( i \) and \( j \), and all \((m_1, m_2, \ldots, m_l)\). In short, for any fixed price vector \( p \), and any commodity \( \ell \), the income effect at \( p \) must be the same for whatever consumer at which we look and whatever his level of income or wealth. If this is the case, the individual demand changes arising from any income redistribution across consumers will cancel out. Geometrically, the condition is equivalent to the statement that all consumers' income expansion paths are parallel, straight lines. One special case in which this property holds is when all consumers have identical homothetic preferences. Another is when all consumers have preferences that are quasi-linear \((\psi(p, m) = \nu(p) + m)\) with respect to the same good. The most general result is given in theorem 5.

**Theorem 5.** A necessary and sufficient condition for the set of consumers to exhibit parallel, straight line income expansion paths at any price vector \( p \) is that preferences admit indirect utility functions of the Gorman polar form with the coefficients on \( m \) the same for every consumer \( i \). That is:

\[ \psi(p, m_i) = \frac{m_i - f_i(p)}{g(p)} \]

\[ = \frac{1}{g(p)} m_i - \frac{f_i(p)}{g(p)} \] (140)

Applying Roy’s identity to equation 140 we obtain

\[ \frac{\partial \psi(p, m_i)}{\partial p^k} = \frac{g(p) \left( - \frac{\partial f_i(p)}{\partial p^k} \right) + (f_i(p) - m_i) \frac{\partial \psi(p)}{\partial p^k}}{g(p)^2} \]

\[ = \frac{f_i(p) \frac{\partial \psi(p)}{\partial p^k} - g(p) \frac{\partial f_i(p)}{\partial p^k} - m_i \frac{\partial \psi(p)}{\partial p^k}}{g(p)^2} \] (141)

\[ \frac{\partial \psi(p, m_i)}{\partial m_i} = \frac{1}{g(p)} \]
Then take the ratio of the derivatives in equation 141 and simplify

\[
\frac{\partial \psi(p, m)}{\partial p} = \frac{f_i(p) \frac{\partial g(p)}{\partial p} - g(p) \frac{\partial f_i(p)}{\partial p} - m_i \frac{\partial g(p)}{\partial p}}{g(p)}
\]

\[
= - \frac{\partial f_i(p)}{\partial p^k} + \frac{\partial g(p)}{\partial p^k} f_i(p) - \frac{\partial g(p)}{g(p)} m_i
\]

(142)

\[
\Rightarrow x_i^k(p, m_i) = \frac{\partial f_i(p)}{\partial p^k} + \frac{\partial g(p)}{g(p)} (m_i - f_i(p))
\]

Now consider aggregate demand for the \(k\)th commodity

\[
x^k \left( p, m = \sum_{i=1}^l m_i \right) = \sum_{i=1}^l x_i^k(p, m_i)
\]

\[
= \sum_{i=1}^l \left( \frac{\partial f_i(p)}{\partial p^k} + \frac{\partial g(p)}{g(p)} (m_i - f_i(p)) \right)
\]

\[
= \sum_{i=1}^l \frac{\partial f_i(p)}{\partial p^k} + \frac{\partial g(p)}{g(p)} \sum_{i=1}^l m_i - \frac{\partial g(p)}{g(p)} \sum_{i=1}^l f_i(p)
\]

(143)

\[
= \sum_{i=1}^l \frac{\partial f_i(p)}{\partial p^k} - \frac{\partial g(p)}{g(p)} \sum_{i=1}^l f_i(p) + \frac{\partial g(p)}{g(p)} \sum_{i=1}^l m_i
\]

\[
= a(p) + b(p) \sum_{i=1}^l m_i
\]

12.3. A Representative Consumer. Consider a representative consumer who faces prices \(p\) and has income \(m = \sum_{i=1}^l m_i\) and has indirect utility function

\[
\psi(p, m) = \frac{m - \sum_{i=1}^l f_i(p)}{g(p)}
\]

(144)

The term \(\sum_{i=1}^l f_i(p)\) just depends on prices and so could be written as \(\theta(p)\) which would give

\[
\psi(p, m) = \frac{1}{g(p)} m - \frac{\theta(p)}{g(p)}
\]

(145)

Applying Roy’s identity to equation 144 we obtain
\[
\frac{\partial \psi(p, m)}{\partial p^k} = -g(p) \left( \sum_{i=1}^I \frac{\partial f_i(p)}{\partial p^k} \right) + \left( \sum_{i=1}^I f_i(p) - m \right) \frac{\partial g(p)}{\partial p^k} \\
= \frac{\sum_{i=1}^I f_i(p) \frac{\partial g(p)}{\partial p^k} - g(p) \sum_{i=1}^I \frac{\partial f_i(p)}{\partial p^k} - m \frac{\partial g(p)}{\partial p^k}}{g(p)^2} 
\]

(146)

\[
\frac{\partial \psi(p, m)}{\partial m} = \frac{1}{g(p)} 
\]

Then take the ratio of the derivatives in equation 146 and simplify

\[
\frac{\partial \psi(p, m)}{\partial m} = \frac{\sum_{i=1}^I f_i(p) \frac{\partial g(p)}{\partial p^k} - g(p) \sum_{i=1}^I \frac{\partial f_i(p)}{\partial p^k} - m \frac{\partial g(p)}{\partial p^k}}{g(p)} \\
= -\sum_{i=1}^I \frac{\partial f_i(p)}{\partial p^k} + \frac{\partial g(p)}{g(p)} \left( \sum_{i=1}^I f_i(p) - m \right) \\
\Rightarrow x^k(p, m) = \sum_{i=1}^I \frac{\partial f_i(p)}{\partial p^k} + \frac{\partial g(p)}{g(p)} \left( m - \sum_{i=1}^I f_i(p) \right) \\
= \sum_{i=1}^I \frac{\partial f_i(p)}{\partial p^k} - \frac{\partial g(p)}{g(p)} \sum_{i=1}^I f_i(p) + \frac{\partial g(p)}{g(p)} m 
\]

Equation 147 is the same as 142. Thus, aggregate demand can be written as a function of aggregate wealth if and only if all consumers have preferences that admit indirect utility functions of the Gorman form with equal income coefficients. This is a fairly restrictive set of preferences.

12.4. Other forms of aggregation. Are there less restrictive conditions can be obtained for aggregation if we consider aggregate demand functions that depend on a wider set of aggregate variables than just the total or, equivalently, mean income level. For example, aggregate demand might be allowed to depend on both the mean and the variance of the statistical distribution of wealth or even on the whole statistical distribution itself. Even conditioning aggregate demand on whole statistical distribution of income is still restrictive. It implies that aggregate demand depends only on how many rich and poor there are, not on who in particular is rich or poor.

12.5. Aggregation when income distribution depends on prices. In some situations, individual income levels may be generated by some underlying process that restricts the set of individual income levels which can arise. If so, it may be possible to write aggregate demand as a function of prices and aggregate income or wealth. For example, in general equilibrium models, individual income is generated by individuals’ shareholdings of firms and by their ownership of given, fixed stocks of commodities. Thus, the individual levels of income are determined as a function of the price vector. Alternatively, individual income levels may be determined in part by various government programs that redistribute wealth across consumers. Again, these programs may limit the set of possible wealth distributions that may arise. We call a family of functions \(w_i(p, w), w_i(p, w), \ldots, w_I(p, w)\) with \(\sum_{i=1}^I w_i(p, w) = w\) for all \((p, w)\) a wealth distribution rule. An income distribution rule is similar. When individual income levels are generated by an income distribution rule, we can always write aggregate demand as a function \(x(p, m) = \sum_{i=1}^I x_i(p, m_i(p, m))\).
CONSUMER CHOICE

13. ROY’S IDENTITY AND DEMAND SYSTEMS

There are three basic techniques for deriving a consistent (theoretically plausible) demand system for use in empirical work.

13.1. Direct derivation. Specify a functional form that is consistent with the axioms of preference and then solve the utility maximization problem given that utility function. This will give a series of functions $x_i(p, m)$ that are consistent with the axioms of preference and the utility maximization problem. The conditions usually imposed on the utility function are that it is continuous, non-decreasing in each of its arguments, and that it is quasi-concave so that indifference surfaces have a negative slope.

1. If there is a closed form solution to (a), then this functional form can be used to estimate the demand system.

2. If there is no closed from solution to (a), then an econometric technique based on estimating a series of first order conditions (GMM for example) must be used.

There are only a small number of restrictive forms for which (1) is possible. If there is no closed from solution, the econometrics of estimating highly non-linear first order conditions often precludes successful implementation of (2) for systems of any size or complexity.

13.2. Imposing constraints on arbitrary system. Specify an arbitrary system of demand equations, one for each good, and then directly impose the conditions in section C on the system through statistical restrictions. These conditions are

1. Adding up or Walras law
2. Homogeneity
4. Negativity

This means that nxn matrix formed by the elements $\frac{\partial h_i(u, p)}{\partial p_j}$ is negative semi-definite.

There are a number of difficulties with this approach. The restrictions that must be imposed on the system often involve multiple equations and non-linear functions. A system that was chosen because it is simple and linear may soon become very complex. Secondly, the imposition of the restrictions on general system may reduce its ability to represent general preferences. For example, a general quadratic system as in (79) becomes much less general when the restrictions are applied in (80). Thirdly, there may be no feasible way to impose all the restrictions of consumer theory on a given system and still have it represent general preferences.

13.3. Duality approach. Specify an indirect utility function that satisfies the appropriate properties and then obtain a functional form for the demand equations using Roy’s identity. This avoids the problems of not finding a close form solution in (a) and the problems with specifying a reasonable system that is still reasonable after imposing the restrictions of consumer theory in (b). This method is particularly valuable if the form specified for the indirect utility function is flexible in the sense that it can approximate well arbitrary preferences.
References


