ECON0106: **Microeconomics**

# 3. Optimal Choice and Consumer Theory*


Duarte Gonçalves†

University College London


## 1. Overview

A classic problem studied in economic theory is that of consumer demand. One standard approach is that consumers as choosing bundles of goods, $x \in \mathbb{R}^k_+$, and are faced with a budget constraint $B(p,w)$ determined by their income $w \geq 0$ and the vector of prices they face, $p \in \mathbb{R}^k_{++}$. Specifically, $B(p,w) := \{x \in \mathbb{R}^k_+ \mid p \cdot x \leq w\}$. We assume that they have preferences over the goods, $\succsim$, and we study the properties of their demand $x(p,w) := \arg\max_{\succsim} B(p,w)$. This is perhaps the model of economics that is most used outside academia.

## 2. Utility Maximisation Problem

As we have seen in previous lectures, when we impose some consistency properties on choices (here, demand), we can actually represent preferences $\succsim$ by way of a utility function $u$, and characterise preference-maximising choices as utility-maximising ones. This motivates the term utility maximisation problem to denote the consumer's problem:

$$x(p,w) := \arg\max_{x \in B(p,w)} u(x) \quad \text{(UMP)}$$

$$v(p,w) := \sup_{x \in B(p,w)} u(x)$$

This section shows how we can use structural properties of preference relations to derive properties on consumer demand.

### 2.1. General Implications

We recall that regardless of which utility representation we choose, insofar as it represents the same preference relation, we have the same set of maximisers.

**Proposition 1.** *Let $\succsim$ be a preference relation on $\mathbb{R}^k_+$ and let $u$ and $v$ be two utility representations of $\succsim$. Then, $x(p,w) = \arg\max_{\succsim} B(p,w) = \arg\max_{x \in B(p,w)} u(x) = \arg\max_{x \in B(p,w)} v(x)$.*

*Last updated: 23 September 2025.


† Department of Economics, University College London; duarte.goncalves@ucl.ac.uk. Please do not share these notes with people outside of this class.

*Proof.* Follows by definition: for any utility representation $u$ of $\succsim$, $x \in \arg\max_{\succsim} B(p,w) \iff x \succsim y \, \forall y \in B(p,w) \iff u(x) \geq u(y) \, \forall y \in B(p,w) \iff x \in \arg\max_{x \in B(p,w)} u(x)$. □

In the sequel, to avoid repetition, $\succsim$ will denote a preference relation on $\mathbb{R}^k_+$, $u : \mathbb{R}^k_+ \to \mathbb{R}$ a utility representation of $\succsim$, $v : (p,w) \mapsto v(p,w) \in \mathbb{R}$ the indirect utility,[1] and $x : (p,w) \mapsto x(p,w) \subseteq B(p,w)$ the set of maximisers.

As one would expect, indirect utility is increasing in income and decreasing in prices. The next proposition shows that it is also quasiconvex in prices and income:

**Proposition 2.** *$v(p,w)$ is quasiconvex[2] in $(p,w)$, weakly decreasing in $p$, and weakly increasing in $w$.*

*Proof.* To show quasiconvexity, take any $(p,w),(p',w') \in \{(p,w) \mid v(p,w) \leq \overline{v}\}$. We want to show that $v(\lambda(p,w)+(1-\lambda)(p',w')) \leq \max\{v(p,w), v(p',w')\}$, for any $\lambda \in [0,1]$. $\forall x'' \in X$ such that $(\lambda p + (1-\lambda)p') \cdot x'' \leq \lambda w + (1-\lambda)w'$, we have that (i) $x'' \in B(p,w)$ or (ii) $x'' \in B(p',w')$. (Supposing otherwise means that $p \cdot x'' - w > 0$ and $p' \cdot x'' - w' > 0$ and, doing a convex combination of these, we get a contradiction.) The result follows.

As for the monotonicity properties, note that $p \geq p', w \leq w' \implies B(p,w) \subseteq B(p',w') \implies v(p,w) \leq v(p',w')$ (where is this last implication coming from?). □

If you scale up prices and income, then the consumer is able to afford exactly the same bundles. This implies that both the indirect utility and the set of maximisers remain the same.

**Proposition 3.** *$v(p,w)$ and $x(p,w)$ are homogeneous of degree zero in $(p,w)$: $\forall \lambda > 0$, $v(\lambda p, \lambda w) = v(p,w)$ and $x(\lambda p, \lambda w) = x(p,w)$.*

*Proof.* As $B(\lambda p, \lambda w) = B(p,w)$, then $\arg\max_{\succsim} B(p,w) = \arg\max_{\succsim} B(\lambda p, \lambda w)$. □

## 2.2. Implications of Continuity

We will apply a result we derived earlier to show that consumer demand is nonempty when preferences are continuous.

**Proposition 4.** *If $\succsim$ is continuous, then $x(p,w)$ is nonempty.*

**Correspondences.** We will need to take a small detour to introduce correspondences in order to make use of a very powerful result in optimisation: Berge's Maximum Theorem.

[1] This is a slight abuse of terminology given we defined $v$ as the supremum instead of the maximum, as the latter need not be well defined, i.e., $x(p,w)$ can be empty. Given $B(p,w)$ is compact, these will be the same if $\succsim$ is continuous, as shown later on.

[2] A function $f : X \to \mathbb{R}$ is quasiconvex if $-f$ is quasiconcave. This is equivalent to having that $\{x \in X \mid f(x) \leq \alpha\}$ is convex for any $\alpha \in \mathbb{R}$.

**Definition 1.** A **correspondence** $F$ from $X$ to $Y$ is a mapping that associates with each element $x \in X$ a subset $A \subseteq Y$. This is typically denoted by $F : X \rightrightarrows Y$ or $F : X \to 2^Y$, with $F(x) \subseteq Y$. For $A \subseteq X$, we define the image of $F$ as $F(A) := \cup_{x \in A} F(x)$.

We will introduce two notions of continuity of correspondences in metric spaces:

**Definition 2.** Let $(X, d_X)$ and $(Y, d_Y)$ be metric spaces and $F : X \rightrightarrows Y$.

(i) $F$ is **upper hemicontinuous (uhc) at** $x_0 \in X$ if for any open set $U \subseteq Y$, such that $F(x_0) \subseteq U$, $\exists \epsilon > 0$ such that $F(B_\epsilon(x_0)) \subseteq U$;

(ii) $F$ is **upper hemicontinuous (uhc)** if it is upper hemicontinuous at any $x_0 \in X$;

(iii) $F$ is **lower hemicontinuous (lhc) at** $x_0 \in X$ if for any open set $U \subseteq Y$, such that $F(x_0) \cap U \neq \emptyset$, $\exists \epsilon > 0$ such that $F(x) \cap U \neq \emptyset$, for any $x \in B_\epsilon(x_0)$;

(iv) $F$ is **lower hemicontinuous (lhc)** if it is lower hemicontinuous at any $x_0 \in X$;

(v) $F$ is **continuous at** $x_0 \in X$ if it is both uhc and lhc at $x_0$;

(vi) $F$ is **continuous** if it is both uhc and lhc.

The next proposition provides sequential characterisations of correspondences that may be easier to interpret:

**Proposition 5.** *Let $(X, d_X)$ and $(Y, d_Y)$ be metric spaces and $F : X \rightrightarrows Y$.*

*(i) $F$ is lhc at $x_0$ if and only if for any sequence $\{x_n\}_n \subseteq X$ converging to $x_0$ and any $y_0 \in F(x_0)$, there is an $N$ and a sequence $\{y_n\}_{n>N}$ with $y_n \in F(x_n)$, such that $y_n \to y_0$.*

*(ii) $F$ is uhc (and* compact-valued*) at $x_0$ if (and only if) for any sequence $\{x_n\}_n \subseteq X$ converging to $x_0$ and any sequence $\{y_n\}_n$ such that $y_n \in F(x_n)$, there is some subsequence $\{y_m\}_m \subseteq \{y_n\}_n$ such that $y_m$ converges to some $y_0 \in F(x_0)$.*

Put loosely, part (i) of Proposition 5 shows that lhc is equivalent to stating that every point $y_0 \in F(x_0)$ can be reached by some sequence $y_n \in F(x_n)$. And part (ii) that uhc and compact-valuedness are equivalent to having that the limit $y_0$ of converging sequences $y_n \in F(x_n)$ is also a point in the limit set $F(x_0)$.

These concepts are difficult to diggest; it is very strongly recommended that you develop your understanding with the following:

**Exercise 1.** *(i) Read the lecture notes on correspondences.*

*(ii) Watch a brief (10min) lecture by Rajiv Sethi on upper and lower hemicontinuity: https://youtu.be/OJfzJhsC3Rc.*

One of the main results that you will reencounter later on to prove other fundamental results is Berge's maximum theorem:

**Theorem 1.** ***(Berge's Maximum Theorem)*** *Let $X$ and $\Theta$ be metric spaces, $f : X \times \Theta \to \mathbb{R}$ be a continuous function, and $B : \Theta \rightrightarrows X$ be a non-empty and compact-valued correspondence. Let $f^*(\theta) := \sup_{x \in B(\theta)} f(x, \theta)$ and $X^*(\theta) := \arg\sup_{x \in B(\theta)} f(x, \theta)$. If $B$ is continuous at $\theta \in \Theta$, then $f^*$ is continuous at $\theta$ and $X^*$ is uhc, nonempty, and compact-valued at $\theta$.*

This is very useful theorem that we can then apply off-the-shelf in many circumstances. One of such applications is the following:

**Proposition 6.** *If $\succsim$ is continuous, then $x(p, w)$ is upper hemicontinuous, nonempty- and compact-valued in $(p, w)$.*
*Further, if $u$ is a continuous utility representation of $\succsim$, $v(p, w)$ is continuous.*

**Exercise 2.** *Prove Proposition 6 by showing the following steps:*

*(i) $B$ is nonempty-valued;*

*(ii) $B$ is closed-valued and bounded for any $(p, w) \in \mathbb{R}^k_{++} \times \mathbb{R}_+$. Appeal to Heine–Borel theorem to show it is compact-valued;*

*(iii) $B$ is uhc at any $(p_0, w_0)$* (The sequential characterisation from Proposition 5 is probably easiest)*;*

*(iv) $B$ is lhc at any $(p_0, w_0)$* (Take any open $U \subseteq \mathbb{R}^k_+ : B(p_0, w_0) \cap U \neq \emptyset$ and construct an $\epsilon > 0$ such that $B(p, w) \cap U \neq \emptyset$, $\forall (p, w) \in B_\epsilon((p_0, w_0))$)*;*

*(v) Argue that there is a continuous utility representation of $\succsim$;*

*(vi) Show that you can apply Berge's maximum theorem.*

## 2.3. Implications of Convexity

The next properties are obtained by applying results we have seen in the previous set of lecture notes:

**Proposition 7.** *If $\succsim$ is convex, then $x(p, w)$ is convex. If $\succsim$ is strictly convex, then $x(p, w)$ contains at most one element.*

And, now, we combine both Berge's maximum theorem and the preceeding result to obtain:

**Corollary 1.** *If $\succsim$ is continuous and strictly convex, then $x(p, w)$ is continuous in $(p, w)$.*

**Exercise 3.** *Let $(X, d_X)$ and $(Y, d_Y)$ be metric spaces and $F : X \rightrightarrows Y$. Prove that if $F$ is singleton-valued and uhc, then $F$ is continuous.*

## 2.4. Implications of Local Non-Satiation

Local nonsatiation, which we defined in the previous lecture, is exactly the condition that we need to show that the consumer always exhausts their budget. This is known as the Walras's law:

**Proposition 8.** ***(Walras's Law)*** *If $\succsim$ is locally non-satiated, then for any $x \in x(p,w)$, and any $(p,w) \in \mathbb{R}^k_{++} \times \mathbb{R}_+$, $p \cdot x = w$.*

*Proof.* Let $x \in x(p,w)$ and suppose that $p \cdot x < w$. Then, $\exists \epsilon > 0$ such that $\forall x' \in B_\epsilon(x)$, $p \cdot x' < w$. By local nonsatiation, $\exists x'' \in B_\epsilon(x)$ such that $x'' \succ x$. As $x'' \in B(p,w)$, then $x \notin \arg\max_{\succsim} B(p,w)$. □

**Proposition 9.** *If $u$ is continuous and locally nonsatiated, then $v(p,w)$ is strictly increasing in $w$.*

*Proof.* $w < w' \implies B(p,w) \subsetneq B(p,w')$. Take any $x \in x(p,w)$ and $x' \in x(p,w')$, which exist, by continuity. As $x \in x(p,w) \subseteq B(p,w)$, then $p \cdot x \leq w < w'$, and therefore it violates Walras's Law. Hence, $x \notin \arg\max_{\succsim} B(p,w') \ni x' \implies x' \succ x \iff v(p,w') = u(x') > u(x) = v(p,w)$. □

## 2.5. Implications of Homotheticity

As mentioned in the last lecture, homothetic preferences ensure that a representative consumer exists. That is, if all consumers face the same prices and share the same preferences (but not necessarily the same incomes), then we can treat aggregate demand — the sum of all individual demands — as the choices of an agent that shares the same preferences and whose income is the sum of individual incomes. While this is not true in general, it holds when preferences are homothetic:

**Proposition 10.** *Let every consumer $i \in I$ have income $w_i \geq 0$ and identical preferences $\succsim$. If $\succsim$ is continuous, homothetic and strictly convex, then $\sum_{i \in I} x(p,w_i) = x(p, \sum_{i \in I} w_i)$.*

*Proof.* As $\succsim$ is homothetic, $x \in x(p,1) \iff w \cdot x \in x(p,w)$. As $\succsim$ is strictly convex, $x(p,w)$ is at most a singleton. Continuity of $\succsim$ implies $x(p,w)$ is nonempty. Combining these results, we get that $\sum_{i \in I} x(p,w_i) = \sum_{i \in I} w_i x(p,1) = x(p, \sum_{i \in I} w_i)$. □

# 3. Expenditure Minimisation Problem

The consumer's utility maximisation problem has a "dual problem": given a utility level $u$, the consumer chooses a bundle to minimise the expenditure incurred, subject to the requirement of attaining at least the prespecified utility threshold. More formally, let $\succsim$ be a preference relation on $X := \mathbb{R}^k_+$ and suppose that it admits a utility representation $u$. Define $U := \text{co}(u(X))$, where $\text{co}(A)$ denotes the **convex hull** of set $A$, i.e., the smallest convex set that contains $A$. For any $u \in U$, the consumer's expenditure minimisation problem is given by

$$h(p,u) := \operatorname*{argmin}_{x \in X \,|\, u(x) \geq u} p \cdot x \tag{EMP}$$

$$e(p,u) := \inf_{x \in X \,|\, u(x) \geq u} p \cdot x$$

The set of minimisers $h(p,u)$ is called the **Hicksian demand**.

## 3.1. General Implications

A few properties follow without needing any further assumption. We start with a simple observation that mimicks Proposition 3:

**Proposition 11.** *$h$ is homogeneous of degree zero in $p$ and $e$ is homogeneous of degree one in $p$.*

*Proof.* It follows by definition that $\forall \lambda > 0$, $h(\lambda p, u) = h(p,u)$ and $e(\lambda p, u) = \lambda e(p,u)$. □

While $v$ was shown to be quasiconvex in $(p,w)$, we find that $e$ is concave in $p$, which will allow us to derive further properties:

**Proposition 12.** *$e$ is concave in $p$.*

*Proof.* This follows from the fact that if $f_i : X \to \mathbb{R}$ is concave for every $i \in I$, then $\inf_{i \in I} f_i$ is also concave in $X$.[3] But let us prove this directly in our case: Take any $p, p' \in \mathbb{R}^k_{++}$, $u \in U$, and $\lambda \in [0,1]$. Let $A := \{x \in X \mid u(x) \geq u\}$. For any $x \in A$, by definition, $p \cdot x \geq \inf_{x \in A} p \cdot x =: e(p,u)$ and, similarly, $p' \cdot x \geq e(p', u)$. Hence, for any $x \in A$, $(\lambda p + (1-\lambda)p') \cdot x \geq \lambda e(p,u) + (1-\lambda) e(p', u)$. Then, $e(\lambda p + (1-\lambda)p', u) := \inf_{x \in A} (\lambda p + (1-\lambda)p') \cdot x \geq \lambda e(p,u) + (1-\lambda) e(p', u)$. □

Now we leverage concavity of $e$ in $p$. For that, we need to introduce the concept of a supergradient.

**Definition 3.** A **supergradient** of $f : X \to \mathbb{R}$ at $x_0 \in X$ is an element $c \in \mathbb{R}^k$ such that $f(y) \leq f(x_0) + c \cdot (y - x_0)$, for all $y \in X$. We denote the set of supergradients of $f$ at $x_0$ by $\partial f(x_0)$ and we call $\partial f(x_0)$ the superdifferential of $f$ at $x_0$.

**Theorem 2.** *Let $X \subseteq \mathbb{R}^k$ be a convex set and $f$ be a real-valued function on $X$. $f$ is concave on $int(X)$ if and only if $\forall x \in int(X)$, $\partial f(x) \neq \emptyset$.*

The intuition is as follows: pick $x, y, z \in X$. For $c \in \partial f(x)$, $f(y) \leq f(x) + c \cdot (y - x)$ and $f(z) \leq f(x) + c \cdot (z - x)$. By a convex combination of the two, with $\lambda \in (0,1)$, $\lambda f(y) + (1-\lambda) f(z) \leq f(x) + c(\lambda y + (1-\lambda)z - x)$. Choosing $x = \lambda y + (1-\lambda)z$ delivers concavity of $f$.

A supergradient — also called superderivative — is generalising the notion of derivative to

[3] Equivalently, the supremum over a family of convex functions $f_i : X \to \mathbb{R}$ is convex in $X$.

functions that are not necessarily differentiable everywhere. For instance, take the function $f:\mathbb{R}\to\mathbb{R}$ such that $f(x)=x$ if $x\leq 0$ and $f(x)=-x$ if $x>0$. This is a concave function and its derivative exists everywhere — $f'(x)=1$ if $x<0$ and $f'(x)=-1$ for $x>0$ — but at zero, where it has a kink. Its supergradient, however, is well-defined everywhere: $\partial f(x)=\{1\}$ for $x<0$, $\partial f(x)=\{-1\}$ for $x>0$ and $\partial f(0)=[-1,1]$.

Given a convex $X\subseteq\mathbb{R}^k$ and a concave function $f:X\to\mathbb{R}$, we can say a lot about it:[4]

(i) For any $x\in \text{relint}X$,[5] $\partial f(x)$ is nonempty, convex, and compact.

(ii) For any $c\in\partial f(x)$ and $c'\in\partial f(x')$, $(c'-c)\cdot(x'-x)\leq 0$.

(iii) If $f$ is continuous at $x$, then the superdifferential $\partial f(x)$ is a singleton if and only if $f$ is differentiable at $x$. In this case, $f'(x)=c\in\partial f(x)=\{c\}$.

(iv) $f''$ exists almost everywhere in $\text{int}(X)$.[6]

(v) If $k=1$, at any $x\in\text{int}(X)$, $\partial f(x)=[f'_+(x),f'_-(x)]$, where $f'_-,f'_+$ denote the left- and right-derivatives of $f$.

Back to consumer demand. The next theorem is called the compensated law of demand and it says that the Hicksian demand is weakly decreasing in prices. We shall prove this result by showing that Hicksian demand is a supergradient of expenditure and then using the properties of supergradients.

**Lemma 1.** *If $x_0\in h(p_0,u)$, then $x_0$ is a supergradient of $e(\cdot,u)$ at $p_0$.*

*Proof.* As $p_0\cdot x_0=e(p_0,u)$ and $p\cdot x_0\geq e(p,u)$ for any $p\in\mathbb{R}^k_{++}$, we have that $e(p,u)\leq e(p_0,u)+x_0\cdot(p-p_0)$. □

**Theorem 3.** ***(Compensated Law of Demand)*** *If $p'\geq p$, $x\in h(p,u)$, and $x'\in h(p',u)$, then $(p'-p)\cdot(x'-x)\leq 0$.*

*Proof.* This is obtained immediately by combining property (ii) of concave functions as listed above and Lemma 1. □

Note that, if $p'$ equals $p$ in every dimension except one, say dimension $i$ for which $p'_i>p_i$, then the theorem is telling us that $x'_i\leq x_i$.

Finally, the last result that we can show without adding any assumption on preferences is a counterpart of the monotonicity properties in Proposition 2:

[4]We can also have counterparts of all of these results for convex functions, as if $f$ is concave, $-f$ is convex. Two references for the future: Boyd and Vandenberghe (2004) and the less well-known Niculescu and Persson (2018).

[5]The relative interior of a convex set $A$, relint$(A)$, is defined as $\text{relint}(A):=\{x\in A\mid\forall y\in A\setminus\{x\},\exists z\in A,\lambda\in(0,1)\text{ s.t. }x=\lambda y+(1-\lambda)z\}$.

[6]This is called **Alexandrov theorem**.

**Proposition 13.** *$e$ is weakly increasing in $p$ and $u$.*

*Proof.* Take any $u' \geq u$ and $p' \geq p$. For any $p'' \in \mathbb{R}^k_{++}$, we have that (by transitivity) $\{x \in X \mid u(x) \geq u\} \supseteq \{x \in X \mid u(x) \geq u'\} \implies e(p'',u) \leq e(p'',u')$. And, for any $u'' \in U$, $p \cdot x \leq p' \cdot x$ $\forall x : u(x) \geq u''$, which implies $e(p,u'') \leq e(p',u'')$. □

## 3.2. Implications of Continuity

We were sneaky in defining $e(p,u)$ as an infimum rather than a minimum, as the infimum will always be well defined (why?), but the minimum may not as $h(p,u)$ may be empty. As you may have anticipated, continuity of the utility function will solve this problem and provide some additional properties on Hicksian demand (courtesy of Berge's maximum theorem).

**Proposition 14.** *If $u$ is a continuous utility representation of $\succsim$, then $e(p,u)$ is continuous and $h(p,u)$ is nonempty, compact-valued, and uhc in $(p,u)$.*

*Proof.* Take an arbitrary point $x_0 \in X$ such that $u(x_0) \geq u$. As $u \in U$, $x_0$ must exist. Let $A := \{x \in \mathbb{R}^k_+ \mid p \cdot x \leq p \cdot x_0\}$

**Claim**: $A$ is compact.

Note that $A$ is closed. Let $\overline{x} \in X$ be such that all its coordinates are equal to the largest coordinate of $x_0$, denoted by $\overline{x}_0$. As $[0,\overline{x}_0]^k$ is a compact subset of $\mathbb{R}^k_+$ (why?) and $A \subseteq [0,\overline{x}_0]^k$ — as $p \cdot x \leq p \cdot x_0 \leq p \cdot \overline{x}_0$ — we have that $A$ is compact.

**Claim**: $B := A \cap \{x \in X \mid u(x) \geq u\}$ is compact.

To see this, note that by continuity of $u$,[7] $\{x \in X \mid u(x) \geq u\}$ is closed. Hence, $B$ is a closed subset of $[0,\overline{x}_0]^k$ and therefore compact.

**Claim**: $\min_{x \in B} p \cdot x = \inf_{x \in B} p \cdot x = \inf_{x \in X \mid u(x) \geq u} p \cdot x$.

The first equality is due to Weierstrass extremum theorem; the second equality is due to the fact that $\forall x \in B$, $y \in \{x \in X \mid u(x) \geq u\} \setminus B$, $y$ induces a higher expenditure than $x$ $p \cdot y > p \cdot x_0 \geq p \cdot x$, and both attain utility weakly higher than $u$.

**Last step:** The remainder of the proof follows by constructing a continuous, nonempty- and compact-valued correspondence that does not entail greater expenditure and then applying Berge's maximum theorem. □

**Exercise 4.** *Complete the proof of Proposition 14.*

In fact, with continuity we get that the lower bound on the utility is actually attained:

**Lemma 2.** *If $u$ is a continuous utility representation of $\succsim$, then $\forall x \in h(p,u)$, $u(x) = u$.*

*Proof.* Suppose instead that $u(x) > u$. Then, continuity implies that $\exists \lambda \in [0,1)$ such that $u(\lambda x) > u$, and as $p \cdot x > p \cdot \lambda x$, $x \notin h(p,u)$, a contradiction. □

[7] This is why we need to assume that $u$ is a continuous utility representation and not just that $\succsim$ is continuous.

This gives sense to the expression "compensated law of demand": by varying the prices $p$, $h(p,u)$ describes how the consumer substitutes across the different goods while attaining the same utility level. The "compensated" term comes from imagining that the consumer is given additional income to compensate the price changes. The next section — dealing with local non-satiation — makes this clearer.

### 3.3. Implications of Local Non-Satiation

**Theorem 4.** *Let $\succsim$ be locally nonsatiated and $u$ be a continuous utility representation of $\succsim$. Then*

*(i) $h(p,v(p,w)) = x(p,w)$ and $e(p,v(p,w)) = w$;*

*(ii) $h(p,u) = x(p,e(p,u))$ and $u = v(p,e(p,u))$.*

**Exercise 5.** *Prove Theorem 4.*

This equivalence between Marshallian ($x(p,w)$) and Hicksian demand ($h(p,u)$) casts light onto why the compensated demand $h(p,u)$ is called compensated. If by increasing prices $v(p,w)$ decreases, in order to keep $v(p,w) = u$ we need to compensate the consumer by increasing income $w$.

### 3.4. Implications of Convexity

Finally, to conclude the overview of the properties of the expenditure minimisation program, we note some implications of convexity of preferences.

**Proposition 15.** *(i) If $\succsim$ is convex, then $h(p,u)$ is convex.*

*(ii) If $\succsim$ is strictly convex and $u$ is a continuous utility representation, then $h(p,u)$ is a singleton, continuous in $(p,u)$, and $h(p,u) = e'_p(p,u)$.*

*Proof.* For (i) take any $x, x' \in h(p,u)$ and any $\lambda \in [0,1]$. Note that $p \cdot (\lambda x + (1-\lambda)x') = e(p,u)$ and that $u(\lambda x + (1-\lambda)x') \geq \min\{u(x), u(x')\} \geq u$. Hence, $\lambda x + (1-\lambda)x' \in h(p,u)$.
For (ii), we note that by Theorem 4, we have that $x(p,e(p,u)) = h(p,u)$, and by Proposition 7, $x(p,e(p,u))$ is a singleton. Continuity follows from Proposition 14. The last bit of (ii) follows from the fact that $h(p,u)$ is the unique supergradient of $e(p,u)$. □

## 4. Solving Optimisation Problems using Calculus

It is expected and assumed that you wil be able to handle constrained optimisation problems using Lagrangian methods and Karush-Kuhn-Tucker conditions — although it is unlikely you will need it in this course. If you are unfamiliar with optimisation using calculus, a concise reference is the Mathematical Appendix in Mas-Colell et al. (1995); in this case, the directly relevant appendices are *M.J: Unconstrained Optimisation* (pp. 954–56), *M.K: Constrained Optimisation* (pp. 956–64), and *M.L: The Envelope Theorem* (pp. 964–66).

## 5. Afriat's Theorem (*)

Suppose we observe data in the form $\{(x_t, p_t)\}_{t\in[T]}$. We want to know when it is the case that our data can be rationalised by positing that the consumer is maximising utility. That is, $\forall t \in [T]$, $x_t \in x(p_t, w_t)$ for some income $w_t$. One issue that is quickly resolved is that we don't observe income. Notice that, if we assume that the consumer's preferences are locally nonsatiated, we should have that $p_t \cdot x_t = w_t$.

Let's recall some definitions on revealed preference, adjusted to the case at hand. We say that $x_t$ is **directly revealed preferred to** $x_s$ if $x_t$ was chosen and $x_s$ was affordable, i.e., $p_t \cdot x_s \leq p_t \cdot x_t$. Bundle $x_t$ is **revealed preferred** to $x_s$ if there is a sequence of bundles $\{x_m\}_{m\in[M]}$ such that $x_t$ is directly revealed preferred to $x_1$, $x_1$ to $x_2$, and so on, with $x_M$ being directly revealed preferred to $x_s$.

To adjust the definition of revealed strict preference, we rely on local nonsatiation (how?): we say that $x_t$ is **revealed strictly preferred to** $x_s$ if it was strictly less expensive than $x_t$ under $p_t$, that is, $p_t \cdot x_s < p_t \cdot x_t$. Finally, our data satisfies **Generalised Axiom of Revealed Preference** (GARP) if there is no pair of bundles $x, y$ such that $x$ is revealed preferred to $y$ and $y$ is revealed strictly preferred to $x$.

Our main result for this section is:

**Theorem 5.** ***(Afriat's (1967) Theorem)*** *Let be* $\{(x_t, p_t)\}_{t=1,\dots,T}$ *be a dataset comprising a collection of chosen bundles* $x_t$ *at prices* $p_t$*. The following statements are equivalent:*

*(i) The dataset can be rationalised by a locally nonsatiated preference relation* $\succsim$ *that admits a utility representation.*

*(ii) There is a continuous, concave, piecewise linear, strictly monotone utility function* $u$ *that rationalises the dataset.*

*(iii) The dataset satisfies GARP.*

*(iv) There exist positive* $\{u_t, \lambda_t\}_{t\in[T]}$ *such that* $u_s \leq u_t + \lambda_t p_t \cdot (x_s - x_t)$*, for all* $t, s \in [T]$.

A proof, while certainly not beyond the scope of this course, is surely beyond its time contraints. However, some comments are in order. First, comparing (i) and (ii) we see that if we can rationalise the data with local nonsatiation, we might as well throw in continuity, concavity, piecewise linearity, and strict monotonicity, as these pose no additional constraint on the (finite) data. Second, less surprisingly, GARP (appropriately redefined) is still exactly what we need to rationalise the data as being originated by preference-maximising behaviour.

Third, while GARP is already saving computing time/cost — and certainly data requirements — when compared to HARP, condition (iv) above is far easier to check as it reduces to a simple linear programming problem. Where does this condition come from? Recall the concept of a supergradient: $\forall q_t \in \partial u(x_t)$ and $\forall x_s$, $u(x_s) \leq u(x_t) + q_t \cdot (x_s - x_t)$. If $u$ is concave, then supergradients always exist, and, as $u$ is differentiable almost everywhere (by concavity), $\partial u(x) = \{u'(x)\}$ almost everywhere. So, almost everywhere we get $\forall x_s$, $u(x_s) \leq u(x_t) + u'(x_t) \cdot (x_s - x_t)$. Now, to get intuition as to why we have $\lambda_t p_t$ in the stead of $u'(x_t)$, just consider that $u$ is in fact differentiable. The Langragian for utility maximisation problem $\max_{x \in B(p,w)} u(x)$ is then given by $u(x) + \lambda \cdot (w - p \cdot x)$, with first-order conditions for an interior optimum $u'(x) = \lambda p$.

# 6. Further Reading

**Standard References**: Mas-Colell et al. (1995, Chapters 2, 3), Rubinstein (2018, Chapter 5), Kreps (2012, Chapters 3, 4), Kreps (1988, Chapter 3).

**Related questions/topics**: This is probably the most canonical topic taught in a microeconomics class. Here, we obtained a number of useful properties on consumer choices without taking a single derivative. This is to illustrate how much we can say even when we make a small number of assumptions. A treatment of the topic with differentiability can be found in any of the above references.

Consumer choice is a natural backdrop for applications for choice models. Nowadays we have a large array of very interesting models (going much beyond the standard theory we covered) that are meant to capture more realistic features of how people choose. For instance, some models (of limited or random attention) represent cases in which agents don't consider all the choices available — say as we often do when faced with the huge number of possibilities when doing groceries. Others models (search, information acquisition) have agents searching and acquiring information prior to choosing, which is possibly particularly relevant for expensive purchases. It is interesting to think about how to model agents switching between habitual consideration and searching new alternatives, or about how properties of (aggregate) demand change when the economy is populated with such behavioural agents.

# References


**Afriat, Sidney N.** 1967. "The Construction of Utility Functions from Expenditure Data." *International Economic Review* 8 (1): 67–77. 10.2307/2525382.

**Boyd, Stephen, and Lieven Vandenberghe.** 2004. *Convex Optimization.* Cambridge University Press.

**Kreps, David M.** 1988. *Notes on the Theory of Choice.* Westview Press.

**Kreps, David M.** 2012. *Microeconomic Foundations I. Choice and Competitive Markets.* Princeton, NJ: Princeton University Press.

**Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green.** 1995. *Microeconomic Theory.* Oxford University Press.
**Niculescu, Constantin P., and Lars-Erik Persson.** 2018. *Convex Functions and Their Applications: A Contemporary Approach.* Springer.
**Rubinstein, Ariel.** 2018. *Lectures in Microeconomic Theory.* Princeton University Press.