# Advanced Microeconomics - Econ 481[1]

# Lecture Notes

Martín Alfaro
University of Alberta

2018

[1]These are the first-version notes for the the course 481 at University of Alberta. They are still preliminary and in beta. Please, if you find any typo or mistake, send it to malfaro@ualberta.ca.

# Table of Contents

# Notation

This is a derivation

This is some comment

This is a comment on advanced topics which are not part of the course (you can ignore it without loss of continuity regarding the text)

- I denote vectors by bold lowercase letters (for instance, $\mathbf{x}$) and matrices by bold capital letters (for instance, $\mathbf{X}$).
- To differentiate between the verb "maximize" and the operator "maximum", I denote the former with "max" and the latter with "sup" (i.e., supremum). The same caveat applies to "minimize" and "minimum", where I use "min" and "inf", with the latter indicating infimum.
- "iff" means "if and only if"
- $\exp(x)$ is the function $e^x$.
- Random variables are denoted with a bar below. For instance, $\underline{x}$.

These notes contain hyperlinks in blue and red text. If you are using Adobe Acrobat Reader, you can click on the link and easily navigate back by pressing Alt+Left Arrow.

# Lecture Note 1

# Math Review

The goal of this note is to review some of optimization techniques and the Envelope Theorem. It complements the slides that cover the Math Review. The explanations are kept to the minimum, and the note only aims at providing cookbook procedures to solve exercises.

## 1.1 Unconstrained Optimization

We establish conditions for a maximization and a minimization.

### 1.1.1 Maximization

Let $f : X_1 \times X_2 \times \Lambda \to \mathbb{R}$ with $X_i := [0, \overline{x}_i]$ for $i = 1, 2$ and $\Lambda \subseteq \mathbb{R}_{++}$. The maximization problem is

$$\max_{(x_1,x_2)\in X_1\times X_2} f(x_1, x_2; \alpha).$$

A solution to the problem is a vector of endogenous variables as a function of all the parameters: $(x_1^*(\alpha), x_2^*(\alpha))$. Sometimes the following notation is used to define the solutions:

$$(x_1^*(\alpha), x_2^*(\alpha)) := \arg\max_{(x_1,x_2)\in X_1\times X_2} f(x_1, x_2; \alpha).$$

The value function is the objective function evaluated at the solution. Since the solution is a function of the parameters, the value function is a function of the parameters: $f^*(\alpha)$. Formally, it is defined by:

$$f^*(\alpha) := f[x_1^*(\alpha), x_2^*(\alpha); \alpha] := \sup_{(x_1,x_2)\in X_1\times X_2} f(x_1, x_2; \alpha).$$

Now, let's establish the conditions to pin down the solutions $(x_1^*(\alpha), x_2^*(\alpha))$. We suppose that the Inada conditions hold, and so we rule out boundary solutions.[1] Since any solution must be interior, the FOCs (first-order conditions) are necessary to obtain

[1] The Inada conditions in this case are $\lim\limits_{x_1\to 0} \frac{\partial f(x_1,x_2;\alpha)}{\partial x_1} = \lim\limits_{x_2\to 0} \frac{\partial f(x_1,x_2;\alpha)}{\partial x_2} = \infty$ and $\lim\limits_{x_1\to \overline{x}_1} \frac{\partial f(x_1,x_2;\alpha)}{\partial x_1} = \lim\limits_{x_2\to \overline{x}_2} \frac{\partial f(x_1,x_2;\alpha)}{\partial x_2} < 0$ for each $\alpha \in \Lambda$.

the solution:

$$\nabla f(x_1, x_2; \alpha) := \begin{pmatrix} \frac{\partial f(x_1,x_2;\alpha)}{\partial x_1} \\ \frac{\partial f(x_1,x_2;\alpha)}{\partial x_2} \end{pmatrix} = 0.$$

Let the Hessian Matrix be

$$\mathbf{H}(x_1, x_2; \alpha) := \begin{pmatrix} \frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_1^2} & \frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_2^2} \end{pmatrix}.$$

and when evaluated at the solution:

$$\mathbf{H}[x_1^*(\alpha), x_2^*(\alpha); \alpha] = \mathbf{H}(x_1, x_2; \alpha)\rfloor_{\substack{x_1 = x_1^*(\alpha) \\ x_2 = x_2^*(\alpha)}}.$$

A sufficient SOC (second-order condition) for an interior (local) maximum is $\mathbf{H}(x_1, x_2; \alpha)$ to be negative definite when evaluated at the solution. For two variables this requires that:

- $\left.\frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_i^2}\right\rfloor < 0$ for $i = 1, 2$, and
- $\det \mathbf{H}[x_1^*(\alpha), x_2^*(\alpha); \alpha] > 0$.

When $\mathbf{H}$ is symmetric, the conditions can be simplified. In the example, we know that the cross derivatives are equal, since $f \in \mathcal{C}^2$ (i.e. $f$ is twice continuously differentiable) and we can apply Young's theorem. Therefore, $\frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_1 \partial x_2} = \frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_2 \partial x_1}$, and the sufficient condition for the SOC to hold is simpler and given by:

- $\left.\frac{\partial^2 f(x_1,x_2;\alpha)}{\partial x_i^2}\right\rfloor < 0$ for *either* $i = 1, 2$ (if it holds for one, it holds for the other), and
- $\det \mathbf{H}[x_1^*(\alpha), x_2^*(\alpha); \alpha] > 0$.

For more than two variables, we can generalize the condition for $\mathbf{H} \in \mathbb{R}^{N \times N}$ symmetric that it is negative definite by checking that

all the leading principal minors have the sign $(-1)^k \Delta_k > 0$ for each $k = 1, 2, ..., N$.

For instance, for a function $f \in \mathcal{C}^2$ of three variables $(x_{1,}, x_2, x_3)$ (so that $\mathbf{H}$ is symmetric), the Hessian is :

$$\mathbf{H}(x_1, x_2, x_3) := \begin{pmatrix} \frac{\partial^2 f(\cdot)}{\partial x_1^2} & \frac{\partial^2 f(\cdot)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\cdot)}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(\cdot)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\cdot)}{\partial x_2^2} & \frac{\partial^2 f(\cdot)}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(\cdot)}{\partial x_3 \partial x_1} & \frac{\partial^2 f(\cdot)}{\partial x_3 \partial x_2} & \frac{\partial^2 f(\cdot)}{\partial x_3^2} \end{pmatrix}.$$

Thus, $\mathbf{H}$ is definite negative at the solution if the following expressions evaluated at the solution hold:

- $\frac{\partial^2 f(\cdot)}{\partial x_1^2} < 0$,
- $\det \begin{pmatrix} \frac{\partial^2 f(\cdot)}{\partial x_1^2} & \frac{\partial^2 f(\cdot)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(\cdot)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\cdot)}{\partial x_2^2} \end{pmatrix} > 0$, and
- $\det \mathbf{H}(\cdot) < 0$.

### 1.1.2 Minimization

Assuming that there are no boundary solutions, the FOCs to identify the minimum function $f$ are given by $\nabla f(x_1, x_2; \alpha) = 0$. Furthermore, a sufficient SOC is that $\mathbf{H}$ is positive definite. When $\mathbf{H} \in \mathbb{R}^{N \times N}$ is symmetric, this requires that

all the leading principal minors $\Delta_k$ satisfy $\Delta_k > 0$ for each $k = 1, 2, ...N$.

Thus, for $f \in \mathcal{C}^2$ and two variables (so that $\mathbf{H}$ is symmetric), $\mathbf{H}$ is definite positive at the solution if the following holds when we evaluate the expressions at the solution:

- $\frac{\partial^2 f(x_1, x_2; \alpha)}{\partial x_i^2} > 0$ for *either* $i = 1, 2$, and
- $\det \mathbf{H}[x_1^*(\alpha), x_2^*(\alpha); \alpha] > 0$.

Likewise, we can generalize this result for three variables. Formally, for $f \in C^2$ and three variables (so that $\mathbf{H}$ is symmetric), $\mathbf{H}$ is definite positive at the solution if the following expressions evaluated at the solution hold:

- $\frac{\partial^2 f(\cdot)}{\partial x_1^2} > 0$,
- $\det \begin{pmatrix} \frac{\partial^2 f(\cdot)}{\partial x_1^2} & \frac{\partial^2 f(\cdot)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(\cdot)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\cdot)}{\partial x_2^2} \end{pmatrix} > 0$, and

- $\det \mathbf{H}(\cdot) > 0$.

## 1.2 Constrained Optimization (Lagrange)

A constrained optimization imposes some restrictions on the values that the domain $X_1 \times X_2$ can take. When the objective function and the constraints are differentiable, the optimization problem can be solved by using Lagrangian techniques. This procedure can be summarized by two steps:

[1] Construct the Lagrangian function

[2] Proceed as any unconstrained problem, but taking the Lagrangian as the objective function. This means you have to optimize with respect to both the original variables and the Lagrange multiplier.

We consider the case of a maximization, since the minimization problem is similar. Let the problem be

$$\max_{x_1,x_2} f(x_1, x_2; \alpha)$$
$$\text{subject to } \kappa = g(x_1, x_2; \alpha),$$

where $\alpha, \kappa$ are parameters. Then, the Lagrangian is

$$\mathscr{L}(x_1, x_2, \lambda; \alpha, \kappa) := f(x_1, x_2; \alpha) + \lambda [\kappa - g(x_1, x_2; \alpha)],$$

and we optimize $\mathscr{L}$ as if it were an unconstrained problem (the same FOCs and SOCs).

Notice how we have written the constraint in the Lagrangian: the parameter $\kappa$ minus the function $g$. We could have defined it the other way round[2] and obtained the same optimal solutions $x_1^*(\alpha, \kappa)$ and $x_1^*(\alpha, \kappa)$. However, writing the constraint in the way we did ensures that the Lagrange multiplier $\lambda^*(\alpha, \kappa)$ has a positive sign. This is only important when $\lambda^*$ has some economic interpretation, as it happens in consumer theory. See Section 1.4.2.

[2] That is, $\mathscr{L}(x_1, x_2, \lambda; \alpha, \kappa) := f(x_1, x_2; \alpha) + \lambda [g(x_1, x_2; \alpha) - \kappa]$

## 1.3 Envelope Theorem

Next, we'll review the Envelope Theorem. This provides an easy way to calculate the impact on the value function of a parameter change.

### 1.3.1 Envelope Theorem for Unconstrained Optima

Let $f(x_1, x_2; \alpha)$ be the objective function, where $x_1$ and $x_2$ are decision variables and $\alpha$ is a parameter. Suppose that the optimization problem is:

$$\max_{x_1,x_2} f(x_1, x_2; \alpha).$$

Assuming there are no boundary solutions, we can characterize the solution through the FOCs:

$$\frac{\partial f(x_1, x_2; \alpha)}{\partial x_1} = 0,$$
$$\frac{\partial f(x_1, x_2; \alpha)}{\partial x_2} = 0.$$

Let the solution be $x_1^*(\alpha)$ and $x_2^*(\alpha)$. The value function $f^*$ is the objective function evaluated at its optimal solutions, which is given by

$$f^*(\alpha) := f[x_1^*(\alpha), x_2^*(\alpha); \alpha].$$

The Envelope Theorem provides a way to identify value of $\frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha}$. To ease notation, we write derivatives as function the values at which they are evaluated. For example,

$$\frac{\partial f[x_1^*(\alpha), x_2^*(\alpha); \alpha]}{\partial \alpha} := \left.\frac{\partial f(x_1, x_2; \alpha)}{\partial \alpha}\right]_{\substack{x_1 = x_1^*(\alpha) \\ x_2 = x_2^*(\alpha)}}.$$

Taking the derivative of $f^*$ with respect to $\alpha$:

$$\frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha} = \underbrace{\frac{\partial f[x_1^*(\alpha), x_2^*(\alpha); \alpha]}{\partial x_1}\frac{\mathrm{d}x_1^*(\alpha)}{\mathrm{d}\alpha} + \frac{\partial f[x_1^*(\alpha), x_2^*(\alpha); \alpha]}{\partial x_2}\frac{\mathrm{d}x_2^*(\alpha)}{\mathrm{d}\alpha}}_{\text{indirect effects}} + \underbrace{\frac{\partial f[x_1^*(\alpha), x_2^*(\alpha); \alpha]}{\partial \alpha}}_{\text{direct effects}}.$$

But, by the FOCs,

$$\frac{\mathrm{d}f^{*}(\alpha)}{\mathrm{d}\alpha}=\underbrace{\frac{\partial f\left[x_{1}^{*}(\alpha),x_{2}^{*}(\alpha);\alpha\right]}{\partial x_{1}}}_{=0}\frac{\mathrm{d}x_{1}^{*}(\alpha)}{\mathrm{d}\alpha}+\underbrace{\frac{\partial f\left[x_{1}^{*}(\alpha),x_{2}^{*}(\alpha);\alpha\right]}{\partial x_{2}}}_{=0}\frac{\mathrm{d}x_{2}^{*}(\alpha)}{\mathrm{d}\alpha}+\frac{\partial f\left[x_{1}^{*}(\alpha),x_{2}^{*}(\alpha);\alpha\right]}{\partial\alpha}.$$

which provides the main result of the Envelope Theorem.

**Result 1.1** *Envelope Theorem for Unconstrained Optimization:*

$$\frac{\mathrm{d}f^{*}(\alpha)}{\mathrm{d}\alpha}=\left.\frac{\partial f(x_{1},x_{2};\alpha)}{\partial\alpha}\right]_{\substack{x_{1}=x_{1}^{*}(\alpha)\\x_{2}=x_{2}^{*}(\alpha)}}.$$

The practical implication of this result is that, if you are interested in the impact of $\alpha$ on the optimal objective function, you only have to look at the direct effect of $\alpha$ on $f^{*}$—the impact on $f^{*}$ due to variations in $x_{1}^{*}$ and $x_{2}^{*}$ are irrelevant.

### 1.3.2 Constrained Optima

The Envelope Theorem can also be applied to the case of unconstrained maxima. Suppose now the optimization problem is

$$\max_{x_{1},x_{2}} f(x_{1},x_{2};\alpha) \text{ subject to } \kappa=g(x_{1},x_{2};\alpha).$$

The Lagrangian is $\mathscr{L}(x_{1},x_{2},\lambda;\alpha):=f(x_{1},x_{2};\alpha)+\lambda\left[\kappa-g(x_{1},x_{2};\alpha)\right]$, which delivers the following FOCs:

$$\text{FOC:}\begin{cases}\frac{\partial\mathscr{L}(x_{1},x_{2},\lambda;\alpha)}{\partial x_{1}}=\frac{\partial f(x_{1},x_{2};\alpha)}{\partial x_{1}}-\lambda\frac{\partial g(x_{1},x_{2};\alpha)}{\partial x_{1}}=0,\\ \frac{\partial\mathscr{L}(x_{1},x_{2},\lambda;\alpha)}{\partial x_{2}}=\frac{\partial f(x_{1},x_{2};\alpha)}{\partial x_{2}}-\lambda\frac{\partial g(x_{1},x_{2};\alpha)}{\partial x_{2}}=0,\\ \frac{\partial\mathscr{L}(x_{1},x_{2},\lambda;\alpha)}{\partial\lambda}=\kappa-g(x_{1},x_{2};\alpha)=0.\end{cases}$$

This enables us to obtain the solutions $x_{1}^{*}(\alpha), x_{2}^{*}(\alpha), \lambda^{*}(\alpha)$ and the value function $f^{*}(\alpha):=f\left[x_{1}^{*}(\alpha),x_{2}^{*}(\alpha);\alpha\right]$. The main implication of the Envelope Theorem is stated in the following.

**Result 1.2** *Envelope Theorem for Constrained Optima:*

$$\frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha} = \frac{\mathrm{d}\mathscr{L}^*(\alpha)}{\mathrm{d}\alpha} = \left.\frac{\partial \mathscr{L}(x_1, x_2; \alpha)}{\partial \alpha}\right|_{\substack{x_1 = x_1^*(\alpha) \\ x_2 = x_2^*(\alpha)}}.$$

**Proof.** (OPTIONAL) $\mathscr{L}^*(\alpha) := \mathscr{L}[x_1(\alpha), x_2(\alpha), \lambda(\alpha); \alpha]$

$\frac{\mathrm{d}\mathscr{L}^*(\alpha)}{\mathrm{d}\alpha} = \frac{\mathrm{d}\mathscr{L}[\cdot]}{\mathrm{d}\alpha} = \frac{\partial\mathscr{L}[\cdot]}{\partial x_1}\frac{\mathrm{d}x_1(\alpha)}{\mathrm{d}\alpha} + \frac{\partial\mathscr{L}[\cdot]}{\partial x_2}\frac{\mathrm{d}x_2(\alpha)}{\mathrm{d}\alpha} + \frac{\partial\mathscr{L}[\cdot]}{\partial \lambda}\frac{\mathrm{d}\lambda(\alpha)}{\mathrm{d}\alpha} + \frac{\partial\mathscr{L}[\cdot]}{\partial \alpha}$

$\Rightarrow \frac{\mathrm{d}\mathscr{L}^*(\alpha)}{\mathrm{d}\alpha} = \frac{\mathrm{d}\mathscr{L}[\cdot]}{\mathrm{d}\alpha} = \frac{\partial\mathscr{L}[x_1(\alpha),x_2(\alpha),\lambda(\alpha);\alpha]}{\partial\alpha}$

Also, let $f^*(\alpha) := f[x_1(\alpha), x_2(\alpha); \alpha]$ and $g^*(\alpha) := g[x_1(\alpha), x_2(\alpha); \alpha]$

$\frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha} = \frac{\partial f[\cdot]}{\partial x_1}\frac{\mathrm{d}x_1(\alpha)}{\mathrm{d}\alpha} + \frac{\partial f[\cdot]}{\partial x_2}\frac{\mathrm{d}x_2(\alpha)}{\mathrm{d}\alpha} + \frac{\partial f[\cdot]}{\partial\alpha}$

and by the FOCs:

$\Rightarrow \frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha} = \left(\lambda\frac{\partial g[\cdot]}{\partial x_1}\right)\frac{\mathrm{d}x_1(\alpha)}{\mathrm{d}\alpha} + \left(\lambda\frac{\partial g[\cdot]}{\partial x_2}\right)\frac{\mathrm{d}x_2(\alpha)}{\mathrm{d}\alpha} + \frac{\partial f[\cdot]}{\partial\alpha}$

The constraint at the optimal value has to satisfy $g^*(\alpha) = \kappa$ and so that if $\alpha$ changes, the constraint cannot change. Formally,

$\frac{\mathrm{d}g^*(\alpha)}{\mathrm{d}\alpha} = 0$, which implies that $\lambda\frac{\mathrm{d}g^*(\alpha)}{\mathrm{d}\alpha} = 0$. Thus,

$\lambda\frac{\partial g[\cdot]}{\partial x_1}\frac{\mathrm{d}x_1(\alpha)}{\mathrm{d}\alpha} + \lambda\frac{\partial g[\cdot]}{\partial x_2}\frac{\mathrm{d}x_2(\alpha)}{\mathrm{d}\alpha} + \lambda\frac{\partial g[\cdot]}{\partial\alpha} = 0$

$\Rightarrow \lambda\frac{\partial g[\cdot]}{\partial\alpha} = -\lambda\frac{\partial g[\cdot]}{\partial x_1}\frac{\mathrm{d}x_1(\alpha)}{\mathrm{d}\alpha} - \lambda\frac{\partial g[\cdot]}{\partial x_2}\frac{\mathrm{d}x_2(\alpha)}{\mathrm{d}\alpha}$

Using this result, now

$$\Rightarrow \frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha} = \underbrace{\left(\lambda\frac{\partial g[\cdot]}{\partial x_1}\right)\frac{\mathrm{d}x_1(\alpha)}{\mathrm{d}\alpha} + \left(\lambda\frac{\partial g[\cdot]}{\partial x_2}\right)\frac{\mathrm{d}x_2(\alpha)}{\mathrm{d}\alpha}}_{=-\lambda\frac{\partial g[\cdot]}{\partial\alpha}} + \frac{\partial f[\cdot]}{\partial\alpha}$$

which implies that $\frac{\mathrm{d}f^*(\alpha)}{\mathrm{d}\alpha} = \frac{\partial f[\cdot]}{\partial\alpha} - \lambda\frac{\partial g[\cdot]}{\partial\alpha}$. Since $\frac{\partial\mathscr{L}[\cdot]}{\partial\alpha} = \frac{\partial f[\cdot]}{\partial\alpha} - \lambda\frac{\partial g[\cdot]}{\partial\alpha}$, the result follows. ■

## 1.4 Applications

I provide two applications of the Envelope Theorem. The first one provides a a cookbook procedure to apply it. The second example uses the Envelope Theorem to provide an interpretation for the Lagrange multiplier.

### 1.4.1 Consumer Theory

Suppose the existence of two goods, 1 and 2, with prices $p_1$ and $p_2$. A consumer has income $Y$, and has to decide on the level of consumption of goods 1 and 2, denoted by $x_1$ and $x_2$. The optimization problem is:

$$\max_{x_1,x_2} U(x_1, x_2) \text{ subject to } Y = p_1 x_1 + p_2 x_2,$$

wich provides the solutions $x_1^*(Y,p_1,p_2)$ and $x_2^*(Y,p_1,p_2)$.

The value function, which in the context of consumer theory is called the indirect utility function, is

$$U^*(Y,p_1,p_2) := U[x_1^*(Y,p_1,p_2), x_2^*(Y,p_1,p_2)].$$

We want to know how $U^*$ changes when there is a ceteris paribus change in either $Y$ or $p_1$. The Envelope Theorem allows us to do this, without the need to solve the whole optimization problem.

**Procedure to apply the Envelope Theorem**

Step 1. Construct the Lagrangian: $\mathscr{L}(x_1,x_2;Y,p_1,p_2) := U(x_1,x_2) + \lambda[Y - p_1x_1 - p_2x_2]$.

Step 2. Take derivatives of the Lagrangian with respect to each parameter of interest (without evaluating them at the optimal values):

- $\frac{\partial \mathscr{L}(x_1,x_2;Y,p_1,p_2)}{\partial Y} = \lambda$, and
- $\frac{\partial \mathscr{L}(x_1,x_2;Y,p_1,p_2)}{\partial p_1} = -\lambda x_1$.

Step 3. Evaluate each derivative at the optimal value to obtain $\frac{\mathrm{d}U^*(Y,p_1,p_2)}{\mathrm{d}Y}$ and $\frac{\mathrm{d}U^*(Y,p_1,p_2)}{\mathrm{d}p_1}$. Formally,

- $\frac{\mathrm{d}U^*(Y,p_1,p_2)}{\mathrm{d}Y} = \lambda^*(Y,p_1,p_2)$, and
- $\frac{\mathrm{d}U^*(Y,p_1,p_2)}{\mathrm{d}p_1} = -\lambda^*(Y,p_1,p_2)\,x_1^*(Y,p_1,p_2)$.

The second step in particular considerably simplifies the calculations, since it is only the partial derivative with respect to the parameter. In other terms, it does not require plugging in the optimal solution into the Lagrangian and evaluating the total differential—this requires more work, since it includes computing the impact of a parameter change on each optimal consumption.

### 1.4.2 An Interpretation of the Lagrange Multiplier

Suppose that the optimization problem has the following form:

$$\max_{x_1,x_2} f(x_1,x_2) \text{ subject to } \kappa = g(x_1,x_2),$$

where the only parameter is then $\kappa$. The Lagrangian is

$$\mathscr{L} := f(x_1,x_2) + \lambda\left[\kappa - g(x_1,x_2)\right].$$

Denote the solution by $x_1^*(\kappa)$, $x_2^*(\kappa)$, and $\lambda^*(\kappa)$. Moreover, let the value function be $f^*(\kappa)$. By applying the Envelope Theorem, we obtain that

$$\frac{\mathrm{d}f^*(\kappa)}{\mathrm{d}\kappa} = \lambda^*(\kappa),$$

and so $\lambda^*$ gives the impact on the value function when the constraint is relaxed in one unit.

To illustrate the interpretation, suppose that we are analyzing a consumer's problem. Given that the constraint represents a consumer's budget constraint, the result implies that

$$\frac{\mathrm{d}U^*(Y,p_1,p_2)}{\mathrm{d}Y} = \lambda^*(Y,p_1,p_2).$$

The term $\lambda^*(\cdot)$ indicates how the maximum utility of the consumer increases when the consumer has one additional dollar. This is why $\lambda^*$ **represents the marginal utility of income in consumer theory**: how the maximum utility increases when there is a unitary increase in the consumer's income.

## 1.5 Elasticities

It takes some time to get used to work with elasticities. However, it is crucial that you do so due to two reasons. First, it allows us to get rid of measure units, by expressing the derivatives in percentage terms. Additionally, while expressions for some derivatives can be cumbersome, expression for elasticities could be quite neat. This is usually the case when a function is log-linear (i.e., it becomes linear after expressing it in logarithms).

In the next section, I explain the concept of elasticity and the different ways to calculate it. Then, I provide some examples.

### 1.5.1 Definition of Elasticity

To illustrate the concept of elasticity, consider some good with demand $q(p)$ and price $p$. The price elasticity of the demand is denoted by $\varepsilon(p)$, and remember that there are three equivalent ways to expressing it:

[1] $-\frac{\mathrm{d}q(p)/q(p)}{\mathrm{d}p/p}$,

[2] $-\frac{\mathrm{d}q(p)}{\mathrm{d}p}\frac{p}{q(p)}$, or

[3] $-\frac{\mathrm{d}\ln q(p)}{\mathrm{d}\ln p}$.

The negative sign of each term reflects that the elasticity is usually expressed in absolute terms. If $\frac{\mathrm{d}q(p)}{\mathrm{d}p}$ were positive, we would dispense with the negative sign.

The first expression is the definition of elasticity, and the most intuitive one to explain the concept. Rewriting it in terms of finite variations:

$$-\frac{\frac{\Delta q(p)}{q(p)}\times 100}{\frac{\Delta p}{p}\times 100} \tag{1.1}$$

Let's analyze what the denominator in (1.1) tells us. The term $\Delta p/p$ means the variation of $p$ relative to the value of $p$. For instance, consider that $p = 200$ and $\Delta p = 2$. This means that the prices have varied in 2 dollars, starting from a total price of 200. Thus, the variation of two dollars represents a 1% variation, since $\frac{\Delta p}{p}\times 100 = \frac{2}{200}\times 100 = 1$. The interpretation of the numerator in (1.1) is similar. Therefore, $\varepsilon(p)$ is the percentage of variation in the quantities when price varies in 1%. For instance, if $\frac{\Delta q(p)}{q(p)}\times 100 = 8$ so that (1.1) equals 8%, we can conclude that a increase in price of 1% results in decrease of an 8% in quantities.

Although the first expression provides a clear-cut interpretation, it is usually the second and the third expressions that we use for computing elasticities. In particular, we will make extensive use of the third one throughout the course. The key to understand expressions like $-\frac{\mathrm{d}\ln x(p)}{\mathrm{d}\ln p}$ is treating the numerator and denominator like differentials.

To see this, define the function $y(x) := \ln q$. A differential is defined by $\mathrm{d}y = y'(x)\,\mathrm{d}x$, determining that $\mathrm{d}y = \frac{\mathrm{d}x}{x}$ since $y'(x) = \frac{1}{x}$. By using in particular that $y(x) := \ln x$, we get $\mathrm{d}\ln x = \frac{\mathrm{d}x}{x}$. This explains why the numerator in (1.1) can be expressed as either $\frac{\mathrm{d}x(p)}{x(p)}$ or $\mathrm{d}\ln x(p)$.

### 1.5.2 Examples

We provide two examples. The first one shows how the use of logs can simplify considerably the elasticity calculations. Then, we show the computation of derivatives with respect to $\ln p$ when the term $\ln p$ does not appear explicitly.

#### 1.5.2.1 Applying Logs Transformations

Every time you have the product of power functions, you can linearize the equation by applying logs. This makes calculations of elasticities easier.

To fix ideas, let's consider the following demand

$$q(p; A) := Ap^{-\sigma},$$

where $A, \sigma > 0$ are parameters.

The price elasticity of demand can be calculated in two steps. First, we take logs of $x$, determining

$$\ln q(p; A) = \ln A - \sigma \ln p.$$

In a second step, we calculate the elasticity by exploiting that the equation is log-linear. To clearly show this, define $y := \ln q$, $x := \ln p$, and $\alpha := \ln A$. Then, the demand function becomes

$$y = \alpha - \sigma x,$$

which is a linear function. Therefore,

$$\frac{\partial y}{\partial x} = \frac{\partial \ln q\left(p; A\right)}{\partial \ln p} = -\sigma,$$

and so the price elasticity of demand is $\sigma$.

The result points out that the power of any variable is its elasticity when we have a function that is the product of power functions. To provide an additional example of this, notice that the term $A$ trivially has a power of 1. Hence, we can immediately realize that the elasticity of $x$ with respect to $A$ is in fact one. This can be proved formally in the same way as we did with $p$:

$$\frac{\partial y}{\partial \alpha} = \frac{\partial \ln q\left(p; A\right)}{\partial \ln A} = 1.$$

#### 1.5.2.2 The Case where $\ln p$ Does Not Appear Explicitly

Even the equation to be derived does not include terms like $\ln x$ and $\ln p$ explicitly, we can express it in terms of them. For example, let's take the following demand:

$$q\left(p\right) := \exp\left(A - bp\right)$$

where $A, b > 0$.

Applying logs, the function becomes

$$\ln q\left(p\right) = A - bp$$

and so

$$\frac{\mathrm{d} \ln q\left(p\right)}{\mathrm{d}p} = -b$$

If we are interested in the expression $\frac{\mathrm{d}\ln q(p)}{\mathrm{d}\ln p}$, we can start from $\frac{\mathrm{d}\ln q(p)}{\mathrm{d}p}$ and use that $\mathrm{d}\ln p = \frac{1}{p}\mathrm{d}p$. Thus, dividing each side of $\frac{\mathrm{d}\ln q(p)}{\mathrm{d}p}$ by $\frac{1}{p}$:

$$\frac{\mathrm{d} \ln q\left(p\right)}{\frac{\mathrm{d}p}{p}} = -\frac{b}{\frac{1}{p}} \Rightarrow \frac{\mathrm{d} \ln q\left(p\right)}{\frac{\mathrm{d}p}{p}} = \frac{\mathrm{d} \ln q\left(p\right)}{\mathrm{d} \ln p} = -pb,$$

thereby implying that the price elasticity of demand is $pb$.

Another example is when we only have the value of $\frac{\mathrm{d}q(p)}{\mathrm{d}p}$, but we want to obtain

$\frac{d\ln q(p)}{\mathrm{d}p}$.[3] For instance, suppose the following demand:

$$q(p) := A - bp,$$

from which we can easily obtain that $\frac{\mathrm{d}q(p)}{\mathrm{d}p} = -b$. There are two ways to get an expression for $\frac{\mathrm{d}\ln q(p)}{\mathrm{d}p}$. One is applying logs to the original function and then taking the derivative. Another is by exploiting that $\frac{\mathrm{d}q(p)}{\mathrm{d}p} = -b$, so that

$$\frac{\mathrm{d}\ln q(p)}{\mathrm{d}p} = \frac{\frac{\mathrm{d}q(p)}{q(p)}}{\mathrm{d}p} = \frac{\mathrm{d}q(p)}{\mathrm{d}p}\frac{1}{q(p)} = -\frac{b}{q(p)} = -\frac{b}{A - bp}.$$

Notice that we can also play this strategy to obtain an expression for $\frac{\mathrm{d}q(p)}{\mathrm{d}\ln p}$:

$$\frac{\mathrm{d}q(p)}{\mathrm{d}\ln p} = \frac{\mathrm{d}q(p)}{\frac{\mathrm{d}p}{p}} = p\frac{\mathrm{d}q(p)}{\mathrm{d}p} = -bp.$$

[3] Notice that the technique used here is also useful to obtain an elasticity when we have a term like $\frac{\mathrm{d}x(p)}{\mathrm{d}\ln p}$.

# Lecture Note 2

# Basics of Consumer Theory

## 2.1 Introduction

In this set of notes, we cover the basics that economists use to model the behavior of consumers. More complex models of demand use the concepts we are about to present in one or another way.

By a consumer, we mean a person who buys different goods at some fixed market prices and under some monetary restriction (income). A consumer is described by two elements: her preferences and her income constraint. The former is captured by means of a *utility function*, while the latter by a *budget constraint*. Each of them will be formally described in these notes.

Since you probably have seen some of the concepts, I will be focusing on how to describe the concepts formally (i.e. mathematically). In particular, rather than an exhaustive review, I concentrate on the techniques we'll use in the course. Once we grasp a thorough understanding of the basics, in the next lecture note we will move to the consumer optimization problem.

## 2.2 The Utility Function

Let's start by defining some basic concepts and assumptions. To keep matters simple, we will deal with an agent in an economy with 2 goods.

A **consumption bundle** for the consumer is denoted by $\mathbf{x} := (x_1, x_2)$. This is also referred to as a **consumption basket**. We assume that each good $i$ is perfectly divisible and can be consumed in zero or positive quantities. This is formalized by $x_i \in X_i$, where $X_i$ is the domain of consumption of good $i$ and we assume that $X_i := \mathbb{R}_+$ for each $i = 1, 2$. Thus, each consumption bundle belongs to the space $X := X_1 \times X_2$. In the literature, $X$ is referred to as the **consumption space**, and describes the set of feasible bundles. Depending on the context, we could define a consumption space different from $\mathbb{R}^2_+$. For instance, for some utility functions, we do not allow for zero quantities and hence suppose that $X = \mathbb{R}^2_{++}$.

To capture how a consumer assesses the different bundles and makes consumption choices, the theory starts by defining a preference relation $\succsim$. This is a primitive of the model (i.e an element belonging to the model setup), and mathematically is an order relation. Given two bundles $\mathbf{x}'', \mathbf{x}' \in X$, the preference relation tells us if the consumer weakly prefers $\mathbf{x}''$ (denoted by $\mathbf{x}'' \succsim \mathbf{x}'$), weakly prefers $\mathbf{x}'$ (denoted by $\mathbf{x}' \succsim \mathbf{x}''$), or is indifferent between both (denoted by $\mathbf{x}'' \sim \mathbf{x}'$, which means that simultaneously $\mathbf{x}'' \succsim \mathbf{x}'$ and $\mathbf{x}' \succsim \mathbf{x}''$).

In general, working with preference relations is not our first choice to analyze a consumer problem. Rather, we use concepts that turn the problem more tractable. Under some conditions, the order of bundles described by $\succsim$ is the same it could be established through a **utility function**, $U : X \rightarrow \mathbb{R}$. This function takes a consumption bundle $\mathbf{x} \in X$ as an input and returns a value $U(\mathbf{x}) \in \mathbb{R}$ as an output. The number $U(\mathbf{x})$ describes the satisfaction the consumer gets from consuming a basket $\mathbf{x}$, and can be used to order all the bundles. Consequently, the utility function $U$ conveys exactly the same information as the preference relation $\succsim$.

Formally, given two bundles $\mathbf{x}''$, $\mathbf{x}' \in X$, we define the utility as a $U$ such that:

$$\mathbf{x}'' \succsim \mathbf{x}' \text{ iff } U(\mathbf{x}'') \geq U(\mathbf{x}'). \tag{2.1}$$

### 2.2.1 Ordinality of the Utility Function

We have said that the utility function is $U : X_1 \times X_2 \rightarrow \mathbb{R}$. Notice that we are assuming that the co-domain of $U$ is $\mathbb{R}$, thus allowing for negative numbers. But what is the interpretation? How is it possible that a consumer gets negative utility from consuming nonnegative quantities of each good? The answer is simple: it does not matter. The only relevant aspect of a utility function is to allow for comparisons between bundles and determine which one she prefers, as in (2.1). Put it differently, the specific number attached to the utility function has no meaning, as long as it describes the same ranking (i.e. preferences) between the bundles. In formal terms, it is said that the utility function is *ordinal*, rather than cardinal.

From (2.1) we can also infer that only inequalities are relevant to rank the bundles. Statements like "the consumer gets twice the utility from the bundle $\mathbf{x}''$ relative to $\mathbf{x}'$ " are not necessary for describing the consumer's preferences. This is quite fortunate, since utility is quite an abstract concept and hardly measurable.

This fact has specific implications regarding utility functions. One of the most important is that the utility function is not uniquely defined. Any strictly increasing transformation of $U$ changes the scale in which utility is measured, but preserves the same order of bundles given by 2.1. Since having the same order of baskets satisfies the definition of a utility function given by 2.1, any monotone transformation defines a new utility function for the consumer. This is why any particular utility function should be actually referred to as ***a*** consumer's utility function, rather than ***the*** consumer's utility function.

Formally, let $g$ be a strictly increasing positive function, which means that $g' > 0$. Defining $V(\mathbf{x}) := g[U(\mathbf{x})]$, if $U(\mathbf{x}'') \geq U(\mathbf{x}')$ then it is necessarily true that $V(\mathbf{x}'') \geq V(\mathbf{x}')$. This follows because $g$ is strictly increasing and so $g[U(\mathbf{x}'')] \geq g[U(\mathbf{x}')]$.

Examples of monotone transformations are $V(\mathbf{x}) := U(\mathbf{x}) + a$ where $a$ is a constant (positive or negative), and $V(\mathbf{x}) = bU(\mathbf{x})$ where $b > 0$. In the first example, we can see more clearly why the utility function could be negative and yet make sense. If $a < 0$, $V$ might be negative, but the only thing that matters is whether $V(\mathbf{x}'') \geq V(\mathbf{x}')$, which is all that matters in the end—defining a utility $V$ with $a < 0$ would not affect the ranking.

But, then, someone might ask: which one is the "original" utility function? is it $U$ or $V$? The question is misdirected, since there is no such a thing as a primitive utility function. $U$, $V$, or any other monotone transformation represents a utility function for the consumer. It is like asking if Bruce Wayne is the Batman or the Batman is Bruce Wayne—both are the same person, there is no such a thing as the "original" person.

## 2.3 The Cobb Douglas

Before delving into several axioms that a consumer could satisfy, we introduce the most common functional form used in the literature: the Cobb Douglas utility function. This will help us illustrate some properties of utility functions, since the Cobb Douglas satisfies all of them.

The Cobb Douglas is defined by

$$U\left(x_1, x_2\right) := \left(x_1\right)^{\alpha_1}\left(x_2\right)^{\alpha_2},$$

where $\alpha_1, \alpha_2 > 0$. The domain is usually $X := \mathbb{R}_{++}{}^2$ to avoid some issues when we apply a log monotone transformation (recall that logs are not defined for zero values). Later in the course, we show this is not an issue at all, since the bundle $(0,0)$ (or any bundle were one good is not consumed) is never an optimal choice for the consumer under this utility function.

When researchers work with the Cobb Douglas, they do not always use the form $U\left(x_1, x_2\right) := \left(x_1\right)^{\alpha_1}\left(x_2\right)^{\alpha_2}$. Rather, they specify the utility function with the logarithmic transformation. The logarithm function represents a monotone transformation due to the following. Since $g\left(U\right) := \ln U$ implies that $g'\left(U\right) = \frac{1}{U}$, and $\left(x_1\right)^{\alpha_1}\left(x_2\right)^{\alpha_2} > 0$ for any $x_1, x_2 > 0$, then $U > 0$ and so $g'\left(U\right) > 0$. Then, we know that the utility represents the same preferences.

Specifically, applying logs to $U$ we obtain

$$\widetilde{U}\left(x_1, x_2\right) := \alpha_1 \ln x_1 + \alpha_2 \ln x_2.$$

Notice that the logarithm function is only defined for positive values. This is why we have assumed $x_1, x_2 > 0$, so that zero quantities of any good are not part of the domain.

Remember two properties of the logarithmic function:

- $\ln\left(x_1 x_2\right) = \ln\left(x_1\right) + \ln\left(x_2\right)$
- $\ln\left(x^{\alpha}\right) = \alpha \ln x$

Researchers go even further and work with the case in which $\alpha_1 + \alpha_2 = 1$, thereby defining the utility function by $\widetilde{U}(x_1, x_2) := \alpha \ln x_1 + (1-\alpha)\ln x_2$ with $\alpha \in (0,1)$. This is possible because we can apply an additional monotone transformation to the logs by dividing $\widetilde{U}$ by $\frac{1}{\alpha_1+\alpha_2}$. Thus, the coefficients of the new utility function become $\frac{\alpha_1}{\alpha_1+\alpha_2}$ and $\frac{\alpha_2}{\alpha_1+\alpha_2}$, whose sum equals one. By calling $\alpha := \frac{\alpha_1}{\alpha_1+\alpha_2}$ and noticing that $1-\alpha = \frac{\alpha_2}{\alpha_1+\alpha_2}$, it is common to define the Cobb Douglas utility function as

$$\widetilde{U}(x_1, x_2) := \alpha \ln x_1 + (1-\alpha)\ln x_2,$$

where $\alpha > 0$ and $X_1 \times X_2 := \mathbb{R}_{++}{}^2$.

## 2.4 Axioms

We present different axioms that a consumer can satisfy. Axioms are properties taken as true within the model, which describe the decision process of consumers.

We break down the axioms into two. First, we present those axioms that a consumer endowed with a utility function always satisfies: completeness and transitivity. These two properties define what we understand by a **rational consumer**.

Then, we proceed to define three axioms that not all the utility functions satisfy. They are differentiability, strong monotonicity, and strict convexity.

### 2.4.1 The Definition of Rationality

We present the axioms of completeness and transitivity. They just follow by the definition of a function and the fact that the function takes values belonging to the real numbers. When the choices of a consumer can be described through a utility function, these two properties are always satisfied. They define what we understand as rationality in consumer theory.

1. Completeness This property means that **the consumer can compare any two bundles and say which one she prefers**. Given a bundle $\mathbf{x} \in X$, the consumer can *always* assign a value $U(\mathbf{x})$ to the satisfaction she derives from it. This just follows by the mere definition of a function: any element of the domain $\mathbf{x} \in X$ has a value $U(\mathbf{x})$

attached. Consequently, the consumer can always say which bundle she prefers between two options. Formally, for any two bundles $\mathbf{x}'$ and $\mathbf{x}''$, it is always true that she prefers one of them (either $U(\mathbf{x}') > U(\mathbf{x}'')$, $U(\mathbf{x}') < U(\mathbf{x}'')$) or she is indifferent between both (that is, $U(\mathbf{x}') = U(\mathbf{x}'')$). In other terms, if you ask the consumer if she prefers $\mathbf{x}'$ and $\mathbf{x}''$, she will never answer that she is not sure.

**2. Transitivity** This is the distinctive feature of the utility function, and justifies that we can use describe a consumer as rational. Mathematically, it states that if $U(\mathbf{x}''') > U(\mathbf{x}'')$ and $U(\mathbf{x}'') > U(\mathbf{x}')$, then it is necessarily true that $U(\mathbf{x}''') > U(\mathbf{x}')$. This property just follows by property of the real numbers (recall that the utility function is real-valued since it has a co-domain $\mathbb{R}$). To see this, define the utility values $c := U(\mathbf{x}''')$, $b := U(\mathbf{x}'')$ and $a := U(\mathbf{x}')$. Since they are numbers, it is always true that if $c > b$ and $b > a$, then $c > a$. Some people consider this axiom a strong assumption. In a lot of circumstances, people could make decisions that violate transitivity. Using the same notation, it could happen that the consumer sometimes prefers $\mathbf{x}'''$ and prefers $\mathbf{x}'$ in a different context, even when she claims that $c > a$. Violations of transitivity have shown up in different lab experiments. However, as a first approximation to consumer theory, transitivity is a reasonable behavioral postulate.

## 2.5 Other Consumer Axioms

There are other properties that can be defined, in addition to completeness and transitivity. Unlike these two properties, the ones we will present might or not be satisfied by a specific utility function.

**3. Differentiability** This assumption allows us to use the tools of differential calculus. It is immediate to note that the Cobb Douglas is differentiable, since it is the product of two power functions. However, as we will see in the course, there are several standard utility functions that are not differentiable.

When differentiability holds, we can define the concept of **marginal utility of good** $i$, given by $\frac{\partial U(x_1,x_2)}{\partial x_i}$. This provides information on how infinitesimal increases in the consumption of good $i$ impact the utility function.

> **Remark** *Remember how you should interpret a derivative. Let's take* $\frac{\partial U(x_1,x_2)}{\partial x_1}$ *to illustrate this. Since it is a partial derivative, you are assuming that* $x_2$ *remains fixed. Thus, the term indicates how* $U$ *varies when* $x_1$ *varies infinitesimally,* but expressed in terms of a unitary change of $x_1$. *Put it differently, if* $\frac{\partial U(x_1,x_2)}{\partial x_1} = 2$, *we let* $x_1$ *vary in a small amount, but express the result as if* $x_1$ *had increased in one unit. In that case, the variation in utility would be 2.*

**4. Strong Monotonicity** It means that **greater consumption of any good increases utility**. The property implies what is known as **non-satiation**: given a specific bundle $\mathbf{x}'$, it is always possible to find another bundle $\mathbf{x}''$ that gives more utility. With strong monotonicity we are ruling out cases where a good could become a "bad" if it is consumed in excess.

When the utility function is differentiable, we can define **strong monotonicity** by the existence of a positive marginal utility for each good. Formally, $\frac{\partial U(x_1,x_2)}{\partial x_1} > 0$ and $\frac{\partial U(x_1,x_2)}{\partial x_2} > 0$ for any $x_1, x_2 \in X_1 \times X_2$. This indicates that when more quantity of each good is consumed, the consumer's utility is higher.

For example, the marginal utility of good 1 in the Cobb-Douglas case is

$$\frac{\partial U}{\partial x_1} = \alpha_1 (x_1)^{\alpha_1 - 1} (x_2)^{\alpha_2} > 0,$$

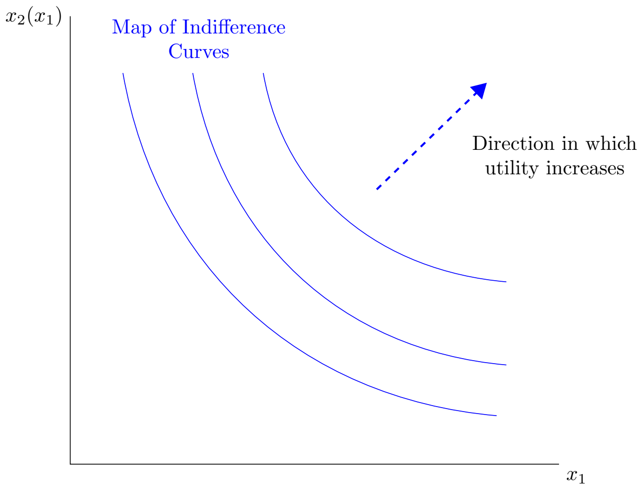
where the sign is positive since $\alpha_1 > 0$. By the same token, $\frac{\partial U}{\partial x_2} = \alpha_2 (x_1)^{\alpha_1} (x_2)^{\alpha_2 - 1} > 0$ when $\alpha_2 > 0$.

**5. Strict Convexity of $\succsim$ (Strict Quasiconvaty of $U$)** This axiom is a necessary condition for **the consumer to prefer diversifying consumption**. A diversified basket has a little bit of every good, rather than a lot of some goods and none of others.

Showing how **strict convexity of preferences** is reflected in the utility function requires a little bit more work, since we first need to introduce some preliminary concepts.

First, we need to define indifference curves: all the combinations of $x_1$ and $x_2$ that provide the same level of utility $U_0$, where $U_0$ is some fixed number. Since $U_0$ can take different values, this allows us to define the indifference map, which is the set of indifference curves. This is shown in Figure IC.

**Figure 2.1.** *Example of Indifference Curve*




You have probably seen the concept of an indifference curve before and analyzed it graphically. However, how can we characterize the indifference curves algebraically? To do this, fix a level of utility at $U_0$. The combinations of $x_1$ and $x_2$ that give the utility $U_0$ are by definition $U_0 = U(x_1, x_2)$. From this, we can establish a function $x_2(x_1)$ that tells us, for a feasible quantity $x_1$, what is the quantity $x_2$ such that the consumption of the bundle $(x_1, x_2(x_1))$ provides a utility $U_0$.

In the case of a Cobb Douglas, we know that $U_0 = (x_1)^{\alpha_1}(x_2)^{\alpha_2}$ for a given level of utility $U_0$. Working out the equation, we obtain that

$$x_2(x_1) = (U_0)^{\frac{1}{\alpha_2}}(x_1)^{-\frac{\alpha_1}{\alpha_2}}.$$

If we drew that function for different values of $U_0$, we would obtain a similar graph to Figure IC.

To characterize indifference curves when we have a general utility function, we proceed by characterizing the slope (first derivative) and the curvature (second derivative) of the function $x_2(x_1)$. To do this, we proceed by totally differentiating $U_0 = U[x_1, x_2(x_1)]$ for a fixed value of $U_0$. In this way, we allow for changes in $x_1$ and $x_2$ under the restriction that $U_0$ does not vary.

For the analysis, assume that strong monotoniticy holds, so that each marginal utility is positive. The differential is:

$$\underbrace{\mathrm{d}U_0}_{=0} = U'_{x_1} + U'_{x_2} \frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1},$$

$$\Rightarrow \frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1} = -\frac{U'_{x_1}[x_1, x_2(x_1)]}{U'_{x_2}[x_1, x_2(x_1)]} < 0.$$

From this, we can see that every time a consumer satisfies strong monotonicity, then $\frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1}$, which is the slope of the indifference curve, is negative. The slope of the indifference curve is also known as the **marginal rate of substitution** (MRS). The MRS provides an answer to the following question: starting from a point $(x_1, x_2(x_1))$, if we increase the consumption of good 1 in one unit, in which amount the good 2 has to vary in order to keep the utility level at the same level?

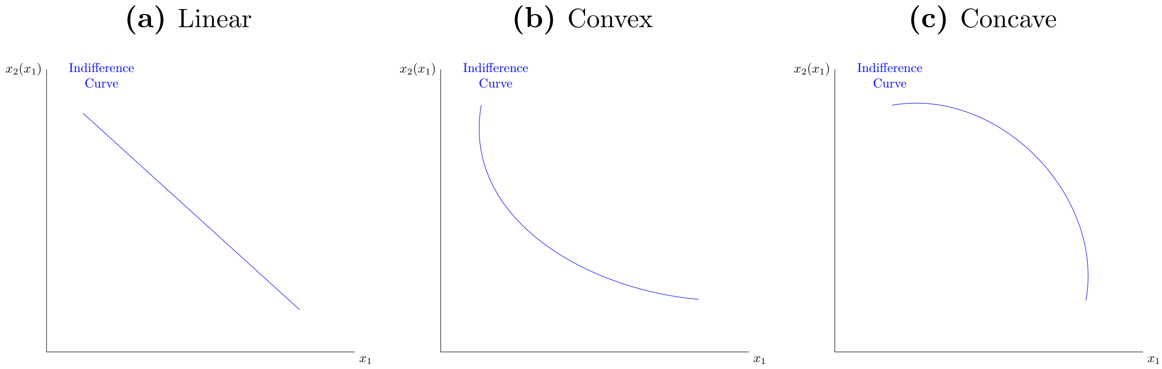
The negative slope indicates that there is a trade-off. By strong monotonicity, a greater consumption of any good increases the utility. By this reason, the consumer would get more utility if $x_1$ increases, and so $x_2$ necessarily has to decrease to restore the level of utility to $U_0$. This is in fact why the slope is called marginal rate of *substitution*: it provides information about how much the consumer is willing to give up of one good, to obtain one more unit of the other good.

For the case of the Cobb Douglas, we already know that $x_2(x_1) = (U_0)^{\frac{1}{\alpha_2}} (x_1)^{-\frac{\alpha_1}{\alpha_2}}$, so that the MRS follows by just taking the derivative with respect to $x_1$:

$$\frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1} = -\frac{\alpha_1}{\alpha_2} (U_0)^{\frac{1}{\alpha_2}} (x_1)^{-\frac{\alpha_1}{\alpha_2}-1} < 0.$$

Once we have defined indifference curves and the MRS, let's come back to what was our original interest: the concept of strict convexity of preferences (SCP). So far, we have only shown that, when preferences are strongly monotone, the MRS is negative. But indifference curves with a negative slope are consistent with any of the three shapes described in Figure 2.2. Depending on the sign of the MRS's slope (i.e. the second derivative of the indifference curve), we can have the three possible shapes of the indifference curve shown in Figure 2.2.

**Figure 2.2.** *Possible Types of Indifference Curves when Slope is Negative*




SCP represents the case where the MRS has a positive slope. Thus, SCP means that the agent has strictly convex indifference curves. This corresponds to the Subfigure 2.3b. Formally,

$$\text{strictly convex preferences: } \frac{\mathrm{d}^2 x_2(x_1)}{\mathrm{d}x_1^2} > 0.$$

For instance, in the case of the Cobb Douglas, we can take the derivative of $\frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1}$ with respect to $x_1$ and obtain

$$\frac{\mathrm{d}^2 x_2(x_1)}{\mathrm{d}x_1^2} = \left(\frac{\alpha_1}{\alpha_2} + 1\right) \frac{\alpha_1}{\alpha_2} (U_0)^{\frac{1}{\alpha_2}} (x_1)^{-\frac{\alpha_1}{\alpha_2} - 2} > 0.$$

What is the implication of strictly convex indifference curves in terms of the utility function? The following can be proven.

**Result 2.1** *The agent has strictly convex indifference curves iff the utility function is strictly quasiconcave.*

We have illustrated in previous classes that quasiconcavity can be proven in different ways. One way is through the sign of the Hessian matrix, but this is not the only way. In fact, directly showing that the indifference curves are strictly convex, as we did for the Cobb Douglas, is sometimes easier.

### 2.5.1 Interpretation of Strictly Convex Preferences (OPTIONAL)

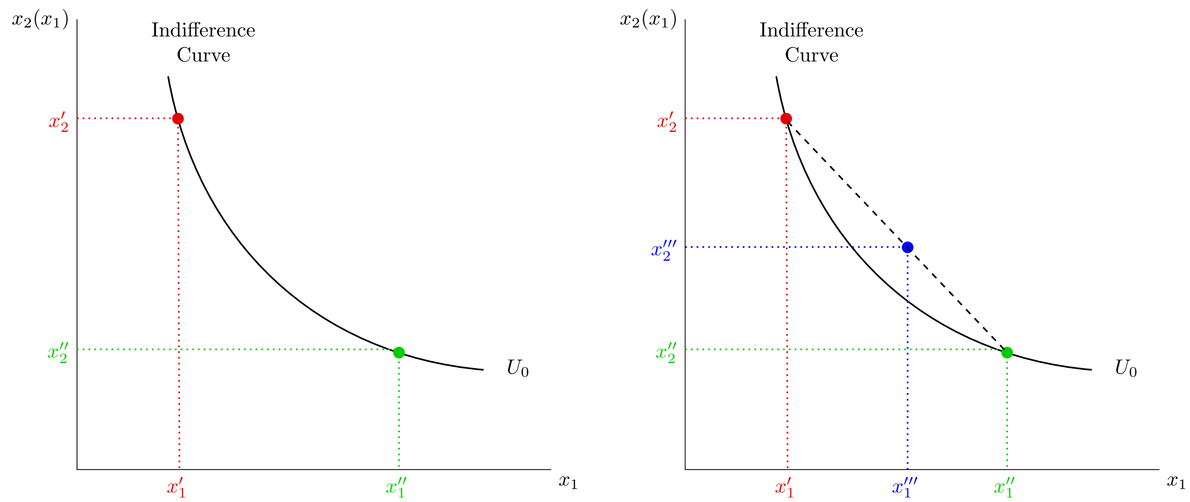
You have probably seen in previous courses why strict convexity of the indifference curves entails that the consumer diversifies consumption. Here, I briefly review this property, just in case you do not remember it.

Suppose an indifference curve with utility $U_0$ and two bundles that belong to that curve. Let's refer to these bundles by $\mathbf{x}' := (x_1', x_2')$ and $\mathbf{x}'' := (x_1'', x_2'')$. Now suppose we construct a third bundle $\mathbf{x}'''$, which is a linear combination of those bundles. Specifically, let $x_1''' := \alpha x_1' + (1-\alpha)\, x_1''$ and $x_2''' := \alpha x_2' + (1-\alpha)\, x_2''$, where $\alpha \in (0,1)$.[1] Since $\alpha \in (0,1)$, the bundle $\mathbf{x}'''$ represents a basket with intermediate consumptions of $\mathbf{x}'$ and $\mathbf{x}''$.

For instance, consider $\mathbf{x}' := (2,1)$ and $\mathbf{x}'' := (1,2)$, and suppose these bundles provide the same utility $U_0$. The basket $\mathbf{x}'''$ would be then $\mathbf{x}''' = (1+\alpha, 2-\alpha)$, and so $x_1''', x_2''' \in (1,2)$. SCP means that the basket $(x_1''', x_2''')$ for any $\alpha \in (0,1)$ will provide more utility than $U_0$, so that it is preferred to $(x_1', x_2')$ and $(x_1'', x_2'')$. Graphically, this can be seen in Figure 2.3.

**Figure 2.3.** *Strictly Convex Indifference Curves*

**(a)** Two Points **(b)** Convex Combination




Let's provide some intuition through the graph, and then see the economics behind it. In Figure 2.3, we are assuming one specifically $\alpha$ to represent $(x_1''', x_2''')$. By choosing

[1]Notice that, in fact, $x_2''' := x_2\,(x_1''')$.

a different $\alpha \in (0,1)$, we can obtain any bundle along the dashed line. Thus, the dashed line represents all the possible bundle $\mathbf{x}'''$ we can think of. Since the dashed line is above the indifference curve of $U_0$, any conceivable bundle $\mathbf{x}'''$ provides more utility than $\mathbf{x}'$ or $\mathbf{x}''$.

To get some intuition for this result, consider the following example. Let good 1 be water (or any type of beverage), while good 2 is food. Suppose an extreme situation where the consumer is considering bundles where she only consumers one good. These bundles are represented by $(x_1', 0)$ and $(0, x_2'')$, where $x_1', x_2'' > 0$. Now consider a basket $\mathbf{x}''' := (\alpha x_1', (1-\alpha) x_2'')$ with $\alpha \in (0,1)$. The basket $\mathbf{x}'''$ captures the situation where she consumes strictly positive quantities of each good. Then, no matter what the values of $x_1'$ and $x_2''$ are, an the agent that has SCP will prefer to consume the basket $\mathbf{x}'''$ for any $\alpha \in (0,1)$. This means that an agent having SCP derives more utility from diversifying consumption, than from consuming bundles with extreme quantities.

## 2.6 Budget Constraint

For our purposes, the description of the different elements involved in Consumer Theory has as a final goal to determine what a consumer chooses in the market. So far, we have described the utility function. If the consumer is not constrained (and she satisfies strong monotonicity), the solution would be trivial: she would always like to consume as much as she can of each good.

However, consumers are always constrained, either by the time or income at their disposal. In consumer theory, we consider the case where each consumer has an income constraint (or, more generally, a wealth constraint). Thus, a consumer has to decide her consumption, knowing that her total expenditure cannot be greater than her income.

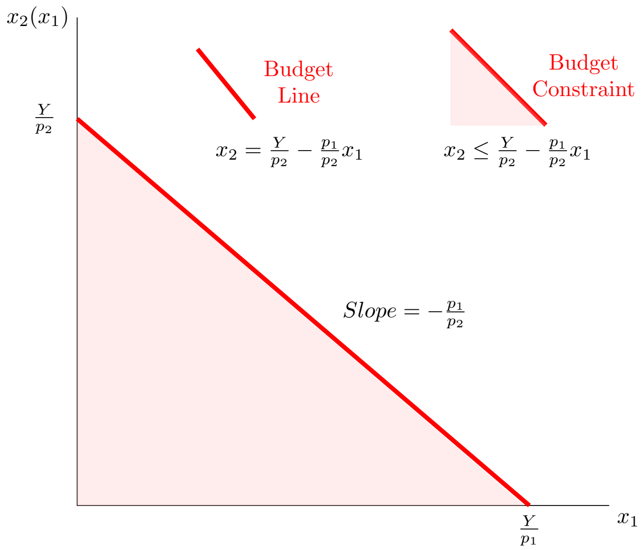
Although, to some extend, income is endogenously determined by each agent through her labor choices, Consumer Theory proceeds in a simple way by assuming that income is exogenous. Given two goods with a vector of prices $(p_1, p_2)$ and income $Y$, the **budget**

**constraint** is defined by

$$\underbrace{Y}_{\text{income}} \geq \underbrace{p_1 x_1 + p_2 x_2}_{\text{expenditure}}. \tag{2.2}$$

The **budget set** defines all the possible combinations $(x_1, x_2)$ that the consumer can afford given her income $Y$. Likewise, we refer to the **budget line** as the combinations $(x_1, x_2)$ when (2.2) holds with equality, which it describes the bundles that exhaust the consumer's income. We can represent these elements as in Figure 2.4.

**Figure 2.4.** *Budget Constraint*




The relevance of the budget line arises since we will assume consumers that always satisfy strong monotonicity for at least one good. Thus, the consumer will always spend all her income, and she will choose a basket among those along the budget line.

The slope of the budget line indicates the *objective* rate at which the market allows her to substitute goods. To determine the slope of the budget line, we consider the case of two goods. The budget line is $Y = p_1 x_1 + p_2 x_2$ and determines a relation $x_2(x_1)$. This relation provides the maximum amount of good 2 that the consumer can afford given her income, given some consumption for good 1. Formally,

$$x_2(x_1) := \frac{Y}{p_2} - \frac{p_1}{p_2} x_1,$$

and so the slope is

$$\frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1} = -\frac{p_1}{p_2}.$$

The term $\frac{Y}{p_2}$ is the value when $x_1 = 0$. Thus, it represents the maximum possible amount of good 2 that the consumer can afford.

By the same token, we could have determined the budget line as $x_1(x_2)$. In this case, the slope would be $\frac{-p_2}{p_1}$. In addition, $\frac{Y}{p_1}$ would the maximum possible quantity of good 1.

There is another way to find out the slope of the budget line. This is more similar to the approach we have used for the derivation of the MRS. It consists in totally differentiating the expression with income fixed, so that $\underbrace{\mathrm{d}Y}_{=0} = p_1\mathrm{d}x_1 + p_2\frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1}$. Thus, $\frac{\mathrm{d}x_2(x_1)}{\mathrm{d}x_1} = -\frac{p_1}{p_2}$.

The slope of the budget constraint indicates how much quantity of the good 2 the consumer has to stop consuming if she wants to get one more unit of good 1. The fact that this is equal to $-\frac{p_1}{p_2}$ reflects two aspects. First, that to buy one more unit of good 1 the consumer needs to get enough money to afford its price $p_1$ (numerator). Second, that for each unit of good 2 not consumed, the consumer has an additional amount $p_2$ of money available (denominator).

### 2.6.1 Taxes in the Budget Constraint

Throughout the course, we will study how taxes impact a consumer's decisions as an application. Next, we show how three typical types of taxes are reflected in the budget constraint.

**1. Income Tax** Income taxes are usually set as a percentage over total income. Specifically, suppose that the government establishes that the agent has to pay a fraction $\tau \in (0,1)$ over the total income she earns. For instance, if the taxes are 15% of the total income, then $\tau = 0.15$. The budget constraint requires that we express this income in terms of disposable income. Thus,

$$Y(1-\tau) = p_1x_1 + p_2x_2.$$

**2. Value-Added Tax** This type of tax is also known as *ad-valorem* tax. It comprises those taxes applied to the consumption of goods with a tax base given by prices (i.e. value). More specifically, every time a consumer buys a unit of the good, she has to pay an additional fraction $\tau$ over the value of the product. One typical example of this tax is the 5% of GST we pay in Edmonton for each product we consume.

In terms of the budget constraint, we can understand how they are incorporated in two different ways. Suppose a value-added tax on good 1. For each unit of good consumed, you have to pay an additional fraction $\tau$ out of the total value. Hence, the consumer pays a total of $p_1 (1 + \tau)$ for each unit consumed. Equivalently, we can think that, out of the total purchase, the consumer has to pay an additional $\tau$ for the total value. Thus, the expenditure of the consumer when she spends $p_1 x_1$ is $p_1 x_1 (1 + \tau)$. Whatever the interpretation we use, the budget line becomes:

$$Y = (1 + \tau) p_1 x_1 + p_2 x_2.$$

**3. Tax per Unit Sold** It is also a tax applied to the consumption of goods, but with the tax base given by quantities. Thus, instead of adding a fraction of the value to the total payment, the consumer pays for each unit consumed. This is captured by an additional expenditure of $\tau$ monetary units, such that the consumer pays $p_1 + \tau$ for each unit consumed. Formally, the budget constraint becomes

$$Y = (p_1 + \tau) x_1 + p_2 x_2.$$

## 2.7 Exercises

[1] Suppose a utility function $U(x_1, x_2) := \ln(x_1) + x_2$, where $x_1 > 0$ and $x_2 \geq 0$.

(a) Is $U$ strongly monotone in each good? What does strongly monotonicity mean?

(b) Starting from an indifference curve with utility $U_0$, find an expression for $x_2(x_1)$.

(c) Determine the slope of the indifference curve and its sign.

(d) Does $U$ represent strictly convex preferences? What does this property imply for the analysis?

[2] Repeat the exercise 1) for the following utility function:

(a) $U(x_1, x_2) := x_1 + x_2$

(b) $U(x_1, x_2) := (x_1)^2 + x_2$

**Some Answer Keys**: 1) strictly convex preferences, 2a) convex preferences 2b) not convex (in fact, strictly concave preferences)

# Lecture Note 3

# The Consumer's Optimization Problems

## 3.1 Introduction

In Lecture Note 1, we covered a consumer's preferences and the concept of budget constraint. This had the goal of preparing the ground to analyze consumer's choices.

In this set of notes, we will study two optimization problems. The first one is the so-called utility maximization problem: choosing a consumption basket that maximizes utility subject to a budget constraint. This gives rise to the Marshallian demands.

After this, we study the expenditure minimization problem. This refers to the quantity demanded that minimizes the expenditure to get a specific level of utility. From this problem, we determine the Hicksian demands. The analysis of this case is more abstract, but it has an intimate relation with the utility maximization problem. In this sense, its main aim is to understand the maximization problem in more detail.

We restrict the analysis to the case of two goods. Also, throughout the analysis, we consider a utility function $U$ that is well behaved. By a well-behaved utility function, we mean that it satisfies all the five properties presented in the previous lecture. In addition, we assume Inada conditions to rule out boundary solutions.

**Definition 3.1:** *A* ***well-behaved utility function*** $U$ *satisfies:*

*[1] Completeness.*

*[2] Transitivity.*

*[3] Differentiability (i.e., $U \in \mathcal{C}^2$).*

*[4] Strong monotonicity (i.e. positive marginal utilities for each good).*

*[5] Strict quasiconcavity (i.e. strictly convex preferences).*

*[6] Inada Conditions:* $\lim\limits_{x_i \to 0} U'_{x_i} = \infty$ *for* each *good* $i$.

Remember that completeness and transitivity are satisfied by any utility function. So, strictly speaking, we are assuming differentiability, strong monotonicity strict quasiconcavity, and Inada conditions, relative to the axioms of the previous lecture note.

Inada conditions is a new property that we did not consider previously. Its goal is to ensure that there are no boundary solutions. Notice we suppose that the Inada condition holds for *each* good $i$. If it happens that it holds for, let's say, good 1 but not for good 2, we cannot rule out that there is a boundary solution in which good 2 is not consumed. In other terms, the Inada condition for $i$ only rules out that $x_i = 0$ is not a solution.

Keep in mind that not all the utility functions are well-behaved. In fact, we will consider several functional forms that are not in the next lecture note.

## 3.2 The Utility Maximization Problem (UMP)

The UMP is defined by:

$$\max_{(x_1,x_2)\in X_1\times X_2} U(x_1,x_2) \text{ subject to } Y = p_1x_1 + p_2x_2,$$

where $Y$ is the consumer's income, and $p_1$ and $p_2$ are the price of each good.[1]

The budget constraint holds with equality by the assumption of strong monotonicity. It reflects that, since consuming more of each good makes the consumer better off, it is never optimal to spend less than the total income. Thus, the consumer will always choose a bundle along the budget line (i.e., a basket that exhausts all her income).

Intuitively, an optimal solution with $Y < p_1x_1 + p_2x_2$ could arise if goods become a "bad" after a certain consumption threshold. In that case, it may be optimal not spending all the income, since a basket along the budget line could actually make the consumer worse off.

The first conclusion we can get from the UMP follows by simple inspection of the income constraint: only relative prices matter for consumption choices. Thus, the consumers' decision would be the same if *all* prices increase an $\alpha\%$, since $\frac{p_1}{p_2}$ would not be affected.[2]

[1] We can assume that each $X_i$ is either $\mathbb{R}_+$ or $\mathbb{R}_{++}$. Since the utility function well behaved, zero consumption of any good is not a solution.

[2] Formally, suppose that all prices increase an $\alpha\%$, such that the price of good $i$ changes from $p_i$ to

To see this, starting from $Y = p_1 x_1 + p_2 x_2$, we can divide both sides by $p_2$ and determine that

$$\frac{Y}{p_2} = \frac{p_1}{p_2} x_1 + x_2.$$

thereby showing that what matters is not $p_1$ and $p_2$, but $\frac{p_1}{p_2}$.

Although the result is mathematically simple, notice that it has deep implications in terms of the consumer's decision process. It says that the consumer has no monetary illusion. To understand what this means, consider a situation where Canada decides to change the name and denomination of its currency. Instead of Canadian dollars (CAD), now the currency is called Maples and 1 CAD has the value of 3 Maples. No monetary illusion means that consumers will not change their decisions by such a policy, since it merely entails a change in all the prices (including income) of 300%. Thus, the budget constraint in CAD would be multiplied by 3,

$$3Y = (3p_1)\, x_1 + (3p_2)\, x_2,$$

$$\Rightarrow Y = p_1 x_1 + p_2 x_2,$$

but the budget constraint would remain the same.

An example of monetary illusion could be following an increment in salary that only represents an inflation adjustment. This type of wage adjustment would actually make the purchasing power of your salary be the same (i.e., the person can buy the same goods with her new wage, given the annual increase in prices). Nonetheless, the person could think she is richer, if she compares the nominal wage before and after the increment. The UMP rules out cases like this.

### 3.2.1 General Solution

Given that we are dealing with a well-behaved utility function, it can be shown that the solution to the UMP is interior and unique. Uniqueness follows by strict quasiconcavity,

$p_i\,(1+\alpha)$. Then, the relative prices would not change, since they are given by $\frac{p_1(1+\alpha)}{p_2(1+\alpha)}$, which equals $\frac{p_1}{p_2}$. Only if prices change in different proportions, and so relative prices vary, is that the consumer would change her consumption decisions.

while Inada conditions rule out corner solutions (more specifically, a zero consumption of any good). Furthermore, since the utility function is differentiable, we can make use of the Lagrange technique to characterize the solution.

**Remark** *It is worth emphasizing that not all the functional forms satisfy differentiability, strong monotonicity, strict quasiconcavity, and Inada type conditions, as we are assuming. In case at least one of these properties does not hold, the solution is not necessarily unique and interior. Moreover, we do not necessarily can characterize its solution by using the Lagrange technique. Keep in mind this for the next lecture notes, where we consider specific utility functions that do not satisfy all of these properties.*

The Lagrangian is:

$$\mathscr{L}(x_1, x_2, \lambda; Y, p_1, p_2) := U(x_1, x_2) + \lambda [Y - p_1 x_1 - p_2 x_2].$$

Remember that the Lagrangian is not only a function of $x_1$ and $x_2$, but also of the artificial variable $\lambda$. Once we construct the Lagrangian, we proceed as in any other optimization problem. Specifically, the FOCs are[3]:

$$\text{FOC:} \begin{cases} \frac{\partial \mathscr{L}(\cdot)}{\partial x_1} = \frac{\partial U(x_1,x_2)}{\partial x_1} - \lambda p_1 = 0, \\ \frac{\partial \mathscr{L}(\cdot)}{\partial x_2} = \frac{\partial U(x_1,x_2)}{\partial x_2} - \lambda p_2 = 0, \\ \frac{\partial \mathscr{L}(\cdot)}{\partial \lambda} = Y - p_1 x_1 - p_2 x_2 = 0. \end{cases}$$

Through the FOCs, we obtain the solution to the UMP. This gives the optimal endogenous variables (including $\lambda$), as functions of the parameters of the model: $x_1^*(Y, p_1, p_2)$, $x_2^*(Y, p_1, p_2)$, and $\lambda^*(Y, p_1, p_2)$. The optimal demands of the UMP are called **Marshallian demands.**

Furthermore, the solution allows us to obtain the optimal utility function. In the optimization terminology, this is called *the value function.* In the context of consumer theory, it is known as the **indirect utility function**, and is given by

$$U^*(Y, p_1, p_2) := U[x_1^*(Y, p_1, p_2), x_2^*(Y, p_1, p_2)].$$

[3]Notice that, since we are assuming that the utility function is strictly increasing (by strong monotonicity) and strictly quasiconcave (by strict convexity of the preferences), there is no need to check the second-order conditions to ensure that we are in fact maximizing.

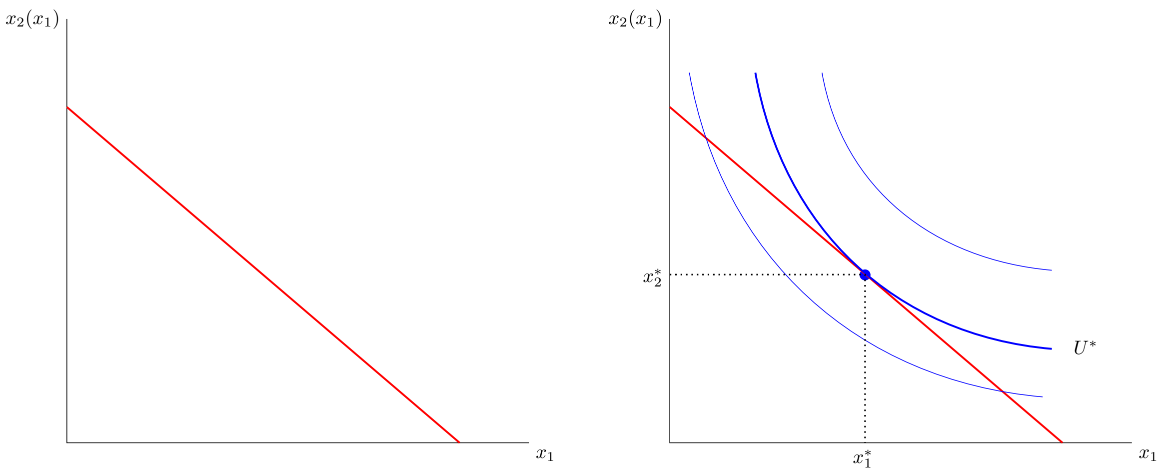
> **Remark** *Keep in mind that $U^*$ and $U$ are two different functions: $U^*$ is a function of the parameters $(Y, p_1, p_2)$, while $U$ of $(x_1, x_2)$.*

### 3.2.2 Interpreting the Optimality Condition

The optimization problem can be graphically depicted in the following way. In Figure 3.1a, there is a given budget line. The optimization problem means finding a basket that provides the highest utility possible. In Figure 3.1b, this is identified by drawing several indifference curves, until we find the one that gives the highest utility possible among those that satisfy the income constraint.

**Figure 3.1.** *UMP*

(a) Constraint (b) Optimization Problem



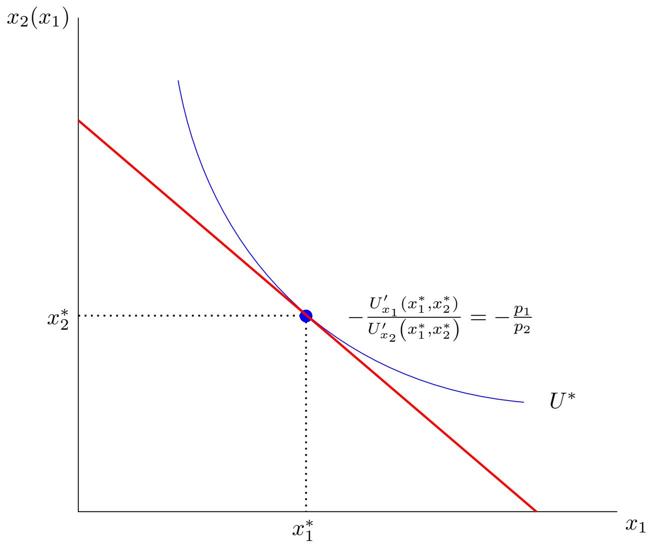

How can we characterize the basket that provides the highest utility? Take the equations $\frac{\partial \mathscr{L}(\cdot)}{\partial x_1} = 0$ and $\frac{\partial \mathscr{L}(\cdot)}{\partial x_2} = 0$, which indicate that $U'_{x_1} = \lambda p_1$ and $U'_{x_2} = \lambda p_2$. Dividing both equations, the optimal demands satisfy the following condition:

$$\frac{U'_{x_1}}{U'_{x_2}} = \frac{p_1}{p_2} \tag{3.1}$$

or, equivalently, $-\frac{U'_{x_1}}{U'_{x_2}} = -\frac{p_1}{p_2}$. Thus, when equation (3.1) is evaluated at the optimal basket, the slope of the indifference curve equals the slope of the budget line. In Figure

3.2, we graphically demonstrate the point at which the optimal solution occurs.

**Figure 3.2.** *Optimal Consumption*




In Lecture Note 1, we introduced one of the most common utility functions used in the literature: the Cobb Douglas. Next, we use this case to illustrate the UMP.

Example

Remember that we can apply a monotone transformation to an utility function, and the new function would still represent the same preferences. In Lecture Note 1, we have shown that a Cobb Douglas utility function can be presented in its log form by

$$U(x_1, x_2) := x_1^{\alpha_1} x_2^{\alpha_2}.$$

where $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$. The optimization problem is then

$$\max_{x_1,x_2} U(x_1, x_2) = x_1^{\alpha_1} x_2^{\alpha_2} \text{ subject to } Y = p_1 x_1 + p_2 x_2.$$

In Lecture Note 1, we have also shown that the Cobb Douglas function satisfies differentiability, strong monotonicity, and strictly quasiconcavity. To show that it is a well-behaved utility function, it rests to show that it also satisfies the Inada conditions for each good:

$$\lim_{x_i \to 0} U'_{x_i} = \infty \text{ for } i = 1, 2.$$

Consider good 1. Then, $\frac{\partial U(x_1,x_2)}{\partial x_1} := \alpha_1 (x_1)^{\alpha_1 - 1}$. Thus, if $x_1 \to 0$, the denominator goes to zero, and hence the whole expression goes to infinite. Thus, the Inada condition for good 1 is satisfied. The result for good 2 is identical.

When the utility function is well behaved, the solution is unique (by strict quasiconcavity) and interior (by Inada). So, we do not need to check the SOC. Moreover, the Marshallian demands can be obtained by use of the Lagrange techniques, giving

$$x_1^* (p_1, Y) = \alpha_1 \frac{Y}{p_1} \text{ and } x_2^* (p_2, Y) = \alpha_2 \frac{Y}{p_2}.$$

To find the optimal solution, we know that the solution is unique and interior. Moreover, the utility function is differentiable and the consumer is constrained by her income. Thus, we can make use of the Lagrange technique. The optimization problem can be solved in two steps. First, we construct the Lagrangian. Second, we take the Lagrangian as the objective function and proceed to its optimization. This is done as if the optimization problem were an unconstrained problem, but with the Lagrange multiplier as one more variable to optimize. Let's do it.
We can optimize a function after applying a monotone transformation, since it would provide the same result. So let's use the log form of the Cobb Douglas. The Lagrangian is given by:
$\mathscr{L} := \alpha_1 \ln (x_1) + \alpha_2 \ln (x_2) + \lambda [Y - p_1 x_1 - p_2 x_2]$.
Now, we proceed to optimize the function by choosing $x_1$, $x_2$ and $\lambda$ as control variables. The FOCs are:
$\mathscr{L}'_{x_1} = \frac{\alpha_1}{x_1} - \lambda p_1 = 0,$
$\mathscr{L}'_{x_2} = \frac{\alpha_2}{x_2} - \lambda p_2 = 0$, and
$\mathscr{L}'_{\lambda} = Y - p_1 x_1 - p_2 x_2 = 0.$
An easy way to solve the system of equations and obtain the Marshallian demands is the following. First consider the equations $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$, and obtain a relation between $x_1$ and $x_2$. Then, plug the relation into the budget constraint (that is, the equation $\mathscr{L}'_{\lambda} = 0$) to obtain the Marshallian demands.
Let's see how this works. First, notice that the equations $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$ can be expressed as $\frac{\alpha_1}{x_1} = \lambda p_1$ and $\frac{\alpha_2}{x_2} = \lambda p_2$, respectively. Dividing both equations, we obtain an expression for $x_2$ as a function of $x_1$:
$\frac{\frac{\alpha_1}{x_1}}{\frac{\alpha_2}{x_2}} = \frac{\lambda p_1}{\lambda p_2} \Rightarrow x_2 = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1.$
Plugging this expression into $\mathscr{L}'_{\lambda} = 0$:
$Y - p_1 x_1 - p_2 x_2 = 0 \Rightarrow Y - p_1 x_1 - p_2 \left( \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1 \right) = 0$
$\Rightarrow Y - p_1 x_1 - \left( \frac{\alpha_2}{\alpha_1} \right) p_1 x_1 = 0 \Rightarrow x_1^* (p_1, Y) = \alpha_1 \frac{Y}{p_1}$
Once we identified the optimal demand of good 1, we can obtain $x_2^*$ by using that $x_2 = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1$, so that
$x_2^* = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1^* \Rightarrow x_2 = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} \left( \alpha_1 \frac{Y}{p_1} \right)$
$\Rightarrow x_2^* (p_2, Y) = \alpha_2 \frac{Y}{p_2}.$

With this solution, we can obtain the indirect utility:

$$U^* (p_1, p_2, Y) = \frac{Y}{\mathbb{P}}, \tag{3.2}$$

where $\mathbb{P} := \left( \frac{p_1}{\alpha_1} \right)^{\alpha_1} \left( \frac{p_2}{\alpha_2} \right)^{\alpha_2}$.

By definition of the indirect utility function, $U^*(p_1, p_2, Y) := U\left[x_1^*(p_1, p_2, Y), x_2^*(p_1, p_2, Y)\right]$.

For the case of the Cobb Douglas this is,

$U^*(p_1, p_2, Y) = \left[x_1^*(p_1, Y)\right]^{\alpha_1}\left[x_2^*(p_2, Y)\right]^{\alpha_2} \Rightarrow U^*(p_1, p_2, Y) = \left(\alpha_1 \frac{Y}{p_1}\right)^{\alpha_1}\left(\alpha_2 \frac{Y}{p_2}\right)^{\alpha_2}$

We reorder the terms in brackets, so that:

$U^*(p_1, p_2, Y) = \left(\frac{\alpha_1}{p_1} Y\right)^{\alpha_1}\left(\frac{\alpha_2}{p_2} Y\right)^{\alpha_2}$

$\Rightarrow U^*(p_1, p_2, Y) = \left(\frac{\alpha_1}{p_1}\right)^{\alpha_1}\left(\frac{\alpha_2}{p_2}\right)^{\alpha_2} Y^{\alpha_1+\alpha_2}$

and using that $\alpha_1 + \alpha_2 = 1$ and that $\left(\frac{\alpha_i}{p_i}\right)^{\alpha_i} = (\alpha_i)^{\alpha_i}(p_i)^{-\alpha_i} = \frac{1}{\left(\frac{p_i}{\alpha_i}\right)^{\alpha_i}}$, then

$U^*(p_1, p_2, Y) = \frac{Y}{\mathbb{P}}$, where $\mathbb{P} := \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$.

### 3.2.3 Some Results Using the Envelope Theorem

We will apply the Envelope Theorem to derive some results regarding the UMP. Remember that the Envelope Theorem is a result that concerns how variations in a parameter affect the value function. In the UMP, this means how variations in either income or one of the prices affect the indirect utility function. Furthermore, recall that the procedure to apply the Envelope Theorem is as follows.

**Procedure to apply the Envelope Theorem**

Step 1. Construct the Lagrangian: $\mathscr{L}(\cdot) := U(x_1, x_2) + \lambda\left[Y - p_1 x_1 - p_2 x_2\right]$

Step 2. Take derivatives of the Lagrangian with respect to the parameter of interest, without embedding the optimal solutions:

- $\frac{\partial \mathscr{L}(x_1, x_2, \lambda; Y, p_1, p_2)}{\partial Y} = \lambda$
- $\frac{\partial \mathscr{L}(x_1, x_2, \lambda; Y, p_1, p_2)}{\partial p_1} = -\lambda x_1$

Step 3. Evaluate the derivatives at the optimal values to obtain $\frac{\partial U^*(Y, p_1, p_2)}{\partial Y}$ and $\frac{\partial U^*(Y, p_1, p_2)}{\partial p_1}$:

- $\frac{\partial U^*(Y, p_1, p_2)}{\partial Y} = \lambda^*(Y, p_1, p_2)$
- $\frac{\partial U^*(Y, p_1, p_2)}{\partial p_1} = -\lambda^*(Y, p_1, p_2)\, x_1^*(Y, p_1, p_2)$

By analyzing how the parameters affect the indirect utility function, we can derive two results.

- **Interpretation of the Lagrange multiplier**: $\frac{\partial U^*(Y,p_1,p_2)}{\partial Y} = \lambda^*(Y,p_1,p_2) > 0$

In consumer theory, the Lagrange multiplier indicates **the marginal utility of income**. This means that $\lambda^*$ indicates the impact of one unit of income on the agent's optimal utility. It can be shown that $\lambda^* > 0$, so that $\lambda^*$ provides the intuitive result that the higher the income of the agent, the better off she is.

> **Remark** *Notice that the Lagrange multiplier $\lambda^*$ is affected by monotone transformations. This can be clearly seen since $\lambda^*$ is obtained by $\frac{\partial U^*(Y,p_1,p_2)}{\partial Y}$. Consequently, depending on which utility function we take as primitive, both $U^*$ and $\lambda^*$ vary. Of course, the Marshallian demands are still the same, and hence independent of monotone transformations.*

- **Roy's Identity**: $\frac{\partial U^*}{\partial p_i} = -\lambda^*(p_1,p_2,Y)\, x_i^*(p_1,p_2,Y) < 0$ for any good $i$.

Since $\lambda^* > 0$ and $x_i^* > 0$, increases in prices determine a lower maximum utility.[4] Roy's identity is not so important on its own, but combined with the result for $\lambda^*$ indicates that we can determine the Marshallian demand of any good $i$ through knowledge of the indirect utility function. Specifically,

$$x_i^*(p_1,p_2,Y) = -\frac{\partial U^*(Y,p_1,p_2)/\partial p_i}{\partial U^*(Y,p_1,p_2)/Y}.$$

The importance of the result is due to two reasons. First, it simplifies some calculations, by establishing a direct link with the expenditure minimization problem we study below. Second, it will become relevant to compute welfare measures, as we will show in subsequent lecture notes.

Example

Suppose an indirect utility function given by

$$U^*(p_1,p_2,Y) := \frac{Y}{\sqrt{p_1 p_2}}$$

First, we can obtain the optimal Lagrange multiplier by using that equals $\frac{\partial U^*}{\partial Y}$, so

[4] Since we are assuming that utility functino is well-behaved, then the quantities demanded cannot be zero. Otherwise, Roy's identify would imply that $x_i^* \geq 0$ and we should allow for the possibility that $\frac{\partial U^*}{\partial p_i} = 0$.

that $\lambda^*(p_1, p_2) := .\frac{1}{\sqrt{p_1 p_2}}$.

Then, we can obtain the Marshallian demands. For example, consider good 1. Then, since $\frac{\partial U^*}{\partial p_1} = -\frac{Y}{2}\sqrt{\frac{1}{p_2}}(p_1)^{-\frac{3}{2}}$ and $\frac{\partial U^*}{\partial Y} = \frac{1}{\sqrt{p_1 p_2}}$, then $x_1^*(p_1, p_2, Y) = \frac{Y}{2}\sqrt{\frac{1}{p_2}}(p_1)^{-\frac{3}{2}}\left(\frac{1}{\sqrt{p_1 p_2}}\right)^{-1}$, which gives $x_1^*(p_1, p_2, Y) = \frac{Y}{2p_1}$ after some simplifications.

Notice that this Marshallian demand is the same that we would obtain from a Cobb Douglas when $\alpha_1 := 0.5 =: \alpha_2$. This arises because the indirect utility function is the same as equation (3.2) with those parameter values.[5]

### 3.2.4 Comparative Statics

So far, we have characterized the optimal solution of the UMP and provided some relations through the use of the Envelope Theorem. Although those matters were interesting on their own, we usually build a model to study a phenomenon. For this reason, we want to establish predictions and/or refutable hypotheses that we can ultimately contrast with the data.

One of the main tools to derive results of this type is through performing comparative statics exercises. Given a model, as a characterization of how agents make decisions, comparative statics entails the following: first shocking a parameter, and then obtaining predictions regarding how the endogenous variables are affected by this. In terms of consumer theory, this means that we either vary prices or income, and then analyze how the consumption decisions are affected.

Originally, consumer theory was conceived to justify the ***uncompensated law of demand***: a negative relation between the price and consumption of a good. Nonetheless, it turned out that all the comparative statics of the UMP have an ambiguous sign. Thus, variations in prices or income can increase or decrease the consumption of goods.

[5] The only difference is a constant $\frac{1}{2}$, but this does not affect the result.

**Result 3.1** *The comparative statics of the UMP provide the following results:*

- $\frac{\partial x_i^*(p_1,p_2,Y)}{\partial p_j} \gtreqless 0$ *for* $i, j = 1, 2$
- $\frac{\partial x_i^*(p_1,p_2,Y)}{\partial Y} \gtreqless 0$ *for* $i = 1, 2$.

**Remark** *I will not formally derive the results, since I do not want to dwell on comparative statics' techniques when we analyze consumer theory. I prefer to focus on other aspects, relegating this technique to other topics that we will see later. If you are interested in the formal derivation of these results, I have some written notes that I can send to you.*

### 3.2.5 A Glance at What We Will Do Later in the Course

We have said that changes in the own price of the good, in the prices of other goods or income can have a positive, negative, or no effect on the Marshallian demands. This is not very reassuring. However, it does not mean that the model is useless. We can still use it to analyze why the results are ambiguous, and what channels have to dominate to get a definite result.

I have not included the comparative static analysis, mainly because it is not really insightful concerning *why* results are ambiguous. However, we will study an alternative approach, which provides some answers on what the signs depend on. This is the so-called Slutsky equation, which decomposes a change in prices into two decisions by the consumer.

## 3.3 Expenditure Minimization Problem (EMP)

So far, we have studied the optimization problem of a consumer when she is endowed with some income and faces some prices. Now, we are going to study a problem intimately related: the EMP. The UMP is usually called the **primal** problem, while the EMP is referred to as the **dual** problem. We first state the EMP, and then explain its relation with the UMP.

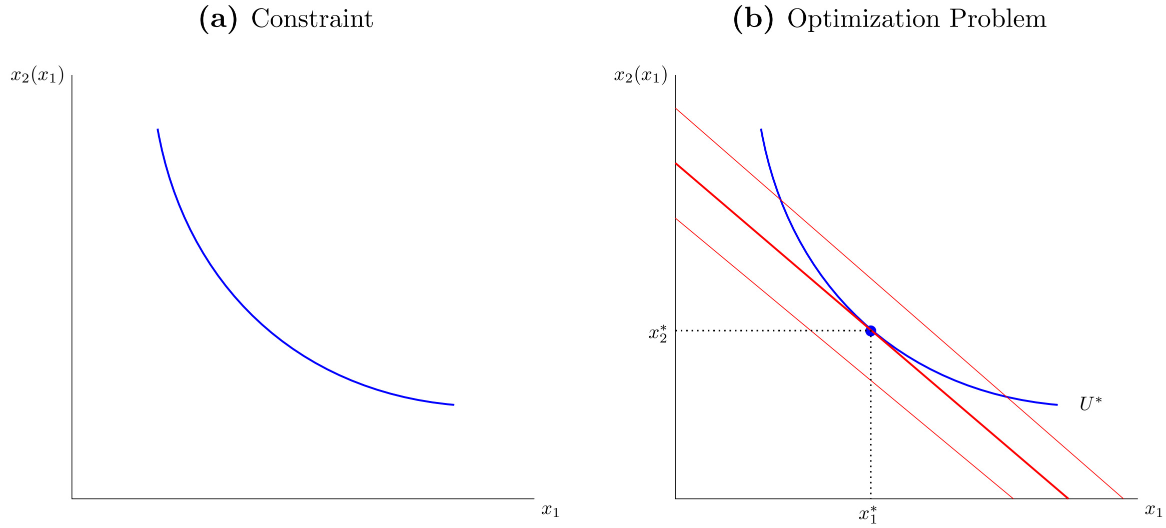
In the EMP, the consumer does not maximize utility. Rather, she has to decide the consumption of each good that achieves a level utility $U_0$, when prices are $p_1$ and $p_2$. Formally, let $E$ be the expenditure allocated. Then, the EMP is

$$\min_{x_1,x_2 \in X_1 \times X_2} E = p_1 x_1 + p_2 x_2 \text{ subject to } U_0 = U(x_1, x_2).$$

We can understand the EMP as the "inverse" analysis of the UMP. To see this, recall that one of the intuitions behind the UMP, outlined through Figure 3.1. Given a budget line, the UMP consists in drawing the map of indifference curves until we find the one that provides the greatest utility.

On the contrary, in the EMP, we fix an indifference curve and find the budget line that minimizes the income to achieve a level of utility. In Figure 3.3, we illustrate this. Given an indifference curve (Figure **??**), then we draw several budget lines, until we find the one that determines the lowest expenditure and provides utility $U_0$ (Figure **??**).

**Figure 3.3.** *Expenditure Minimization Problem*



### 3.3.1 General Solution

In case the utility is well-behaved, it can be shown that the EMP has a unique and interior solution. Moreover, this can be characterized by the Lagrange technique. The Lagrangian is given by

$$\mathscr{L}(x_1, x_2, \mu; p_1, p_2, U_0) := p_1 x_1 + p_2 x_2 + \mu \left[U_0 - U(x_1, x_2)\right]$$

I have denoted the Lagrange multiplier by $\mu$ to emphasize that this is not the same multiplier $\lambda$ from the UMP.

The FOCs are given by:

$$\text{FOCs:} \begin{cases} \frac{\partial \mathscr{L}(\cdot)}{\partial x_1} = p_1 - \mu \frac{\partial U(x_1,x_2)}{\partial x_1} = 0, \\ \frac{\partial \mathscr{L}(\cdot)}{\partial x_2} = p_2 - \mu \frac{\partial U(x_1,x_2)}{\partial x_2} = 0, \\ \frac{\partial \mathscr{L}(\cdot)}{\partial \mu} = U_0 - U(x_1, x_2) = 0. \end{cases}$$

Through these FOCs, we obtain solutions $h_1^*(p_1, p_2, U_0)$, $h_2^*(p_1, p_2, U_0)$, and $\mu^*(p_1, p_2, U_0)$. The optimal demands of the expenditure minimization problem, $h_1^*$ and $h_2^*$, are referred to as **Hicksian demands**.

Evaluating the expenditure at the optimal solutions, we also get the **minimum expenditure function**:

$$\begin{aligned} E^*(p_1, p_2, U_0) :=& E\left[h_1^*(p_1, p_2, U_0), h_2^*(p_1, p_2, U_0)\right], \\ =& p_1 h_1^*(p_1, p_2, U_0) + p_2 h_2^*(p_1, p_2, U_0) \end{aligned}$$

### 3.3.2 Interpreting the Optimality Conditions

The optimality condition of the EMP can be characterized through the FOCs. Remarkably, the EMP has the same tangent condition as the UMP. By using $\frac{\partial \mathscr{L}(\cdot)}{\partial x_1} = 0$ and $\frac{\partial \mathscr{L}(\cdot)}{\partial x_2} = 0$, we obtain $U'_{x_1} = \mu p_1$ and $U'_{x_2} = \mu p_2$. Dividing both expressions:

$$\frac{U'_{x_1}}{U'_{x_2}} = \frac{p_1}{p_2}.$$

Notice this does not imply that the Marshallian and Hicksian demands are equal. While they both have the same tangent condition, they differ in the constraints they have, which are given by $\frac{\partial \mathscr{L}(\cdot)}{\partial \mu}$ in the UMP and $\frac{\partial \mathscr{L}(\cdot)}{\partial \mu} = 0$ in the EMP. Nonetheless, it reveals that there is some relation between both. We explore this relation in the next section.

Example

We keep using the Cobb Douglas to illustrate how to solve the EMP. The optimization problem is

$$\min_{x_1,x_2} E = p_1x_1 + p_2x_2 \text{ subject to } U_0 = x_1^{\alpha_1}x_2^{\alpha_2},$$

and the Hicksian demands for the case $\alpha_1 + \alpha_2 = 1$ are

$$h_1^*(p_1, p_2, U_0) = U_0 \left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2} \text{ and } h_2^*(p_1, p_2, U_0) = U_0 \left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}.$$

The Lagrangian is:

$\mathscr{L} := p_1x_1 + p_2x_2 + \mu\left[U_0 - x_1^{\alpha_1}x_2^{\alpha_2}\right]$

and the FOCs are:

$\mathscr{L}'_{x_1} = \alpha_1 x_1^{\alpha_1-1}x_2^{\alpha_2} - \mu p_1 = 0$

$\mathscr{L}'_{x_2} = x_1^{\alpha_1}\alpha_2 x_2^{\alpha_2-1} - \mu p_2 = 0$

$\mathscr{L}'_{\mu} = U_0 - x_1^{\alpha_1}x_2^{\alpha_2} = 0$

Just like in the case of UMP, an easy way to solve the system of equations and obtain the optimal Hicksian demands is to first consider $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$ and obtain a relation between $x_1$ and $x_2$. Then we plug the relation into the constraint (that is, the equation $\mathscr{L}'_{\mu} = 0$) to obtain the solution.

From the two equations we obtain the same tangent condition as in the UMP, so that $\frac{\alpha_1}{\alpha_2}\frac{x_2}{x_1} = \frac{p_1}{p_2}$.

From this, we obtain an expression for $x_2$ as a function of $x_1$: $x_2 = \frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}x_1$.

Plugging in this expression into $\mathscr{L}'_{\mu} = 0$:

$U_0 - x_1^{\alpha_1}x_2^{\alpha_2} = 0 \Rightarrow U_0 - x_1^{\alpha_1}\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}x_1\right)^{\alpha_2} = 0$

$\Rightarrow U_0 - x_1^{\alpha_1+\alpha_2}\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_2} = 0$ which using that $\alpha_1 + \alpha_2 = 1$, then

$h_1^*(p_1, p_2, U_0) = U_0\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2}$.

Likewise, we use that $x_2 = \frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}x_1$ and so that $h_2^*(p_1, p_2, U_0) = \left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)h_1^*(p_1, p_2, U_0)$. This determines that $h_2^*(p_1, p_2, U_0) = U_0\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}$.

Likewise, the minimum expenditure is

$$E^*(p_1, p_2, U_0) = U_0\left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} = U_0\mathbb{P}$$

By definition, $E^*(p_1,p_2,U_0) = p_1 h_1^*(p_1,p_2,U_0) + p_2 h_2^*(p_1,p_2,U_0)$. Hence,
$E^*(p_1,p_2,U_0) = p_1 U_0 \left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2} + p_2 U_0 \left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}$
$\Rightarrow E^*(p_1,p_2,U_0) = U_0 \left[p_1 \left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2} + p_2 \left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}\right]$
$\Rightarrow E^*(p_1,p_2,U_0) = U_0 \left[(p_1)^{1-\alpha_2} \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} (p_2)^{\alpha_2} + (p_2)^{1-\alpha_1} \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1} (p_1)^{\alpha_1}\right]$
By using that $\alpha_1+\alpha_2=1$, then $\alpha_2 = 1-\alpha_1$ and $\alpha_1 = 1-\alpha_2$. Therefore,
$E^*(p_1,p_2,U_0) = U_0 \left[(p_1)^{\alpha_1} \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} (p_2)^{\alpha_2} + (p_2)^{\alpha_2} \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1} (p_1)^{\alpha_1}\right]$
$\Rightarrow E^*(p_1,p_2,U_0) = U_0 (p_1)^{\alpha_1} (p_2)^{\alpha_2} \left[\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1}\right]$
Finally, using that $\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1} = \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{1-\alpha_2}$ we can reexpress the RHS
$\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} \Rightarrow \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2}\left(1+\frac{\alpha_2}{\alpha_1}\right) \Rightarrow \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2}\left(\frac{\alpha_1+\alpha_2}{\alpha_1}\right) \Rightarrow (\alpha_1)^{\alpha_2-1}(\alpha_2)^{-\alpha_2}$ which is just $\left(\frac{1}{\alpha_1}\right)^{\alpha_1}\left(\frac{1}{\alpha_2}\right)^{\alpha_2}$.
Thus, $E^*(p_1,p_2,U_0) = U_0 (p_1)^{\alpha_1} (p_2)^{\alpha_2} \left[\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1}\right]$ becomes
$E^*(p_1,p_2,U_0) = U_0 (p_1)^{\alpha_1} (p_2)^{\alpha_2} \left(\frac{1}{\alpha_1}\right)^{\alpha_1}\left(\frac{1}{\alpha_2}\right)^{\alpha_2}$ which gives the result.

### 3.3.3 Some Results Using the Envelope Theorem

Just like we did with the indirect utility function, we can obtain some relations by applying the Envelope Theorem to the minimum expenditure function.

- **Interpretation of the Lagrange multiplier**: $\frac{\partial E^*(p_1,p_2,U_0)}{\partial U_0} = \left.\frac{\partial \mathscr{L}}{\partial p_i}\right]_{h_1^*,h_2^*} = \mu^*(p_1,p_2,U_0)$

As we might suspect, $\mu^* > 0$. This indicates that achieving a higher level of utility requires increasing the minimum expenditure.

- **Shephard's Lemma**: $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_i} = \left.\frac{\partial \mathscr{L}}{\partial p_i}\right]_{h_1^*,h_2^*} = h_i^*(p_1,p_2,U_0)$

Notice that this is similar to Roy's identity, but with a subtle and important difference: we can recover the Hicksian demand without the need to know the Lagrange multiplier.

### 3.3.4 Comparative Statics

Unlike the case of the UMP, comparative statics in the EMP have definite signs. I summarize the results and relegate their explanations for when we analyze the Slutsky equation in Lecture Note 4. The fact that we obtain definite signs makes it clear that the Marshallian and Hicksian demands are different functions—keeping different parameters fixed ($Y$ vs $U_0$) affects the results.

**Result 3.2** *In the EMP, the comparative statics exercises provide the following results:*

- $\frac{\partial h_i^*(p_1,p_2,U_0)}{\partial p_i} < 0$ *for* $i = 1, 2$.
- $\frac{\partial h_j^*(p_1,p_2,U_0)}{\partial p_i} > 0$ *for* $i = 1, 2$ *and* $j \neq i$.
- $\frac{\partial h_i^*(p_1,p_2,U_0)}{\partial U_0} \gtreqless 0$ *for any* $i = 1, 2$.

**Remark** *The result* $\frac{\partial h_j^*(p_1,p_2,Y)}{\partial p_i} > 0$ *for* $i = 1, 2$ *and* $j \neq i$ *is only true for the case of two goods, while* $\frac{\partial h_j^*(p_1,p_2,Y)}{\partial p_i} \gtreqless 0$ *with more than two goods. Nonetheless, this will not matter to us. The most important result that you should keep in mind is that an increase of a good's own price reduces the Hicksian demand.*

## 3.4 Duality: Relations between UMP and EMP

What is the relation between both optimization problems? Suppose we start from a situation where a consumer makes choices when parameters are $(p_1, p_2, Y)$. This determines a Marshallian $x_i^*(p_1, p_2, Y)$ for good $i = 1, 2$ and an indirect utility function $U^*(p_1, p_2, Y)$. Suppose now, that we solve the EMP *for the particular case* in which $U_0 = U^*(p_1, p_2, Y)$. By the graphs we have shown, we could suspect that $E(p_1, p_2, U_0) = Y$, given that the UMP and the EMP have the same tangent condition. This is indeed the case. The intuition is that if a consumer is maximizing utility, she will use the income she has efficiently. Thus, the income $Y$ is the minimum that a rational consumer needs to achieve utility $U^*(p_1, p_2, Y)$.

Although the intuition is easy to grasp, let's show this formally. One way to prove a statement of the type "If $A$ then $B$" is by a technique called contradiction. This consists in assuming that $A$ happens but $B$ does not occur, and then show that this leads us to a contradiction. If that is the case, then we conclude that if $A$ happens, then $B$ necessarily happens too.

To keep the notation simple, let's refer to $U^*(p_1, p_2, Y)$ and $E^*(p_1, p_2, U^*)$ by $U^*$ and $E^*$ respectively Towards a contradiction, let's assume that $U_0 = U^*$ but $E^* \neq Y$. The fact that $E^* \neq Y$ requires analyzing two cases: $E^* < Y$ and $E^* > Y$.

If $E^* < Y$ then a consumer maximizing could increase her utility by spending $E^*$ to achieve $U^*$, and then spending the rest of income given by $Y - E^*$ on some goods. Thus $U^*$ would not maximize utility when the income is $Y$, which was how we defined $U^*$ in first place. Hence, a contradiction.

If $E^* > Y$, then a consumer minimizing expenditure to achieve $U^*$ could have chosen the same bundle composed of the Marshallian demands and would have spent $Y$. Thus, $E^*$ does not minimize expenditure when the utility to achieve is $U^*$, which contradicts the definition of $E^*$.

All this means that if $U_0 = U^*$ then $E^* = Y$.

We have said that if $U_0 = U^*(p_1, p_2, Y)$ then $E(p_1, p_2, U_0) = Y$. Once we know this proposition holds, we can establish a relation between Marshallian and Hicksian demands: $x_i^*(p_1, p_2, Y) = h_i^*[p_1, p_2, U^*(p_1, p_2, Y)]$ for $i = 1, 2$.

In fact, the relation also holds the other way round: if $E(p_1, p_2, U_0) = Y$, then $U_0 = U^*(p_1, p_2, Y)$. Hence, it is also true that $x_i[p_1, p_2, E(p_1, p_2, U_0)] = h_i^*(p_1, p_2, U_0)$. We formalize this in the following.

**Result 3.3** *Relation between the UMP and EMP:*

- *if* $U_0 = U^*(p_1, p_2, Y)$ *then*
  - $E^*(p_1, p_2, U_0) = Y$
  - $x_i^*(p_1, p_2, Y) = h_i^*[p_1, p_2, U^*(p_1, p_2, Y)]$ *for* $i = 1, 2.$
- *if* $E^*(p_1, p_2, U_0) = Y$ *then*
  - $U_0 = U^*(p_1, p_2, Y)$
  - $x_i^*[p_1, p_2, E(p_1, p_2, U_0)] = h_i^*(p_1, p_2, U_0).$

A key implication of this result is that, once we solve one of the optimization problems (UMP or EMP), we do not need to solve the other one. Specifically, we will say that we *use duality* when we establish the solution of one of the problems through the result of the other problem (see the example below).

**Remark** *The Marshallian and Hicksian demands have the same value at the specific point where income is defined by the minimum expenditure, and at the specific points where the level of utility is defined as the indirect utility. However, these demands do not coincide for the rest of the values. It is only under some stringent conditions under which both demands are the same at all points. We explore this in the next subsection.*

Example

Next, we illustrate how to recover the solution of the EMP when we only have information about the indirect utility function of the UMP. Suppose that the indirect utility function is given by

$$U^*(p_1, p_2, Y) := \frac{Y}{\sqrt{p_1 p_2}}.$$

To obtain the minimum expenditure function, we use that that if $U_0 = U^*(p_1, p_2, Y)$ then $E^*(p_1, p_2, U_0) = Y$. Therefore, substituting in the indirect utility function, we obtain that $U_0 = \frac{E^*(p_1,p_2,U_0)}{\sqrt{p_1 p_2}}$, which determines that

$$E^*(p_1, p_2, U_0) = U_0\sqrt{p_1 p_2}.$$

Thus, we have recovered the minimum expenditure.

Using the indirect utility function, we can also apply Roy's identity and recover the Marshallian demands. For good 1, this is is given by $x_1^*(p_1, p_2, Y) = \frac{Y}{2p_1}$, and we can determine the Hicksian demand by using duality. We know that $x_1^*[p_1, p_2, E^*(p_1, p_2, U_0)] = h_1^*(p_1, p_2, U_0)$, and so $x_1^*(p_1, p_2, Y) = \frac{E^*(p_1,p_2,U_0)}{2p_1}$. Replacing for the minimum expenditure we have obtained, this gives $h_1^*(p_1, p_2, U_0) = \frac{U_0\sqrt{p_1 p_2}}{p_1}$, or simply $h_1^*(p_1, p_2, U_0) = U_0\sqrt{\frac{p_2}{p_1}}$.

Notice that, alternatively, once we have recovered the minimum expenditure, we could have determined the Hicksian demands by using Shephard's Lemma.

### 3.4.1 Conditions for Equivalence between Marshallian and Hicksian Demands (OPTIONAL)

Even though in general $h_1^* \neq x_1^*$, under some specific conditions both demands are equal. The conditions are quite stringent, but nonetheless they apply to one pervasive case: the quasilinear utility function, which we explore later in the course.

I begin by presenting one result that is of some importance itself. It establishes that Marshallian demands do not depend on income iff the Hicksian demands do not depend on the utility level.

**Result 3.4** *$h_1^*$ does not depend on $U_0$ iff $x_1^*$ does not depend on $Y$.*

We start from the relation between the primal and dual problem given by $E[p_1, p_2, U^*(p_1, p_2, Y)] = Y$. Taking derivatives wrt $Y$,

$$\frac{\partial E^*}{\partial U_0}\frac{\partial U^*}{\partial Y} = 1 \tag{3.3}$$

This implies that $\mu^*[p_1, p_2, U^*(p_1, p_2, Y)]\,\lambda^*(p_1, p_2, Y) = 1$.
Taking derivative of equation (3.3) wrt $p_1$,

$$\frac{\partial^2 E^*}{\partial U_0 \partial p_1}\frac{\partial U^*}{\partial Y} + \frac{\partial E^*}{\partial U_0}\frac{\partial^2 U^*}{\partial Y \partial p_1} = 0 \tag{3.4}$$

First, keep in mind that $\frac{\partial E^*}{\partial p_1} = h_1^*(p_1, p_2, U_0)$ and so

$$\frac{\partial^2 E^*}{\partial p_1 \partial U_0} = \frac{\partial h_1^*(p_1, p_2, U_0)}{\partial U_0}$$

Moreover, $x_1^*(p_1, p_2, Y) = -\frac{\partial U^*(Y, p_1, p_2)/\partial p_1}{\partial U^*(Y, p_1, p_2)/Y}$ and so $\frac{\partial x_1^*(p_1, p_2, Y)}{\partial Y} = -\frac{\frac{\partial^2 U^*}{\partial p_1 \partial Y}\frac{\partial U^*}{\partial Y} - \frac{\partial^2 U^*}{\partial p_1 \partial Y}\frac{\partial U^*}{\partial p_1}}{[\partial U^*/Y]^2}$ which determines that

$$\frac{\partial x_1^*(p_1, p_2, Y)}{\partial Y} = -\frac{\partial^2 U^*}{\partial p_1 \partial Y}\frac{\frac{\partial U^*}{\partial Y} - \frac{\partial U^*}{\partial p_1}}{[\partial U^*/Y]^2}$$

Now we have all the elements to provide a proof. Since we have a proposition that is and "if and only if", we need to prove two statements. First, that if $h_1^*$ does not depend on $U_0$ then $x_1^*$ does not depend on $Y$. Second, if $x_1^*$ does not depend on $Y$ then $h_1^*$ does not depend on $U_0$.
Regarding the first statement, This follows because if $h_1^*$ does not depend on $U_0$ then $\frac{\partial^2 E^*}{\partial p_1 \partial U_0} = \frac{\partial h_1^*(p_1, p_2, U_0)}{\partial U_0} = 0$ which implies by equation (3.4) that

$$\frac{\partial E^*}{\partial U_0}\frac{\partial^2 U^*}{\partial Y \partial p_1} = 0$$

since $\frac{\partial E^*}{\partial U_0} = \mu^* \neq 0$, then $\frac{\partial^2 U^*}{\partial Y \partial p_1} = 0$ and so $\frac{\partial x_1^*(p_1, p_2, Y)}{\partial Y} = 0$.
Concerning the second statement, if $x_1^*$ does not depend on $Y$ then, since $\frac{\partial U^*}{\partial Y} - \frac{\partial U^*}{\partial p_1} > 0$, then $\frac{\partial^2 U^*}{\partial p_1 \partial Y} = 0$. Therefore, by using equation (3.4),

$$\frac{\partial^2 E^*}{\partial U_0 \partial p_1}\frac{\partial U^*}{\partial Y} = 0$$

Since $\frac{\partial U^*}{\partial Y} = \lambda^* \neq 0$ then $\frac{\partial^2 E^*}{\partial U_0 \partial p_1} = 0$ which implies that $\frac{\partial h_1^*(p_1, p_2, U_0)}{\partial U_0} = 0$.

Notice that the result also establishes that Marshallian demands depend on income iff the Hicksian demands depend on the utility level. By using this result, we can establish the conditions under which both demands are identical.

**Result 3.5** *If $h_1^*$ does not depend on $U_0$ or $x_1^*$ does not depend on $Y$, then $h_1^* = x_1^*$.*

This statement follows almost trivially by the relation between primal and dual. We know that $x_1^*[p_1, p_2, E^*(p_1, p_2, U_0)] = h_1^*(p_1, p_2, U_0)$ (remember that both are equal only at that point). But if $x_1^*$ does not depend on $Y$ so that $h_1^*$ does not depend on $U_0$, then $x_1^*(p_1, p_2) = h_1^*(p_1, p_2)$ for all $(p_1, p_2)$.

As we indicated, the main application of this result will be for the quasilinear utility function. We will show in the next lecture note that when income is high enough, the Marshallian demand of one of the goods does not exhibit income effects. Consequently, we do not have to solve the EMP if we want to obtain the Hicksian demand of good 1—it is going to be identical to the Marshallian demand.

## 3.5 Exercises

[1] John is a student at University of Alberta. He spends half of his income on books (good 1) and half on food (good 2). He does not consume any other good since he studies all time, leaving him with no time for any other activity.

(a) A person $A$ suggests that, given the information at our disposal, the utility function of John could be represented by $U(x_1, x_2) := \sqrt{x_1 x_2}$. Explain why someone has arrived at this conclusion.

(b) A person $B$ says that $A$ is wrong and argues that John's utility function should actually be represented by $U(x_1, x_2) = 2\ln x_1 + 2\ln x_2$. What would you say about it?

(c) From now on, suppose that the utility function of John is the one given in a). Solve John's maximization problem and find his Marshallian demands.

(d) John's mother is worried that his son is studying so much that he's neglecting his diet. For this reason, she decides to give him some money. Specifically, she'll give him enough money to increase his income in 1%. The mother expects that John will completely devote that 1% of additional income to the consumption of food. Can you predict if this will end up happening?

[2] Let's continue with John, but now as a representative student in Canada. As in the previous exercise, he has a utility function $U(x_1, x_2) := \sqrt{x_1 x_2}$, and now we set specific values for the parameters: $Y := 2000$ CAD, and prices $p_1 = 2$ and $p_2 = 1$. Alberta's authorities are planning to build new roads. To finance the project, they are studying the possibility of levying a tax on textbooks, which are exempt from taxes. Given Canada's number of students, they estimate that the project could be financed by collecting 200 CAD per student. Consider the following situations.

(a) Suppose Alberta sets a tax of $t$ CADs per book sold. This means that every time a bookstore sells a book, the consumer has to pay $p_1$ and the tax $t$.

i. Write the budget constraint incorporating the tax.

ii. What has to be the value of $t$ so that Alberta can collect 200 CAD per student?

(b) Answer the questions in a), but now assuming that the tax is a $\tau\%$ of each book sold's value. This means that every time a book is sold, the consumer has to pay the price and a $\tau\%$ more over the price (this is known as an *ad valorem* or value-added tax, similar to the GST tax that we pay in Edmonton)

(c) Answer the question in a), but now assuming that the tax consists of a $w\%$ of a student's income.

(d) By construction, the three tax schemes provide the same tax revenue. Someone from Alberta's government asks you which scheme you'd choose. What would you recommend? (Hint: the key here is the criteria you'd use for the answer. Think equal tax revenue does not mean that the consumer is indifferent between the tax schemes).

[3] This exercise is just for you to practice how to use the Envelope Theorem and duality. It'll become important in the next problem sets.

Suppose that the indirect utility function of Mona is given by $U^*(p_1, p_2, Y) := \frac{Y}{p_1+p_2}$. Determine:

(a) Mona's Marshallian demands

(b) Mona's minimum expenditure function

(c) Mona's Hicksian demands

**Some answer keys**:

2a) $\tau = 0.5$ CAD, 3b) $\tau = 25\%$ c) $\tau = 10\%$

3a) $x_1^*(p_1, p_2, Y) = \frac{Y}{p_1+p_2}$, 1b) $E^*(p_1, p_2, U_0) = U_0(p_1 + p_2)$, 1c) $h_1^*(p_1, p_2, U_0) = U_0$

‘

# Lecture Note 4

# The Slutsky Equation

## 4.1 Introduction

When we studied the UMP, our main conclusion comparative statics do not have unambiguous results. In particular, the uncompensated law of demand, which states a negative relation between the consumption of a good and its own price, does not necessarily hold. The Slutsky equation provides some insight regarding why this is case, by decomposing the effects of a price change into channels. Additionally, it will helps us understand why the uncompensated law of demand does not hold, and under what conditions it is satisfied.

Specifically, a variation of price triggers two changes: one on relative prices and another on real income. **The substitution effect** corresponds to the change in relative prices while **the income effect** to the change in real income. There are two ways to define the substitution effect, depending on the experiment used to isolate the change in relative prices.

On the one hand, we have the original derivation of Slutsky. Starting from the UMP, this requires compensating the consumer with enough income, such that the optimal bundle consumed before the variation in price is still affordable at the new prices. On the other hand, we have the derivation by Hicks, which is based on the EMP. It consists in compensating the consumer with enough income, such that the initial level of utility is achievable at the new prices.

When the variation in price is infinitesimal, both definitions of substitution effect provide the same result. But, for an arbitrary change in price (i.e. not necessarily small) they may differ. We will respectively refer to each way to identify the substitution effect as the **Slutsky** and **Hicks compensations**.

In this lecture note, we proceed as follows. First, we derive the Slutsky equation by using duality. This derivation implicitly assumes a compensation à la Hicks. Then, we provide some interpretations of the substitution effect by using the Slutsky and Hicks compensations. In particular, we illustrate how each effect can be calculated through the Cobb-Douglas case.

## 4.2 Derivation of the Slutsky Equation

We provide a formal derivation of the Slutsky equation based on duality. Given a level of utility $U_0 := U^*(p_1, p_2, Y)$, consider good 1 and a variation in its own price. Then, the Slutsky equation is defined by:

$$\underbrace{\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1}}_{\text{total price effect}} = \underbrace{\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1}}_{\text{substitution effect } (<0)} \underbrace{- x_1^*(p_1,p_2,Y)\frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}}_{\text{income effect } \left(\gtreqless 0\right)}.$$

By duality, we know that when $U_0 = U^*(p_1, p_2, Y)$, then $E(p_1, p_2, U_0) = Y$. In turn, this implies that for the good 1, $h_1^*(p_1, p_2, U_0) = x_1^*[p_1, p_2, E^*(p_1, p_2, U_0)]$. We use this relation between the Marshallian and Hicksian demands to derive the Slutsky equation.

From the relation $h_1^*(p_1, p_2, U_0) = x_1^*[p_1, p_2, E(p_1, p_2, U_0)]$, we take the derivative of each side with respect to $p_1$: $\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1} = \frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1} + \frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}\frac{\partial E(p_1,p_2,U_0)}{\partial p_1}$.

By Shepard's Lemma, we know that $\frac{\partial E(p_1,p_2,U_0)}{\partial p_1} = h_1^*(p_1,p_2,U_0)$, and so $\frac{\partial E(p_1,p_2,U_0)}{\partial p_1} = x_1^*(p_1,p_2,Y)$. This results in the expression $\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1} = \frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1} + \frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}x_1^*(p_1,p_2,Y)$. Reordering the terms, we arrive to the Slutsky equation: $\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1} = \frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1} - x_1^*(p_1,p_2,Y)\frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}$.

Actually, the Slutsky equation is more general, and even applies to variations in the price of the other good:

$$\underbrace{\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_2}}_{\text{total price effect}} = \underbrace{\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_2}}_{\text{substitution effect } (<0)} \underbrace{- x_2^*(p_1,p_2,Y)\frac{\partial x_2^*(p_1,p_2,Y)}{\partial Y}}_{\text{income effect } \left(\gtreqless 0\right)}.$$

### 4.2.1 Intuition behind the Slutsky Equation

Suppose an initial situation with prices $p_1'$ and $p_2'$, and consider a new situation where the price of the good 1 changes. Specifically, let prices in the new situation be $p_1''$ and $p_2''$ where $p_1'' > p'_1$ and $p_2' = p_2''$. We denote the equilibrium values with superscripts ' and " depending on the situation under consideration. Thus, $U'$ and $U''$ are the indirect utility functions obtained in each case, and $x_1'$ and $x_1''$ are the Marshallian demands of good 1, and $x_2'$ and $x_2''$ the Marshallian demands of good 2.

The Slutsky equation analyzes the total effect of a variation $\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1}$, noticing that a change in price triggers changes in relative prices (substitution effect) and real income (income effect).

To provide some intuition, let's rewrite the budget constraint. Starting from $Y = p_1x_1 + p_2x_2$ and dividing both sides by $p_1$, we obtain that $\frac{Y}{p_1} = x_1 + \frac{p_2}{p_1}x_2$. Expressing the budget constraint in this way, the optimization problem of the consumer has two parameters $\left(\frac{Y}{p_1}, \frac{p_2}{p_1}\right)$, instead of three parameters $(Y, p_1, p_2)$.

The parameter $\frac{Y}{p_1}$ is the real income expressed in terms of good 1, i.e., the total quantity that income $Y$ can buy of good 1. Moreover, $\frac{p_2}{p_1}$ represents the relative price. It indicates how many units of good 1 the consumer could buy when she does gives up one unit of good 2 and spends that additional income (i.e., $p_2$) on good 1.

When there is an increase in $p_1$, there are two effects working simultaneously. On the one hand, the relative prices $\frac{p_2}{p_1}$ change. Thus, if the consumer stops buying one unit of good 1, now she can consume more of good 2, in comparison to what was happening before the increase in $p_1$. This is the **substitution effect**—when there is an increase in $p_1$, consuming more of good 2 becomes a more attractive option than before.

The other effect is given by the change in real income. An increase in $p_1$ makes the consumer poorer, since she can afford less amount of goods. Thus, the **income effect** captures how the consumption of good 1 varies when her real income is lower.

It is important to distinguish between the meaning of "real income" and "real income in terms of good 1." We have provided an intuition about these terms by just rewriting the budget constraint in terms of the former. But real income is the purchasing power that the income has, and so it is expressed in terms of *both* good 1 and good 2. This requires defining what is called a price index, which reflects the price of one basket of goods that comprises both goods. In contrast, the real income in terms of good 1 is given by $\frac{Y}{p_1}$, which is the real income exclusively in terms of how much good 1 income can buy. However, both are related: if the real income in terms of good 1 changes, then it is necessarily true that real income varies.

Why is this distinction important? Because agents make consumption decisions according to real income. Thus, the income effect incorporates the fact that after an increase in $p_1$, for a given quantity of good 1 consumed, the consumer can afford fewer units of good 2. This is because the consumer has a lower disposable income, after she has spent money on the now more expensive good 1. This is the reason why a variation in $\frac{Y}{p_1}$ also affects the quantity that the consumer can afford of good 2, even when $p_2$ has not changed.

## 4.3 Hicks' Experiment

While we have provided an interpretation of what each effect comprises, we have not said why specifically the substitution and income effects are given by the terms we stated. Next, we do so.

To analyze how a consumer varies the quantities of good 1 when only relative prices change, we need to isolate the change in real income. Put it differently, the substitution effect is computed ceteris paribus the change in real income. To accomplish this, we can conceive two experiments.

The first way is the one envisioned by Hicks. This was in fact the basis for our derivation of the Slutsky equation. The intuition is the following. An increase in $p_1$ reduces the consumer's real income, and so the consumer can afford less of both goods. This implies that she would end up with a lower utility. To isolate the pure change in relative prices, Hicks conceived the experiment of compensating the consumer with enough income to achieve the same level of utility she was having before the change in prices. Thus, the **substitution effect à la Hicks** is given by the adjustment in the consumption of good 1 in order to achieve the same level utility previous to the change in relative prices. This explains why the substitution effect is captured by the term $\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1}$.

The income effect is obtained just as a residual: the difference between the total change and the variation due to the substitution effect. But, why is this given specifically by $x_1^*(p_1, p_2, Y) \frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}$?

To understand this, remember the notion of a derivative. $\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1}$ entails a small variation in $p_1$ (i.e. infinitesimal), with its result expressed in terms of a unit variation of $p_1$. Put it different, we are expressing an infinitesimal variation as if the variation were $\mathrm{d}p_1 = 1$. How does this impact real income? Since there is an increase in one unit of $p_1$, it is like if the agent had seen reduced his consumption in an amount $x_1^*$. Specifically, $\mathrm{d}Y = x_1^* \,\mathrm{d}p_1$ with $\mathrm{d}p_1 = 1$ provides the decrease in income the consumer faces. Thus, we can provide an interpretation of each term in the following way:

$$\underbrace{-x_1^*(p_1, p_2, Y)}_{=\mathrm{d}Y} \underbrace{\frac{\partial x_1^*(p_1, p_2, Y)}{\partial Y}}_{\text{effect per unit of income}} .$$

## 4.4 Slutsky's Experiment

The other way to derive the Slutsky equation is the one originally conceived by Slutsky. Although we have not derived the Slutsky equation using that method, the Hicks and Slutsky methods arrive at the same result when the variation in $p_1$ is infinitesimal.

Slutsky isolates the substitution effect by compensating the consumer in a different way than Hicks. Rather than giving the consumer enough money to stay at the same level of utility, the monetary compensation is such that the consumer can afford the original bundle. To see this more clearly, consider original prices $(p_1', p_2')$ with an optimal consumption $(x_1', x_2')$. Suppose now that the consumer faces prices $(p_1'', p_2')$ with $p_1'' > p_1'$. Slutsky compensates the agent such that the original basket $(x_1', x_2')$ is affordable at the new prices $(p_1'', p_2')$.[1]

Thus, the **substitution effect à la Slusky** is given by the adjustment in consumption of good 1 when the consumer is compensated with enough money to afford the original bundle. Just like with Hicks' method, the income effect can be obtained as a

[1] Let's denote the compensation as $\Delta Y$ and the total income once the agent is compensated by $Y^C$. To determine what each term is equal to, we know that $p_1' x_1' + p_2' x_2' = Y$. So now we need to give the consumer $Y^C := p_1'' x_1' + p_2' x_2'$, determining that $\Delta Y := x_1' (p_1'' - p_1')$.

residual: the difference between the total change in consumption and the variation due to the substitution effect.

I have emphasized that the Slutsky's and Hicks' compensations provide the same value of substitution effect for infinitesimal variations in $p_1$. On the contrary, substitution effects depend on the method used to compute them if we do not consider small variations in $p_1$. Why is this the case? Try to think about it in the following way. Suppose good 1 represents hardcover books and good 2 digital books. Usually, physical books are more expensive than digital ones.
Suppose that the price of hardcover books increase and the university offers you two deals. First, the university could give you enough money to buy the same amount of physical and digital books that you were buying before the change in prices. The second deal is that, whatever you end up buying in terms of physical and digital books, the university gives you enough money to get the same utility as you were having before the increase in $p_1$. What would you choose?
With the second alternative, you won't be either worse off or better off, irrespective of your choice. By definition, you will stay in the same indifference curve. On the other hand, the first alternative allows for some scope to exploit the type of compensation. For instance, rather than buying the same original bundle that you were consuming before the increase in prices, you could reduce the quantity consumed of hardcover books and buy these books in a digital form. In this way, you would end up saving money, which you could spend on other goods and hence increase your utility. In other terms, when you are compensated à la Slutsky, you might adjust your consumption's choices to reach an indifference curve that provides more utility.

## 4.5 Giffen Goods

The main goal of the Slutsky equation is to study the effect on consumption following variations in prices, by decomposing the impact into channels. Based on it, next we provide conditions under which the sign of the total price effect can be determined. Remember that the Slutsky equation for good 1 when there is a change in its own price is

$$\underbrace{\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1}}_{\text{total price effect}} = \underbrace{\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1}}_{\text{substitution effect } (<0)} - \underbrace{x_1^*(p_1,p_2,Y)\frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}}_{\text{income effect } \left(\gtreqless 0\right)}.$$

Thus, we might attain different results depending on the sign of the income effect.

Before studying this expression, let's state some definitions. First, we distinguish between two concepts related to total effects on prices.

- Good $i$ satisfies the **uncompensated law of demand** (henceforth, ULD) if $\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1} < 0$.

- Good $i$ is a **Giffen good** when $\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1} > 0$.

As far as income effects go, goods can be classified into three categories.

- Good $i$ is a **normal good** when the income effect is positive, $\frac{\partial x_i^*(p_1,p_2,Y)}{\partial Y} > 0$.
- Good $i$ is an **inferior good** when the income effect is negative, $\frac{\partial x_i^*(p_1,p_2,Y)}{\partial Y} < 0$.
- Good $i$ has **a zero income effect** when $\frac{\partial x_i^*(p_1,p_2,Y)}{\partial Y} = 0$.

Let's start by inquiring upon conditions under which the total price effect will be negative, so that the ULD holds. Without loss of generality, we state the results in terms of good 1.

**Result 4.1** *If good 1 is normal or has a zero income effect, then the ULD for good 1 holds.*

It follows by the fact that

$$\underbrace{\frac{\partial x_1^*(p_1,p_2,Y)}{\partial p_1}}_{\text{total price effect}} = \underbrace{\frac{\partial h_1^*(p_1,p_2,U_0)}{\partial p_1}}_{(-)} - \underbrace{x_1^*(p_1,p_2,Y)\frac{\partial x_1^*(p_1,p_2,Y)}{\partial Y}}_{\text{normal good }(+)} < 0.$$

As a corollary, if the ULD does not hold for good 1, then it is necessarily true that good 1 is neither a normal or a zero income effect good. Put it differently, if a good is Giffen, then it is necessarily inferior.

In real life, observing goods that are Giffen is rare. What actually matters about Giffen goods is that variations in prices encompass an income effect. This means that increases in prices reduce consumption when the ULD holds, but the decrease in consumption is even more so than what the substitution effect dictates— higher prices not only affect relative prices, but also reduce a consumer's real income.

From this result we can also derive an important lesson: when you work with a model, it is important that you identify the channels operating in the model. This allows you to know what type of story you are representing through the model used. Thus, to identify a Giffen good, it is **not** enough to find a good for which there is a positive relation between the consumption and its own price. You also need that the explanation for such a relation is the one dictated by the model you are using. In consumer theory, it means that the income effect has to be driving the result.

For instance, an example of a Giffen good is not given by Apple increasing the price of its new iPhone and observing that its sales increase. This is due to two reasons. First, we have not included quality as a concept that influences demand. Presumably, Apple sells more when it launches a new cell phone because, among other things, it improves its features. Thus, the effect is not explained by an income effect, but by the overhauls introduced—this aspect is not covered by the baseline model we have presented.

Second, keep in mind that the total price effect is a partial derivative. Hence, we need to analyze a change in the price ceteris paribus, such that any other aspect of the framework remains unchanged. If Apple provides a product with greater quality (or even if people find the iPhone more appealing because it is trendy), then we would be observing two effects taking place simultaneously: the effect from an increase in price and the effect from a change in quality.

Given the remark provided, an example of a Giffen good requires not only finding a negative relation between a good's own price and its consumption. We need to think of a situation where the price of the good changes ceteris paribus, with the fall in real income explaining the variation in consumption.

One example of a Giffen good could be the following. Suppose a poor person living with 2 CADs per day, who spends all her income on rice and meat. Assume also that the price of one unit of rice is 0.50 CAD and the price of meat is 1.50 CAD, and that she consumes 1 unit of rice and 1 of meat. Suppose now that the price of rice increases to 1 CAD. In this situation, her income is not enough to afford both goods at the same time. Hence, if we observe that she ends up consuming two units of rice and no meat at all, rice behaves as a Giffen Good at that level of prices and income.

To be more precise, let's think about the changes in terms of the Slutsky equation, assuming that rice is a good inferior for all levels of prices and income. We could expect that if she were compensated by the government for the increase in the price of rice, she would consume more meat and less rice. But the reduction of real income is substantial enough that she cannot afford both goods simultaneously. Thus, the income effect dominates: the reduction of income translates into an increase in the consumption of rice.

## 4.6 Example: Cobb Douglas

Suppose a consumer has a Cobb Douglas utility function with parameters $\alpha_1 = \alpha_2 = 0.5$. Formally, this means that $U(x_1, x_2) := \sqrt{x_1 x_2}$. We consider two situations. In Situation A, the consumer has income $Y' := 2000$ and faces prices $p_1' := 4$ and $p_2' := 1$. In Situation B, $Y$ and $p_2$ do not vary, but the price of good 1 becomes $p_1'' := 25$. To keep notation simple, we denote the variables of each situation by using superscripts $'$ and $''$, respectively.

Since $p_2' = p_2'' = 1$ and $\alpha_1 = \alpha_2 = \frac{1}{2}$, the Marshallian demands are $x_1^*(p_1, Y) = \frac{1}{2}\frac{Y}{p_1}$ and $x_2^*(p_2, Y) = \frac{1}{2}Y$, and the indirect utility function is $U^*(p_1, p_2, Y) = \frac{Y}{2\sqrt{p_1}}$. We have expressed the results as functions of $Y$ and $p_1$, because we will eventually change their values throughout the example.

Using these results and substituting for the values $p_1$ and $Y$ in each situation, we obtain the following results:

| **Situation** | $(p_1, p_2, Y)$ | $(x_1^*, x_2^*)$ | $U^*$ |
|---|---|---|---|
| Original (Situation A) | $(4, 1, 2000)$ | $(250, 1000)$ | 500 |
| $\Delta p_1$ (Situation B) | $(25, 1, 2000)$ | $(40, 1000)$ | 200 |

The total variation in consumption is $(\Delta x_1, \Delta x_2) := (-210, 0)$. The goal is now to decompose the total change in consumption into substitution and income effects. Since there are two ways to calculate the substitution effect (Hicks vs Slutsky), next we consider each case separately.

### 4.6.1 Slutsky Compensation

Suppose we compensate the agent à la Slutsky. This means that the agent receives additional income at the new prices, such that the original bundle is affordable. More specifically, the agent needs to get enough income such that she can afford the bundle $(x_1', x_2') := (250, 1000)$ at the prices $(p_1'', p_2'') := (25, 1)$.

Originally, she was having an income $Y' = 2000$. To afford the basket of situation A under the prices of situation B, she needs a total income $Y^S$ equal to:

$$\begin{aligned} Y^S &:= p_1''x_1' + p_2''x_2', \\ &:= 25 \times 250 + 1 \times 1000 = 7250. \end{aligned}$$

Since she was already having an income of 2000, the compensation $\Delta Y^S$ has to equal $\Delta Y^S := 5250$.

From this, we determine the Marshallian demands in case the consumer is compensated. We denote each with a superscript $S$:

$$x_1^S := x_1^* (25, 1, 7250) = 145$$

$$x_2^S := x_2^* (25, 1, 7250) = 3625$$

With this income, she gets a utility $U^S := 725$. Given the remarks made before, this occurs because when the variation in prices is not infinitesimal and the agent is compensated à la Slutsky, she ends up with a greater utility relative to what she was getting before the compensation.

| **Situation** | $(p_1, p_2, Y)$ | $(x_1^*, x_2^*)$ | $U^*$ |
|---|---|---|---|
| Original | $(4, 1, 2000)$ | $(250, 1000)$ | 500 |
| $\Delta p_1$ | $(25, 1, 2000)$ | $(40, 1000)$ | 200 |
| Slutsky | $(25, 1, 7250)$ | $(145, 3625)$ | 725 |

From this, we conclude that, when there is a change in $p_1$ from $p_1' = 4$ to $p_2'' = 25$, the total reduction in the consumption of good 1 is 210 units $(40 - 250)$. Moreover, this can be broken down into a decrease in 105 units $(145 - 250)$ due to the Slutsky substitution effect, and 105 units due to the income effect $(40 - 145)$.

It is also interesting to split the total effect on the quantities of good 2. In that case, the variation in $p_1$ determines an increase in the consumption of good 2 due to the Slutsky substitution effect of 2625 units $(3625 - 1000)$, but a reduction in the consumption due to the income effect of 2625 units $(3625 - 1000)$. Overall, the quantity consumed of good

2 does not change, because each effect cancels out with each other.

| **Compensation** | $(x_1', x_2')$ | $(x_1'', x_2'')$ | $(x_1^C, x_2^C)$ | $(\Delta x_1, \Delta x_2)$ | **Subst. Effect** | **Income Effect** |
|---|---|---|---|---|---|---|
| Slutsky | $(250, 1000)$ | $(40, 1000)$ | $(145, 3625)$ | $(-210, 0)$ | $(-105, +2625)$ | $(-105, -2625)$ |

### 4.6.2 Hicks Compensation

Let's consider now a compensation à la Hicks. This requires obtaining the total income needed to achieve the utility level of Situation A, but with the prices of Situation B. We can do this by making use of the minimum expenditure function. For the Cobb Douglas, and taking $U_0$ and $p_1$ as parameters, this is given by $E^*(p_1, p_2, U_0) = 2U_0\sqrt{p_1}$. Hence, the total income $Y^H$ that she needs to achieve the utility $U' = 500$ at the prices $(p_1'', p_2'') := (25, 1)$ is:

$$
\begin{aligned}
Y^H &:= E^*(p_1'', p_2'', U') \\
&:= E^*(25, 1, 500) = 5000
\end{aligned}
$$

Since she was already having 2000 units of money, the Hicks compensation is then $\Delta Y^H := 3000$. We denote the Marshallian demands when the agent is Hicks compensated with a superscript $S$. Hence,

$$
\begin{aligned}
x_1^H &:= x_1^*(25, 1, 5000) = 100 \\
x_2^H &:= x_2^*(25, 1, 5000) = 2500
\end{aligned}
$$

Notice we could have obtained those demands in a different way. By using duality, we know that for any good $i$, $x_1^*[p_1, p_2, E^*(p_1, p_2, U_0)] = h_1^*(p_1, p_2, U_0)$. Hence, $x_i^H$ is equal to the value of the Hicksian demand $h_i^*(25, 1, 500)$.

The final table is then:

| **Situation** | $(p_1, p_2, Y)$ | $(x_1^*, x_2^*)$ | $U^*$ |
|---|---|---|---|
| Original | $(4, 1, 2000)$ | $(250, 1000)$ | 500 |
| $\Delta p_1$ | $(25, 1, 2000)$ | $(40, 1000)$ | 200 |
| Slutsky | $(25, 1, 7250)$ | $(145, 3625)$ | 725 |
| Hicks | $(25, 1, 5000)$ | $(100, 2500)$ | 500 |

Calculating the substitution and income effects as we did in the case of the Slutsky compensation, we end up with the following results:

| **Compensation** | $(x_1', x_2')$ | $(x_1'', x_2'')$ | $(x_1^C, x_2^C)$ | $(\Delta x_1, \Delta x_2)$ | **Subst. Effect** | **Income Effect** |
|---|---|---|---|---|---|---|
| Slutsky | $(250, 1000)$ | $(40, 1000)$ | $(145, 3625)$ | $(-210, 0)$ | $(-105, +2625)$ | $(-105, -2625)$ |
| Hicks | $(250, 1000)$ | $(40, 1000)$ | $(100, 2500)$ | $(-210, 0)$ | $(-150, +1500)$ | $(-60, -1500)$ |

### 4.6.3 Differences Between Compensations

Suppose now that the price of good 1 changes to $p_1'' = 5$.

| **Situation** | $(p_1, p_2, Y)$ | $(x_1^*, x_2^*)$ | $U$ |
|---|---|---|---|
| Original | $(4, 1, 2000)$ | $(250, 1000)$ | 500 |
| $\Delta p_1$ | $(5, 1, 2000)$ | $(200, 1000)$ | 447.2 |
| Slutsky | $(5, 1, 2250)$ | $(225, 1125)$ | 503.1 |
| Hicks | $(5, 1, 2236.1)$ | $(223.6, 1118)$ | 500 |

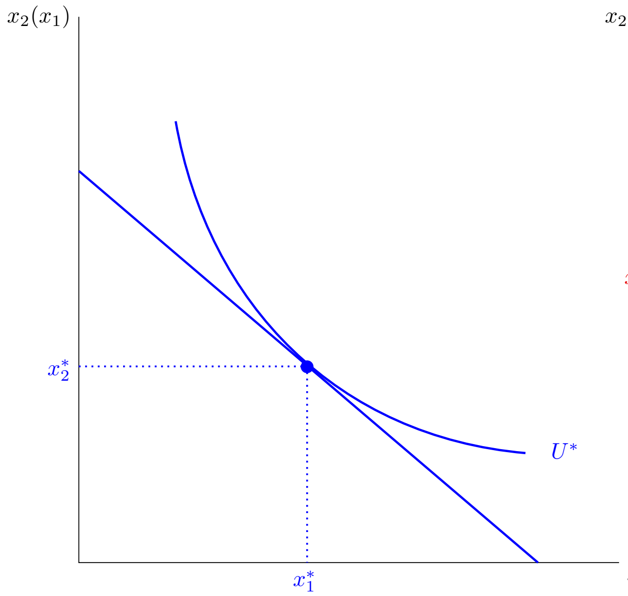
Calculating the substitution and income effects as we did in the case of the Slutsky compensation, we get the following results:

| **Compensation** | $(x_1', x_2')$ | $(x_1'', x_2'')$ | $(x_1^C, x_2^C)$ | $(\Delta x_1, \Delta x_2)$ | **Subst. Effect** | **Income Effect** |
|---|---|---|---|---|---|---|
| Slutsky | $(250, 1000)$ | $(200, 1000)$ | $(225, 1125)$ | $(-50, 0)$ | $(-25, +125)$ | $(-25, -125)$ |
| Hicks | $(250, 1000)$ | $(200, 1000)$ | $(223.6, 1118)$ | $(-50, 0)$ | $(-26.4, +1500)$ | $(-23.6, -1500)$ |

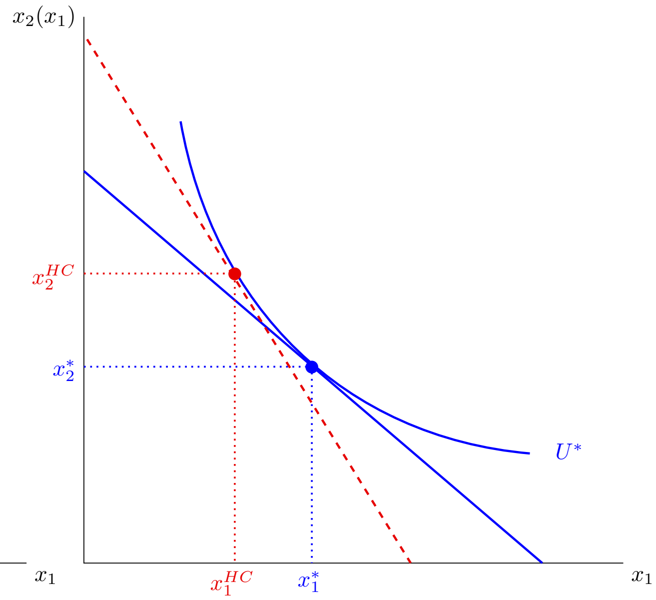
The differences in quantities are just 0.62% after an increase in price of 25%, so the ratio is 2.48. On the contrary, the differences in quantities before were 45% after an increase in price of 525%, with hence a ratio of 8.57.

**Figure 4.1.** *Slutsky Equation - Hicks Compensation with* $\uparrow p_1$

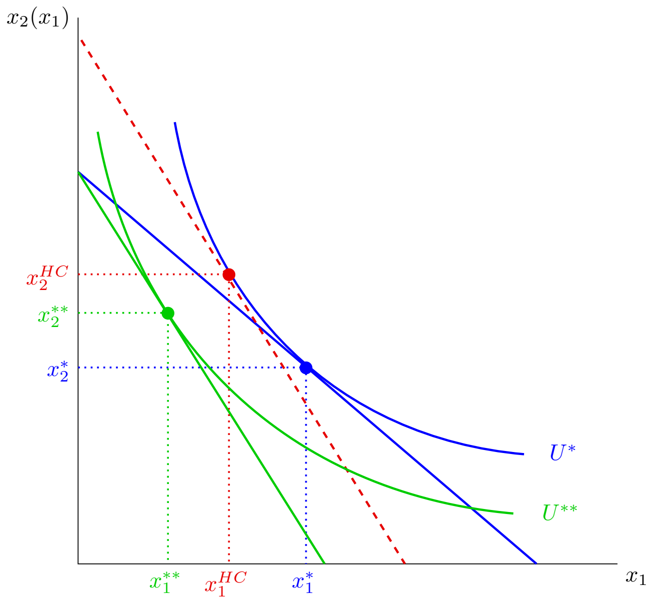
**(a)** Original Situation




**(b)** Hicks Compensation

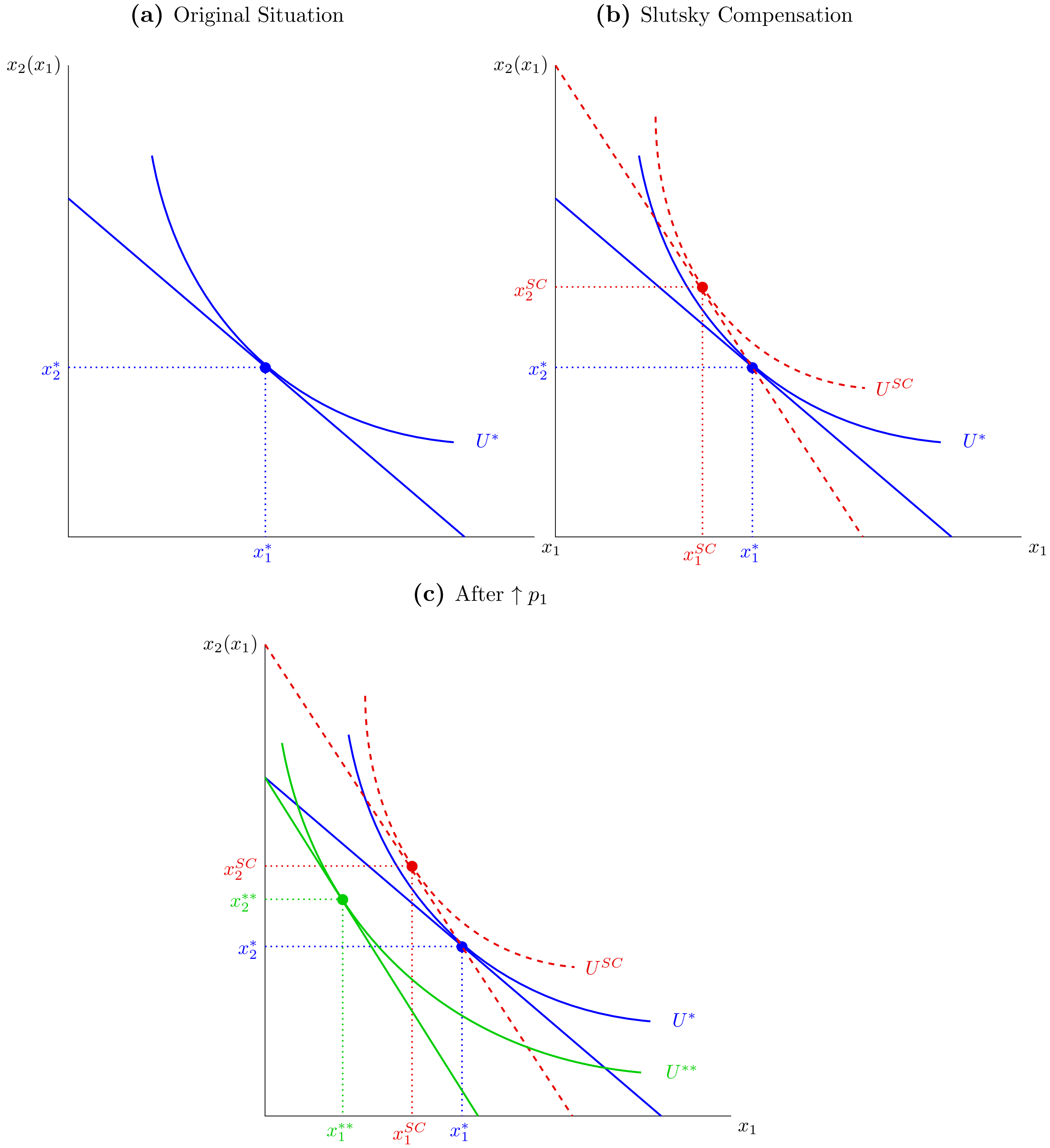



**(c)** After $\uparrow p_1$




**Note**: The Hicks compensation keeps the level of utility fixed at the level of original indifference curve, but taking the change in prices into account (reflected in the slope of the dashed line). Since the price of good 1 has increased, the substitution effect determines that less is consumed of good 1 and more of good 2 is consumed. Given how we have drawn the graph, we are implicitly assuming that the income effect is positive for both goods. This can be noticed since the changes in quantities from the Hicks compensation to the final consumption make the consumption of both goods decrease. That is, the decrease in real income determines that the consumer reduces the consumption of both goods.

**Figure 4.2.** *Slutsky Equation - Slutsky Compensation with* $\uparrow p_1$




**Note**: The Slutsky compensation allows the consumer to afford the original bundle $(x_1', x_2')$ at the new prices (reflected in the slope of the dashed line). Since the price of good 1 has increased, the substitution effect determines that less is consumed of good 1 and more of good 2. Notice that the graph is not considering an infinitesimal variation of the price, and so the consumer can actually be better off with a Slutsky compensation (i.e. can reach an indifference curve with a greater utility). Also, given how we have drawn the graph, we are implicitly assuming that the income effect is positive for both goods. This is reflected in through changes in quantities with the Slutsky compensation such that the final consumption entails a decrease in the consumption of both goods.

## 4.7 Exercises

[1] Laura usually takes his little son Julian to the park. When they go there, she usually gives him $Y' := 2$ CAD. With this money, Julian can go to the nearest convenience store and buy some lollipops (good 1) and candies (good 2). The price of each good is given by $(p_1', p_2') = (1, 1)$. However, the last time, the price of lollipops had increased to $p_1'' := 2$ CAD. Suppose Julian's utility function is $U(x_1, x_2) := \sqrt{x_1 x_2}$ (hint: since you'll have to calculate solutions for different values of parameters, I recommend you to solve the UMP and EMP using income and prices as parameters, and then use a Google SpreadSheet or Excel, to calculate the solution for specific values).

(a) Determine the Marshallian demands and indirect utility function of Julian in the original situation, where $(Y', p_1', p_2') := (2, 1, 1)$

(b) Determine the Marshallian demands and indirect utility function of Julian in the new situation, where $(Y'', p_1'', p_2'') := (2, 2, 1)$

(c) Julian became really anguished by the price increase. Due to this, Laura decided that she'll give him one additional CAD the next time they go to the park, so he can buy one unit of each good. This mean that $\left(Y^S, p_1'', p_2''\right) := (3, 2, 1)$. To her surprise, Julian does not end up consuming that amount of each good.

i. Calculate the new consumption levels and explain Julian's behavior.

ii. If Julian now could choose between the initial situation with $(Y', p_1', p_2') := (2, 1, 1)$, and the situation in which the mother gave him more money and so $\left(Y^S, p_1'', p_2''\right) := (3, 2, 1)$, what would he choose? Explain.

(d) Laura knows Julian's preferences (after all, he's her son). So, instead of giving him enough money to buy the initial bundle, she decides to make a different deal: she'll give Julian enough money to be as happy as in the original situation, considering that $(p_1'', p_2'') := (2, 1)$. How much money will she give him? What will Julian's demands will be?

[2] (I'll solve this one in class) Since 2011, the exchange rate between USD and CAD has been depreciating. Nowadays, 1 CAD equals 0.75 USD, rather than 1 USD as before. One consequence of a depreciation is that firms using imported inputs suffer an increase in their costs. This is because firms need to spend more CADs to buy the same amount of imported inputs.

Consider the following situation. Suppose that there are two types of firms in an industry of Canada, which only sell their products domestically. Type-1 firms (henceforth T1) import some of their inputs, while type-2 firms (T2) rely on local inputs. Since T1 have been facing an increase in their costs, they had to increase the price of their products. This has put them at a disadvantage relative to its competition (i.e. T2). For this reason, some of these firms have gone bankrupt.

The government is worried that the unemployment rate could increase due to the disappearance of T1 firms. For this reason, it has decided to intervene in the market. However, rather than subsidizing T1 firms, it decided to follow a different path: give additional income to Canada's consumers, so that they can afford the same bundle they were consuming at the new prices set by T1 (i.e. after their increase in price).

The policy was a complete failure. Can you imagine reasons why this happened?

**Answer Keys** for Some of the Exercises:

1a) $x_1' = x_2' = U' = 1$, 1b) $x_1'' = 0.5$, $x_2'' = 1$ and $U'' = 0.71$, 1ci) $x_1'' = 0.75$, $x_2'' = 1.5$ and $U'' = 1.07$. 1d) $Y = 2.83$ with $x_1 = 0.71$ and $x_2 = 1.41$

# Lecture Note 5

# Well-Behaved Utility Functions

## 5.1 Introduction

We begin the study of specific utility functions by considering well-behaved functional forms. This means that the solutions to the UMP and EMP are unique and interior, under certain conditions. The presentation will cover two of the most pervasive utility functions used in Economics: the Cobb Douglas and the quasilinear utility function.

## 5.2 Cobb Douglas

We have already used the Cobb Douglas utility function to illustrate some of the concepts in consumer theory. Here, we provide a complete treatment of this case.

### 5.2.1 UMP

Consider two goods, 1 and 2, with prices $p_1$ and $p_2$. Moreover, the consumer has income $Y$ and make choices on quantities consumed $(x_1, x_2) \in \mathbb{R}^2_{++}$. The utility is given by

$$U\left(x_1, x_2\right) := x_1^{\alpha_1} x_2^{\alpha_2},$$

where $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$. Remember that normalizing the sum of coefficients to one is without loss of generality, since it arises by applying a monotone transformation.

The optimization problem the consumer is

$$\max_{x_1, x_2} U\left(x_1, x_2\right) = x_1^{\alpha_1} x_2^{\alpha_2} \text{ subject to } Y = p_1 x_1 + p_2 x_2.$$

We have shown in a previous lecture that that Cobb Douglas function is strictly quasiconcave, satisfies strong monotonicity, and the Inada condition hold for each good. This ensures that the consumer's problem has a unique interior solution, which can be characterized through the Lagrange technique. It determines the following Marshallian demands:

$$x_1^*(p_1, Y) = \alpha_1 \frac{Y}{p_1},$$
$$x_2^*(p_2, Y) = \alpha_2 \frac{Y}{p_2}.$$

Remember that we can apply a monotone transformation and the new utility function would still represent the same preferences. So, for instance, a Cobb Douglas utility function can be presented in its log form:

$$U(x_1, x_2) := \alpha_1 \ln(x_1) + \alpha_2 \ln(x_2)$$

where $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$.

I find the optimal solutions by making use of the log representation.Then, the Lagrangian is given by:

$\mathscr{L} := \alpha_1 \ln(x_1) + \alpha_2 \ln(x_2) + \lambda [Y - p_1 x_1 - p_2 x_2]$

The FOCs are:

$\mathscr{L}'_{x_1} = \frac{\alpha_1}{x_1} - \lambda p_1 = 0$

$\mathscr{L}'_{x_2} = \frac{\alpha_2}{x_2} - \lambda p_2 = 0$

$\mathscr{L}'_{\lambda} = Y - p_1 x_1 - p_2 x_2 = 0$

An easy way to solve the equations and obtain the optimal Marshallian demands is to first consider $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$ and obtain a relation between $x_1$ and $x_2$. Then we plug the relation into the budget constraint (that is, the equation $\mathscr{L}'_{\lambda} = 0$) to obtain the solution.

Let's apply these steps. First, notice that the equations $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$ can be expressed as $\frac{\alpha_1}{x_1} = \lambda p_1$ and $\frac{\alpha_2}{x_2} = \lambda p_2$, respectively. Dividing both equations, we obtain an expression for $x_2$ as a function of $x_1$:

$\frac{\frac{\alpha_1}{x_1}}{\frac{\alpha_2}{x_2}} = \frac{\lambda p_1}{\lambda p_2} \Rightarrow x_2 = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1.$

Plugging in this expression into $\mathscr{L}'_{\lambda} = 0$:

$Y - p_1 x_1 - p_2 x_2 = 0 \Rightarrow Y - p_1 x_1 - p_2 \left( \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1 \right) = 0$

$\Rightarrow Y - p_1 x_1 - \left( \frac{\alpha_2}{\alpha_1} \right) p_1 x_1 = 0 \Rightarrow x_1^*(p_1, Y) = \alpha_1 \frac{Y}{p_1}$

Once we have the optimal Marshallian demand for good 1, we can obtain the optimal $x_2$ by using that $x_2 = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1$, so that

$x_2^* = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} x_1^* \Rightarrow x_2^* = \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} \left( \alpha_1 \frac{Y}{p_1} \right)$

$\Rightarrow x_2^*(p_2, Y) = \alpha_2 \frac{Y}{p_2}.$

The Marshallian demands of a Cobb Douglas exhibit a feature that makes it ideal for empirical analyses. To derive this property, take good 1. Its demand can be reexpressed as

$$\frac{p_1 x_1^*(p_1, Y)}{Y} = \alpha_1.$$

Notice that $p_1 x_1^*(p_1, Y)$ is the consumer's expenditure allocated to good 1, which becomes $\frac{p_i x_i^*(p_1, Y)}{Y}$ if we divide it by $Y$. The expression is the expenditure share of good 1 relative to total income, and equals $\alpha_1$ for the Cobb Douglas. Therefore, the expenditure share of good $i$ under a Cobb Douglas utility function is constant and equal to $\alpha_i$. The result requires that the sum of coefficients equals one, otherwise the expenditure share equals $\frac{\alpha_i}{\alpha_1+\alpha_2}$.

This property of the Cobb Douglas holds for any number of goods. To see this, suppose there is a set $\mathcal{I} := \{1, 2, ..., M\}$ of goods. The optimization problem then becomes

$$\max_{(x_i)_{i\in\mathcal{I}}} U = \sum_{i\in\mathcal{I}} \alpha_i \ln(x_i) \text{ subject to } Y = \sum_{i\in\mathcal{I}} p_i x_i,$$

where we assume that $\sum_{i\in\mathcal{I}} \alpha_i = 1$. Although this is optimization problem is more complicated, the solution is straightforward by using the property stated. Thus, the Marshallian demand of good $i$ is such that its expenditure share equals $\alpha_i$, so that $\frac{p_i x_i^*(p_i,Y)}{Y} = \alpha_i$. Therefore, the solution is still

$$x_i^*(p_i, Y) = \alpha_i \frac{Y}{p_i},$$

for any $i \in \mathcal{I}$.

#### 5.2.1.1 Comparative Statics

Once we assume a specific functional form, we can perform comparative statics in an easy way. This requires simply taking derivatives of the Marshallian demands with respect to the different parameters of the model. We show the results for good 1, since good 2 exhibits the same signs.

The Cobb Douglas determines that:

- $\frac{\partial x_1^*(p_1,Y)}{\partial p_1} = \frac{-\alpha_1 Y}{(p_1)^2} < 0$, and so the uncompensated law of demand holds,
- $\frac{\partial x_1^*(p_1,Y)}{\partial p_2} = 0$, and so there is no relation between goods in terms of prices,
- $\frac{\partial x_1^*(p_1,Y)}{\partial Y} = \frac{\alpha_1}{p_1} > 0$, and so the good is normal.

There is an alternative way that allows us to provide the same characterization, through elasticities. In the Math Review, you can find how elasticities can be easily obtained by applying logs. They determine the following:

- $\frac{\partial \ln x_1^*(p_1,Y)}{\partial \ln p_1} = -1$,
- $\frac{\partial \ln x_1^*(p_1,Y)}{\partial \ln Y} = 1$, and
- $\frac{\partial \ln x_1^*(p_1,Y)}{\partial \ln p_2} = 0$.

All these elasticities are independent of the parameters $\alpha_i$. Thus, the own-price elasticity and income eIasticity are always equal to one under a Cobb Douglas.

#### 5.2.1.2 Indirect Utility Function

Substituting in the Marshallian demands, the indirect utility is

$$U^*(p_1, p_2, Y) = \frac{Y}{\mathbb{P}},$$

where $\mathbb{P} := \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$.

$U^*(p_1, p_2, Y) := U\left[x_1^*(p_1, p_2, Y), x_2^*(p_1, p_2, Y)\right]$
$\Rightarrow U^*(p_1, p_2, Y) = \left[x_1^*(p_1, Y)\right]^{\alpha_1}\left[x_2^*(p_2, Y)\right]^{\alpha_2}$
$\Rightarrow U^*(p_1, p_2, Y) = \left(\alpha_1 \frac{Y}{p_1}\right)^{\alpha_1}\left(\alpha_2 \frac{Y}{p_2}\right)^{\alpha_2}$
we reorder the terms in brackets so that:
$U^*(p_1, p_2, Y) = \left(\frac{\alpha_1}{p_1} Y\right)^{\alpha_1}\left(\frac{\alpha_2}{p_2} Y\right)^{\alpha_2}$
$\Rightarrow U^*(p_1, p_2, Y) = \left(\frac{\alpha_1}{p_1}\right)^{\alpha_1}\left(\frac{\alpha_2}{p_2}\right)^{\alpha_2} Y^{\alpha_1+\alpha_2}$
and using that $\alpha_1 + \alpha_2 = 1$ and that $\left(\frac{\alpha_i}{p_i}\right)^{\alpha_i} = (\alpha_i)^{\alpha_i}(p_i)^{-\alpha_i} = \frac{1}{\left(\frac{p_i}{\alpha_i}\right)^{\alpha_i}}$ then
$U^*(p_1, p_2, Y) = \frac{Y}{\mathbb{P}}$ where $\mathbb{P} := \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$.

Furthermore, the Lagrange multiplier is given by $\frac{\partial U^*(p_1,p_2,Y)}{\partial Y}$ by the Envelope Theorem, and so

$$\lambda^*(p_1, p_2) = \frac{1}{\mathbb{P}}.$$

### 5.2.2 EMP

Next, we obtain the solution of the EMP. We first do it directly, by solving the EMP through the Lagrange technique. After this, we show that there is a more straightforward way to it, if we use the solution to the UMP and apply duality.

The EMP is

$$\min_{x_1,x_2} E = p_1 x_1 + p_2 x_2 \text{ subject to } U_0 = x_1^{\alpha_1} x_2^{\alpha_2}$$

where we keep assuming that $(x_1, x_2) \in \mathbb{R}^2_{++}$ and $\alpha_1, \alpha_2 > 0$ with $\alpha_1 + \alpha_2 = 1$.

The Hicksian demands are:

$$h_1^*(p_1, p_2, U_0) = U_0\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2} \text{ and } h_2^*(p_1, p_2, U_0) = U_0\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}$$

The Lagrangian is:
$\mathscr{L} := p_1 x_1 + p_2 x_2 + \mu \left[U_0 - x_1^{\alpha_1} x_2^{\alpha_2}\right]$
and the FOCs are:
$\mathscr{L}'_{x_1} = \alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} - \mu p_1 = 0$
$\mathscr{L}'_{x_2} = x_1^{\alpha_1} \alpha_2 x_2^{\alpha_2 - 1} - \mu p_2 = 0$
$\mathscr{L}'_{\mu} = U_0 - x_1^{\alpha_1} x_2^{\alpha_2} = 0$
Just like with the UMP, an easy way to solve this system of equations is to first consider $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$, to obtain a relation between $x_1$ and $x_2$. Then, we plug the relation into the constraint (that is, the equation $\mathscr{L}'_{\mu} = 0$) to obtain the solution.
The tangent condition is the same as in the UMP, so that $\frac{\alpha_1}{\alpha_2}\frac{x_2}{x_1} = \frac{p_1}{p_2}$.
From this, we obtain an expression for $x_2$ as a function of $x_1$: $x_2 = \frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}x_1$.
Plugging in this expression into $\mathscr{L}'_{\mu} = 0$:
$U_0 - x_1^{\alpha_1} x_2^{\alpha_2} = 0 \Rightarrow U_0 - x_1^{\alpha_1}\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}x_1\right)^{\alpha_2} = 0$
$\Rightarrow U_0 - x_1^{\alpha_1+\alpha_2}\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_2} = 0$, which using that $\alpha_1 + \alpha_2 = 1$, then
$h_1^*(p_1, p_2, U_0) = U_0\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2}$.
Likewise, if we use $x_2 = \frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}x_1$, then $h_2^*(p_1, p_2, U_0) = \left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right) h_1^*(p_1, p_2, U_0)$. This determines that $h_2^*(p_1, p_2, U_0) = U_0\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}$.

The minimum expenditure is

$$E^*(p_1, p_2, U_0) = U_0\left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} = U_0\mathbb{P}$$

By definition, $E^*(p_1, p_2, U_0) = p_1 h_1^*(p_1, p_2, U_0) + p_2 h_2^*(p_1, p_2, U_0)$. Hence,
$E^*(p_1, p_2, U_0) = p_1 U_0\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2} + p_2 U_0\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}$
$\Rightarrow E^*(p_1, p_2, U_0) = U_0\left[p_1\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2} + p_2\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}\right]$
$\Rightarrow E^*(p_1, p_2, U_0) = U_0\left[(p_1)^{1-\alpha_2}\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2}(p_2)^{\alpha_2} + (p_2)^{1-\alpha_1}\left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1}(p_1)^{\alpha_1}\right]$
By using that $\alpha_1 + \alpha_2 = 1$, then $\alpha_2 = 1 - \alpha_1$ and $\alpha_1 = 1 - \alpha_2$. Therefore,
$E^*(p_1, p_2, U_0) = U_0\left[(p_1)^{\alpha_1}\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2}(p_2)^{\alpha_2} + (p_2)^{\alpha_2}\left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1}(p_1)^{\alpha_1}\right]$
$\Rightarrow E^*(p_1, p_2, U_0) = U_0 (p_1)^{\alpha_1}(p_2)^{\alpha_2}\left[\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1}\right]$
Finally, using that $\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1} = \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{1-\alpha_2}$ we can reexpress the RHS $\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} \Rightarrow \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2}\left(1 + \frac{\alpha_2}{\alpha_1}\right) \Rightarrow \left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2}\left(\frac{\alpha_1+\alpha_2}{\alpha_1}\right) \Rightarrow (\alpha_1)^{\alpha_2 - 1}(\alpha_2)^{-\alpha_2}$ which is just $\left(\frac{1}{\alpha_1}\right)^{\alpha_1}\left(\frac{1}{\alpha_2}\right)^{\alpha_2}$.
Thus, $E^*(p_1, p_2, U_0) = U_0 (p_1)^{\alpha_1}(p_2)^{\alpha_2}\left[\left(\frac{\alpha_1}{\alpha_2}\right)^{\alpha_2} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\alpha_1}\right]$ becomes
$E^*(p_1, p_2, U_0) = U_0 (p_1)^{\alpha_1}(p_2)^{\alpha_2}\left(\frac{1}{\alpha_1}\right)^{\alpha_1}\left(\frac{1}{\alpha_2}\right)^{\alpha_2}$ which gives the result.

By using the Envelope Theorem, we can also obtain the the optimal Lagrange multiplier $\mu^*$, since $\frac{\partial E^*(p_1, p_2, U_0)}{\partial U_0} = \mu^*(p_1, p_2, U_0)$:

$$\mu^*(p_1, p_2, U_0) = \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} = \mathbb{P}$$

#### 5.2.2.1 Application of Duality

The Hicksian demands have been derived by solving the EMP. However, we could have also obtained the same result by using duality.

We know that the indirect utility function is $U^*(p_1,p_2,Y) = \frac{Y}{\mathbb{P}}$. Moreover, the relation between both optimization problems determines that if $U_0 = U^*(p_1,p_2,Y)$ then $Y = E^*(p_1,p_2,U_0)$. Applying this to the case of a Cobb Douglas determines that $U_0 = \frac{E^*(p_1,p_2,U_0)}{\mathbb{P}}$, which implies that

$$E^*(p_1,p_2,U_0) = U_0\mathbb{P}$$

The Hicksian demands can thus be obtained in two different ways. First, once we have recovered $E^*$, we can apply Shepard's Lemma to $E^*(p_1,p_2,U_0) = U_0\mathbb{P}$.

We have that $\mathbb{P} := \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$ and so $\frac{\partial \mathbb{P}}{\partial p_1} = (\alpha_1)^{1-\alpha_1}(p_1)^{\alpha_1-1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$. Shepard's Lemma establishes that $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_1} = h_1^*(p_1,p_2,U_0)$ and so $h_1^*(p_1,p_2,U_0) = U_0(\alpha_1)^{1-\alpha_1}(p_1)^{\alpha_1-1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$. By using that $\alpha_1+\alpha_2 = 1$, then $1-\alpha_1 = \alpha_2$ which implies that $h_1^*(p_1,p_2,U_0) = U_0\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2}$. By the same token, $h_2^*(p_1,p_2,U_0) = U_0\left(\frac{\alpha_2}{\alpha_1}\frac{p_1}{p_2}\right)^{\alpha_1}$.

Alternatively, we can start from the Marshallian demands and then use duality.

The Marshallian demand of good 1 is $x_1^*(p_1,Y) = \alpha_1\frac{Y}{p_1}$. Duality implies that $x_1^*[p_1,p_2,E(p_1,p_2,U_0)] = h_1^*(p_1,p_2,U_0)$. Hence, replacing in the MArshallian demand, $h_1^*(p_1,p_2,U_0) = \alpha_1\frac{E(p_1,p_2,U_0)}{p_1}$ and so $h_1^*(p_1,p_2,U_0) = \alpha_1\frac{U_0\mathbb{P}}{p_1}$. By using that $\mathbb{P} := \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$ then

$h_1^*(p_1,p_2,U_0) = U_0\alpha_1\frac{\left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}}{p_1}$

$\Rightarrow h_1^*(p_1,p_2,U_0) = U_0\frac{\left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}}{\frac{p_1}{\alpha_1}} \Rightarrow h_1^*(p_1,p_2,U_0) = U_0\left(\frac{p_1}{\alpha_1}\right)^{\alpha_1-1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$.

By using that $\alpha_1+\alpha_2 = 1$, then

$h_1^*(p_1,p_2,U_0) = U_0\left(\frac{p_1}{\alpha_1}\right)^{-\alpha_2}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} \Rightarrow h_1^*(p_1,p_2,U_0) = U_0\left(\frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}\right)^{\alpha_2}$.

Similar procedure for $h_2^*$.

## 5.3 Quasilinear Utility Function

The quasilinear is well-behaved when income is high enough to afford both goods. This is the usual assumption. However, we will see that there are corner solutions when income is low. Intuitively, the reason is that the quasilinear utility function satisfies Inada conditions for one of the goods, but not the other.

We say that $U$ is a quasilinear function when

$$U(x_1, x_2) := u(x_1) + x_2,$$

where $(x_1, x_2) \in \mathbb{R}^2_+$.

Good 2 enters linearly into the utility function and is usually referred to as the numéraire. This term is just a fancy way to say that $p_2 := 1$. The numéraire good is usually interpreted as a composite good representing "the rest of the goods", with the analysis focusing on good 1. Even when we will assume that $p_2 = 1$, we will keep track of $p_2$ in the solutions to show its role in the results.

With the goal of having a well-defined optimization problem, we suppose that $u$ is strictly increasing, strictly concave, and satisfies Inada conditions. Formally, $u$ is such that $\frac{\mathrm{d}u(x_1)}{\mathrm{d}x_1} > 0$ (strictly increasing) and $\frac{\mathrm{d}^2u(x_1)}{\mathrm{d}x_1^2} < 0$ (strictly concave), and Inada conditions refer to $\lim_{x_1 \to 0} \frac{\mathrm{d}u(x_1)}{\mathrm{d}x_1} = \infty$ and $\lim_{x_1 \to \infty} \frac{\mathrm{d}u(x_1)}{\mathrm{d}x_1} = 0$. Notice that Inada condition only apply to good 1, which rules out the corner solution with $x_1 = 0$ and all income spent on good 2. However, we do not assume Inada condition for good 2. In fact, this is not possible, since $\frac{\partial U(x_1,x_2)}{\partial x_2} = 1$ for all $x_1$ when we should actually have that this term is infinite for $x_2 \to 0$. Thus, it is possible to have a corner solution where only good 1 is consumed, and we show below that this occurs for low level of incomes.

To illustrate the results as clearly as possible, we assume a specific functional form for $u$ that satisfies all the conditions imposed. This is given by $u(x_1) := \ln(x_1)$.

### 5.3.1 Intuitions for The UMP

The UMP is

$$\max_{x_1,x_2} U(x_1, x_2) = \ln(x_1) + x_2$$

$$\text{subject to } Y = p_1 x_1 + p_2 x_2.$$

Strictly speaking, this optimization problem should be solved by using the Kuhn-Tucker procedure. In this way, we would cover the possibility that $x_2$ is zero. Instead of using this technique, we proceed more intuitively. We first show what the possible solutions are, and then under which conditions they arise.

Suppose that $x_1$ is glasses of water (a good that represents food), while $x_2$ is number of movies (a good representing entertainment). We want to know how the consumer allocates her expenditure on each good. To keep matters simple, suppose that $p_1 = p_2 = 1$, with income unspecified for now.

Under a quasilinear utility function, the UMP can be understood as a sequential allocation of each dollar. Given prices equal to one, each cent of dollar allows the consumer to buy exactly the same quantity of water or movies. The consumer will choose one of these options, according to which one provides more utility to her.

To identify the optimal decision between this two, we determine the marginal utility of consuming one unit of each good:

$$\frac{\partial U(x_1, x_2)}{\partial x_1} = \frac{1}{x_1}, \tag{5.1a}$$

$$\frac{\partial U(x_1, x_2)}{\partial x_2} = 1. \tag{5.1b}$$

The equations show that the marginal utility of good 1 depends on how much water she is already consuming. On the contrary, she always get one unit of utility per movie, irrespective of the number of movies already watched.

Suppose the scenario where she is not purchasing anything, and so has to decide how to spend her first cents of a dollar. Her choice can be identified by comparing (5.1a) and (5.1b) when the consumption of each good is zero. Formally,

$$\lim_{x_1, x_2 \to 0} \frac{\partial U(x_1, x_2)}{\partial x_1} = \infty \text{ vs } \lim_{x_1, x_2 \to 0} \frac{\partial U(x_1, x_2)}{\partial x_2} = 1,$$

where the term on the left reflects the increase in utility from buying water, and the term on the right from watching movies. From this, we infer that consuming water gives her a higher utility than watching movies. In fact, the utility attached to consuming water is infinite, consistent with the idea that food is an indispensable item to live.

The analysis can be repeated to identify how she would spend her subsequent cents of dollars. Eventually, we would conclude that her first cents of income will be entirely devoted to the consumption of water; no movies will be watched. Specifically, this way to allocate income holds as long as

$$\frac{\partial U(x_1, x_2)}{\partial x_1} = \frac{1}{x_1} > 1 = \frac{\partial U(x_1, x_2)}{\partial x_2}. \tag{5.2}$$

Equation (8.2) shows that she prefers to keep allocating cents to consuming water, until she is buying a basket with $x_1 = 1$ and $x_2 = 0$. At that point, both expressions in (8.2) are equal. This arises since the more water she consumes the lower the utility she gets marginally, reflecting that she gets more satiated. Mathematically, this follows since $\frac{\partial U(x_1,x_2)}{\partial x_1} = \frac{1}{x_1}$, which satisfies that $\frac{\partial^2 U(x_1,x_2)}{\partial x_1^2} < 0$. In particular, while she gets an infinite utility when $x_1 \to 0$, she only gets $\frac{\partial U(1,x_2)}{\partial x_1} = 1$ when she is already consuming one glass of water.

Once $x_1 = 1$ and $x_2 = 0$, the marginal utilities of both goods become equal. Buys more water would have a contribution to utility lower than 1, since the marginal utility is decreasing. For example, if she spends ten additional cents on water:

$$\frac{\partial U(1.1, 0)}{\partial x_1} = \frac{1}{1.1} < 1 \text{ vs } \frac{\partial U(x_1, x_2)}{\partial x_2} = 1.$$

Hence, from that point on, she will start consuming movies. In fact, since the marginal utility of movies is constant, *she will spend all her remaining income on movies.*

What lessons can we derive from the example considered? First, the solution will depend on the income the consumer has. Although we provided a solution for any level of income, the analysis could have stopped at a point where she is only consuming water. This is in fact the case when she has an income lower than one dollar.

Second, the quasilinear utility is particularly useful when the consumer needs a minimum threshold of a good. This is why we have used the example of water for good 1. She needs to eat to live.

The final conclusion is that when income is really high, any additional income is spent on good 2. This determines that good 1 displays no income effects for high levels of income. Thus, after a certain income threshold, increases in income do not affect the consumption of good 1.

### 5.3.2 UMP

Next, we formally derive the solution to the UMP. With this goal, let's start considering that the consumer has an income high enough that she consumes both goods.

When we only focus on a scenario like this, we can identify the solution by using the Lagrange technique, as we did with the Cobb Douglas. Alternatively, given the structure of the problem, it is easier to plug in the constraint into the utility function, and hence reduce the problem to a maximization with one good. Then, by using the budget constraint, we can recover the solution of the other good.

Either method provides identical results, and we use the latter as is standard in the literature. To do this, we start by writing the budget constraint as $\frac{Y-p_1x_1}{p_2} = x_2$. Plugging it into the utility function, the optimization problem becomes

$$\max_{x_1} U(x_1) = \ln(x_1) + \frac{Y - p_1x_1}{p_2}.$$

Since we are assuming that the consumer is not constrained by her income, we can find its optimal demand of $x_1$ by the FOC:

$$\frac{\mathrm{d}u(x_1)}{\mathrm{d}x_1} = \frac{1}{x_1} - \frac{p_1}{p_2} = 0$$
$$\Rightarrow x_1^* = \frac{p_2}{p_1}.$$

Maybe you feel more comfortable using Lagrange to pin down the solution. This technique can used when income is high enough so that both goods are consumed. Next, I add the derivation, just in case you want to use this alternative.

The Lagrangian is given by:

$\mathscr{L} := \ln(x_1) + x_2 + \lambda[Y - p_1x_1 - p_2x_2]$

and assuming that both goods are consumed, we can identify the solutions through the FOCs, which are

$\mathscr{L}'_{x_1} = \frac{1}{x_1} - \lambda p_1 = 0$

$\mathscr{L}'_{x_2} = 1 - \lambda p_2 = 0$

$\mathscr{L}'_{\lambda} = Y - p_1x_1 - p_2x_2 = 0$

Similar to the Cobb Douglas case, we divide the first two equations, and then substitute in the result into the budget constraint. Whenever you solve the UMP or EMP through Lagrange, I recommend you to try these steps. They usually simplify the calculations considerably.
Applying these steps, we use the equations $\mathscr{L}'_{x_1} = 0$ and $\mathscr{L}'_{x_2} = 0$, determining that $\frac{1}{x_1} = \lambda p_1$ and $1 = \lambda p_2$. Dividing both equations, we obtain that
$\frac{1}{x_1} = \frac{p_1}{p_2}$
and so the Marshallian demand of good 1 is $x_1^*(p_1, p_2) = \frac{p_2}{p_1}$. Embedding the solution $x_1^*$ into the budget constraint, we get:
$Y - p_1\left(\frac{p_2}{p_1}\right) - p_2 x_2 = 0$
from which we obtain $x_2^*(p_1, p_2, Y) = \frac{Y - p_2}{p_2}$.

The solution $x_1^* = \frac{p_2}{p_1}$ provides a specific meaning to the expression "high level of income". It is the income that allows the consumer to get $x_1^* = \frac{p_2}{p_1}$. Put it differently, since the expenditure on good 1 would be $p_1 x_1^* = p_2$, income has to satisfy $Y \geq p_2$ to be sure that she spends $p_2$ on the good 1.

What about the good 2? Once she spends $p_2$ on good, she would spend the rest of her income on good 2. In other words, good 2 acts as a residual variable that absorbs all the income not spent on good 1. Formally, the expenditure on good 2 would be $p_2 x_2^* = Y - p_1 x_1^*$, which allows us to determine that $x_2^* = \frac{Y - p_2}{p_2}$ by using $p_1 x_1^* = p_2$.

We have established the solution by supposing that the consumer has income enough to afford as much as she want of good 1. However, by the intuition provided in the previous section, we need to consider the possibility that the consumer has a low income, and so only good 1 is consumed. A low income means in particular that $Y < p_2$. It represents the case where the marginal utility of good 1 is greater than that of good 2, and so all the income is spent on good 1, nothing on good 2. This solution prevails until income becomes $Y = p_2$. After this, the marginal utility of good 2 becomes greater than that of good 1. Thus, she will not consume additional quantities of good 1, and all the income is spent on the good 2.

Overall, the Marshallian demands are:

$$\begin{aligned} x_1^*(p_1,p_2,Y) &= \begin{cases} \frac{Y}{p_1} & \text{if } Y < p_2 \\ \frac{p_2}{p_1} & \text{if } Y \geq p_2 \end{cases}, \\ x_2^*(p_1,p_2,Y) &= \begin{cases} 0 & \text{if } Y < p_2 \\ \frac{Y-p_2}{p_2} & \text{if } Y \geq p_2 \end{cases}, \end{aligned} \tag{5.3}$$

which formally shows one important feature of the quasilinear utility function: when income is high enough, good 1 displays no income effects.

> **Remark** *Note that the Marshallian demand are not piecewise for specific values of* $(Y, p_1, p_2)$. *If we know these values, the solution will be given by the choices when* $Y \geq p_2$ *or the choices when* $Y < p_2$. *Equation (5.3) is piecewise because it establishes the demand for each possible values of the parameters.*

Finally, by plugging the Marshallian demands into the utility function, we can also obtain the indirect utility function:

$$U^*(p_1,p_2,Y) := \begin{cases} \ln\left(\frac{Y}{p_1}\right) & \text{if } Y < p_2 \\ \ln\left(\frac{p_2}{p_1}\right) + \frac{Y-p_2}{p_2} & \text{if } Y \geq p_2 \end{cases}.$$

### 5.3.3 EMP

Although we could obtain the Hicksian demands by solving the EMP, it is easier to obtain the solution by using duality. This procedure determines that the minimum expenditure function $E^*$ is

$$E^*(p_1,p_2,U_0) := \begin{cases} \exp(U_0)\, p_1 & \text{if } Y < p_2 \\ \left[U_0 + 1 - \ln\left(\frac{p_2}{p_1}\right)\right] p_2 & \text{if } Y \geq p_2 \end{cases}.$$

We use the duality relation between the indirect utility function and the minimum expenditure. Unlike the Cobb-Douglas case, we have a piecewise indirect utility function. Therefore, we need to separately recover the minimum expenditure for $Y < p_2$ and for $Y \geq p_2$.
If $Y < p_2$, then $U^*(p_1,p_2,Y) = \ln\left(\frac{Y}{p_1}\right)$. Hence, by duality, $E^*$ has to satisfy that $U_0 = \ln\left(\frac{E^*(p_1,p_2,U_0)}{p_1}\right)$ which

determines that $E^*(p_1, p_2, U_0) = \exp(U_0)\, p_1$. In case $Y \geq p_2$, then $U_0 = \ln\left(\frac{p_2}{p_1}\right) + \frac{E^*(p_1,p_2,U_0)-p_2}{p_2}$ and so $E^*(p_1, p_2, U_0) = \left[U_0 + 1 - \ln\left(\frac{p_2}{p_1}\right)\right] p_2$.

To obtain the Hicksian demands, we can apply Shepard's Lemma (or use duality again), getting

$$\begin{aligned} h_1^*(p_1, p_2, U_0) &= \begin{cases} \exp(U_0) & \text{if } Y < p_2 \\ \frac{p_2}{p_1} & \text{if } Y \geq p_2 \end{cases}, \\ h_2^*(p_1, p_2, U_0) &= \begin{cases} 0 & \text{if } Y < p_2 \\ U_0 - \ln\left(\frac{p_2}{p_1}\right) & \text{if } Y \geq p_2 \end{cases}. \end{aligned} \tag{5.4}$$

Let's start by deriving the result through Shepard's Lemma. If $Y < p_2$, $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_1} = \exp(U_0)$ and $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_2} = 0$. If $Y \geq p_2$, then $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_1} = \frac{p_2}{p_1}$ and $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_2} = \left[U_0 + 1 - \ln\left(\frac{p_2}{p_1}\right)\right] - 1 \Rightarrow \frac{\partial E^*(p_1,p_2,U_0)}{\partial p_2} = U_0 - \ln\left(\frac{p_2}{p_1}\right)$.

Deriving the solution by duality requires using the Marshallian demands. Then, if $Y < p_2$, $x_1^*(p_1, p_2, Y) = \frac{Y}{p_1}$ and $x_2^*(p_1, p_2, Y) = 0$. Therefore, $h_1^*(p_1, p_2, U_0) = \frac{E^*(p_1,p_2,U_0)}{p_1} \Rightarrow h_1^*(p_1, p_2, U_0) = \frac{p_1 \exp(U_0)}{p_1}$ and $h_2^*(p_1, p_2, U_0) = 0$. If $Y \geq p_2$, then $x_1^*(p_1, p_2, Y) = \frac{p_2}{p_1}$ and $x_2^*(p_1, p_2, Y) = \frac{Y-p_2}{p_2}$. From this, $h_1^*(p_1, p_2, U_0) = \frac{p_2}{p_1}$ and $h_2^*(p_1, p_2, U_0) = \frac{E^*(p_1,p_2,U_0)-p_2}{p_2} \Rightarrow h_2^*(p_1, p_2, U_0) = \frac{\left[U_0+1-\ln\left(\frac{p_2}{p_1}\right)\right]p_2 - p_2}{p_2}$ and dividing by $p_2$ numerator and denominator, $h_2^*(p_1, p_2, U_0) = U_0 - \ln\left(\frac{p_2}{p_1}\right)$.

### 5.3.4 Remarks on The Case Where Income is High

**First Remark**: The quasilinear utility is used in scientific articles assuming that income is high enough that the consumer can afford both goods. Under this scenario, $x_1^*$ does not depend on income.

The fact that good 1 displays no income effects is convenient in several cases. For instance, it is appropriate when a researcher is studying a good that is negligible on a consumer's income, which we would expect to be insensitive to income. This justifies that good 2 is usually interpreted as a composite good that represents the rest of the goods consumed. To illustrate this, consider a researcher analyzing the demand for lollipops. For this good, it is reasonable to assume that an increase in income has an effect close to zero on a person's demand.

The fact that the demand does not depend on income is also useful for models where real income is determined within the model. For instance, it would make sense to think that a lower price of lollipops does not make the consumer richer in real terms. Thus, a quasilinear utility function allows us to ignore feedback effects between the behavior of the industry under analysis and the rest of the economy, where the latter determines the real income of a consumer.

**Second Remark**: this is related to a topic included as optional in the previous lecture note: when are the Marshallian and Hicksian demands equal? Suppose that income is high enough that both goods are consumed, and the Marshallian demand of good 1 is independent of income. In the previous lecture note, we showed that this property holds iff the Hicksian demand does not depend on $U_0$. In fact, we established that zero income effects make the Marshallian and Hicksian demands be identical.

You can directly check this by comparing the demands of good 1 in (5.3) and (5.4). Notice that this result only applies to the good that enters non-linearly into the utility function. Since good 2 displays income effects, the Marshallian and Hicksian demands of good 2 do not coincide.

### 5.3.5 A Parameter Reflecting Intensity of Preferences

I conclude the presentation of quasilinear utility functions by formalizing a scenario through its use. The aim is to show that consumer theory is flexible enough to model a real-life phenomenon.

Suppose the cell phone industry, and that new iPhone is about to be launched. You receive information regarding the new features that will have and conclude that this will substantially increase its demand. How can we capture that a good has become more appealing for consumers?

Let's modify the baseline quasilinear utility function with this goal. Suppose that good 1 is the iPhone and good 2 is a composite good that represents the rest of the goods in the economy. Consider that the utility function is

$$\max_{x_1, x_2} U(x_1, x_2) = A \ln(x_1) + x_2.$$

where $A > 0$.

The parameter $A$ is incorporated in a way that it affects the marginal utility of good 1:

$$\frac{\partial U(x_1, x_2)}{\partial x_1} = \frac{A}{x_1},$$

so that a greater $A$ results in a higher utility when good 1 is consumed. Thus, $A$ represents the intensity in which a consumer likes the iPhone.

## 5.4 Exercises

[1] Alberta's government is designing a policy to reduce the consumption of addictive substances (henceforth, ASs), among the population of heavy users. The ASs comprise illegal drugs, but also legal ASs like cigarettes. You've been hired to assess some of the policies they are considering. To this end, you build a model that represents a heavy consumer's behavior. You represent the consumer's choices in terms of consumption per day.

The utility function you suppose is $U(x_1, x_2) := A\sqrt{x_1} + x_2$ with $A > 0$, where good 1 represents ASs and good 2 is a composite good representing the rest of goods.

To quantify each policy's magnitude, you find out after some research that a typical heavy user has per-day income $Y := 10$ CAD, and that preferences can be described by $A := 8$. To keep matters simple, also suppose that $p_1 := 2$, and normalize good 2's price so that $p_2 := 1$.

(a) How can you justify the choice of this utility function to model ASs?

(b) Solve the consumer's maximization problem. Then, determine each good's Marshallian demands for the parameters given. In particular, calculate how much the representative agent spends on each good. (*hint*: since later you'll have to assume different values for the parameters, solve the problem parametrically and then replace for the values given).

(c) What is the role of the parameter $A$ in this model? In particular, how do increases in $A$ affect the consumption decisions and why?

(d) A public servant is worried about the consequences in case heavy users have higher income. Her concern is that these people would consume ASs more heavily. What would you say about it?

(e) One of the proposals is levying a tax on legal ASs (for instance, cigarettes). In fact, prohibitive taxes for legal ASs has been quite common worldwide in the last decades. Suppose in particular that this is done through a value-added tax: 100% over the price paid by the consumer. Will the policy be effective? Considering the set of parameters given, how many units will the consumer stop buying?

(f) Another proposal has the illegal ASs as the target. This posits an issue: you can not levy a tax on the consumption of those goods, since they are sold in the black market. For this reason, one (not really popular) policy under study is to set a 10% income tax. Using the set of parameters given, will this policy be effective? Calculate each good's consumption under this policy, and then interpret the result.

(g) Another policy under consideration consists of a marketing campaign against the use of ASs. This will diminish the consumption of heavy users by changing their preferences. How would you capture this policy in terms of the model?

[2] Bart has an astonishing Krusty doll that has caught Milhouse's attention. Milhouse is quite interested in buying that toy to Bart. The goal of the exercise is to determine under what conditions there will be a transaction.

Bart and Milhouse receive a monthly allowance, which is $Y^B > 0$ for Bart and $Y^M > 0$ for Milhouse. They use that money to buy comics (good 2), whose price is $p_2 := 2$. Besides, the Krusty doll (good 1) was a Homer's gift that gives Bart a utility equal to $\ln(2)$. If Bart did not have the doll, he would derive zero utility from that good. This information is captured by the following utility function for Bart:

$$U^B := \begin{cases} x_2 & \text{if } x_1 = 0 \\ \ln(2) + x_2 & \text{if } x_1 = 1 \text{ (i.e. if he keeps the doll)} \end{cases}$$

Regarding Milhouse, he would derive a utility of $\ln(1 + A)$ if he gets the toy, and zero utility if he has not have it (in which case, he'd only derive utility from good

2). This is captured by the following the utility function:

$$U^M := \begin{cases} x_2 & \text{if } x_1 = 0 \\ \ln(1+A) + x_2 & \text{if } x_1 = 1 \text{ (i.e., if he gets the doll)} \end{cases}$$

(a) Calculate the indirect utility function of Bart and Milhouse, given the allowances they will use to buy comics.

(b) Suppose that Milhouse offers Bart an amount of money equal to $M$. What is the range of values for $M$ that induces Bart to sell the doll to Milhouse? (hint: compare Bart's utility of keeping the doll, in contrast to selling it and getting $M$)

(c) Suppose that Bart asks for an amount $B$ to Milhouse. What is the range of values $B$ that induce Milhouse to accept the offer? (Notice that $B$ will be a function of $A$) (hint: compare Milhouse's utility of getting the toy and paying $B$, relative to not getting it).

(d) Bart and Milhouse cannot agree on the doll price, and they ask Marge to set a price. What restriction of values does $A$ need to satisfy such that Marge can find a price where Bart is willing to sell the good and Milhouse is willing to buy it? Interpret the result.

**Answer Keys for Some of the Exercises**:

1b) $x_1^* = 4$ and $x_2^* = 2$, 1e) $x_1^* = 1$ so it'll decrease the consumption in 3 units, 1f) $x_1^* = 4$ and $x_2^* = 1$.

2b) $M \geq 1.38$, 2c) $B \leq 2\ln(1+A)$, 2d) $A \geq 1$.

# Lecture Note 6

# Non-Well-Behaved Utility Functions

## 6.1 Introduction

So far, we have analyzed consumer theory by using well-behaved utility functions. In particular, we studied the Cobb Douglas and the quasilinear utility function. In this set of notes, we study three utility functions that do not satisfy at least one of the axioms of well-behaved utility functions. These cases are referred to as perfect substitutes (linear utility), perfect complements (Leontief function), and the max utility. As we will see, they capture interesting phenomena that arise naturally in many circumstances.

The treatment of these utility functions is separate from the rest, since we cannot use the FOC to solve the optimization problems related. Instead, we will have to analyze each function's behavior to determine their solutions. Throughout the analysis, we assume that the consumption space is $X_1 \times X_2 := \mathbb{R}^2_+$.

## 6.2 Leontief Function (Perfect Complements)

The Leontief utility function is given by

$$U(x_1, x_2) := \inf\left\{\frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}\right\}$$

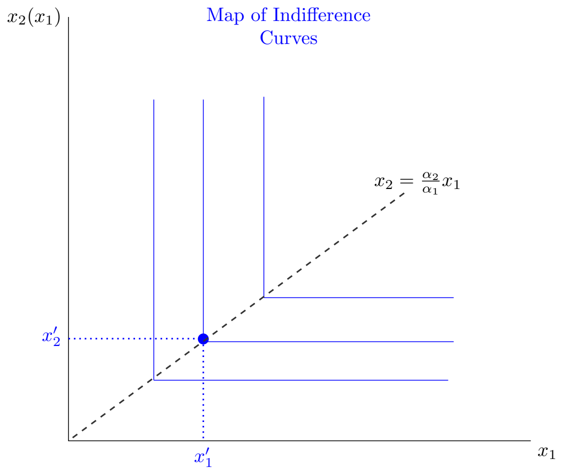
where $\alpha_1, \alpha_2 > 0$.

Since the Leontief function gives the minimum of the two arguments as an output, it can also be reexpressed as:

$$U(x_1, x_2) := \begin{cases} \frac{x_1}{\alpha_1} & \text{if } \frac{x_1}{\alpha_1} \leq \frac{x_2}{\alpha_2} \\ \frac{x_2}{\alpha_2} & \text{if } \frac{x_1}{\alpha_1} > \frac{x_2}{\alpha_2} \end{cases}$$

By expressing the function in this way, we can show that the indifference curves have the shape depicted in Figure 6.1. The dashed line passes through the vertex of each L-shaped curve. This indicates that all the points along the dashed line satisfy $\frac{x_1}{\alpha_1} = \frac{x_2}{\alpha_2}$, and so are represented by the line $x_2 = \frac{\alpha_2}{\alpha_1} x_1$.

**Figure 6.1.** *Indifference Curves - Leontief Function*




To understand what the Leontief function entails, $\alpha_i$ can be interpreted as the units of good $i$ that make the argument $i$ of the inf function equal to one. We can see this by setting $x_1 = \alpha_1$ and $x_2 = \alpha_2$, so that each argument becomes equal to one, and so the consumer obtains one unity of utility.

But what happens if, starting from a basket with $x_1 = \alpha_1$ and $x_2 = \alpha_2$, we increase the consumption of one good in isolation? Then the infimum would still be one, providing a level of utility equal to one. Graphically, when $x_1 = \alpha_1$ and $x_2 = \alpha_2$, we are at the vertex of the indifference curve. This implies that the indifference curve are horizontal if $x_1$ increases, and are vertical if $x_2$ increases.

Intuitively, the Leontief function represents situations where there is no possibility of substitution between goods. Thus, if we start from a situation where $\frac{x_1}{\alpha_1} = \frac{x_2}{\alpha_2}$, increasing the consumption of one of the goods in isolation would not increase the utility—both goods are essential for the consumer. The consequence of this is that the consumption of both goods needs to increase simultaneously and in a proportion $\alpha_1$ and $\alpha_2$ to increase utility.

An extreme example of perfect complements would arise if a clothing store sold left shoes (good 1) and right shoes (good 2) separately. The consumer then would only derive utility if $x_1 = x_2$, otherwise she would not form a pair of shoes.

### 6.2.1 Intuition for the UMP

To grasp some intuition before solving the UMP formally, we provide an illustration by giving specific values to the parameters. The UMP is,

$$\max_{x_1,x_2} U(x_1,x_2) = \inf\left\{\frac{x_1}{\alpha_1},\frac{x_2}{\alpha_2}\right\} \text{ subject to } Y = p_1x_1 + p_2x_2.$$

To fix ideas, let good 1 be cups of coffee and good 2 packets of sugar.

**Figure 6.2.** *Example of a Case Captured by the Leontief Function*

Suppose that the preferences of the consumer are such that she neither enjoys drinking coffee without sugar nor eating sugar alone. She only derives utility when she drinks coffee with some sugar.

More specifically, we will suppose that she prefers to have two sugar packets per cup of coffee. In terms of the Leontief function, the proportion in which she prefers to consume each good is reflected through the parameters $\alpha_1$ and $\alpha_2$. Their values in the example are $\alpha_1 := 1$ and $\alpha_2 := 2$. This determines that if $x_1 = 1$ (one cup of coffee) and $x_2 = 2$ (two sugar packets), then she obtains one unit of utility.

Suppose her income is $Y := 2$. Moreover, each packet of sugar costs 50 cents, while one cup of coffee 1 dollar. Formally, $p_1 = 1$ and $p_2 = 0.5$. Then, her maximization problem is

$$\max_{x_1,x_2} U(x_1,x_2) = \inf\left\{x_1, \frac{x_2}{2}\right\} \text{ subject to } 2 = x_1 + \frac{x_2}{2}.$$

Consider now that she uses her two dollars of income to consume one cup of coffee and two sugar packets. Then, her utility would be

$$U(1,2) = \inf\left\{1, \frac{2}{2}\right\} = 1.$$

What does occur if she has two additional dollars available? Let's first consider the utility she gets by spending those additional two dollars in different ways. If she buys two cups of coffee, without increasing the amount of sugar, she would get

$$U(3,2) = \inf\left\{3, \frac{2}{2}\right\} = 1.$$

so her utility would remain the same. This arises since a Leontief utility function reflects that she would consume one cup of coffee with two sugar packets, and throw away the two other cups of coffee. The reason is that she has no sugar for these two cups of coffee.

Suppose that, instead, she decides to exclusively spend the extra two dollars on sugar, increasing her number of packets to a total of four units. Then,

$$U(1,6) = \inf\left\{1, \frac{6}{2}\right\} = 1,$$

and the utility remains the same, because she does not get any additional utility from consuming sugar without coffee. We could also rationalize this result as she throwing those four additional packets away, because she has no additional coffee. Consequently, she still gets one unit of utility, because she consumes one cup of coffee with two sugar packets.

From this, we can conclude that she only increases her utility if she simultaneously buys both more coffee and sugar. In fact, the way in which she can spend the additional income more efficiently is by buying two sugar packets and one cup of coffee. More generally, every time her income increases, she needs to increase the consumption of both goods in a proportion of two sugar packets per cup of coffee. This implies that the optimal way to spend the extra two dollars is by buying one more cup of coffee and two more sugar packets, yielding

$$U(2,4) = \inf\left\{2, \frac{4}{2}\right\} = 2$$

What happens if she spends her additional income in a different proportion? Although she would increase her utility relative to consuming only one cup of coffee with two sugar packets, *there would always be an excess of one of the goods, which she would not consume.* For instance, suppose she buys 1.5 cups of coffee and 1 packet of sugar. Then,

$$U(2.5,3) = \inf\left\{2.5, \frac{3}{2}\right\} = 1.5$$

which is lower than the utility of the basket $(2,4)$. The intuition is that she only has 3 sugar packets available, and so she would only drink one a half cups of coffee. In this context, it is optimal not to buy that additional amount of coffee, and instead spend the additional money to buy sugar and coffee in a proportion of 2-to-1.

**Remark** *Two goods could be perfect complements in some contexts, but not in others—ultimately, it depends on the situation we are analyzing. For example, a couple could find that having dinner in a restaurant before watching a movie at a theater are complementary activities. However, they could prefer only to watch a movie if they do so at home.*

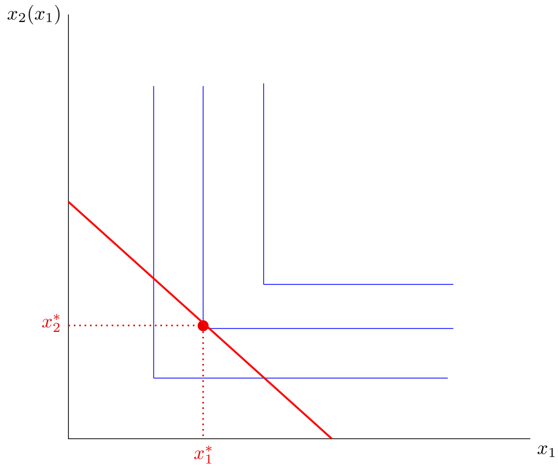
### 6.2.2 UMP

Let's now formally solve the UMP when a consumer is described by a Leontief function. As a necessary condition, the solution to the optimization problem requires that both arguments of the min function are equal:

$$\frac{x_1}{\alpha_1} = \frac{x_2}{\alpha_2}$$

This determines a relation between $x_1$ and $x_2$, given by $x_1 = \frac{\alpha_1}{\alpha_2} x_2$.

**Figure 6.3.** *Optimal Solution- Leontief Function*




**Note**: The red line represents the budget line. The blue lines are the indifference curves which are convex but not strictly convex. The red dot is the optimal bundle.

Plugging in $x_1 = \frac{\alpha_1}{\alpha_2} x_2$ into the budget constraint, we obtain $Y = p_1 \left( \frac{\alpha_1}{\alpha_2} x_2 \right) + p_2 x_2$, which provides

$$x_2^* \left(p_1, p_2, Y\right) = Y \frac{\alpha_2}{\alpha_1 p_1 + \alpha_2 p_2}.$$

Likewise, by using that$x_1^* \left(p_1, p_2, Y\right) = \frac{\alpha_1}{\alpha_2} x_2^* \left(p_1, p_2, Y\right)$, we can recover the Marshallian demand of good 1:

$$x_1^* \left(p_1, p_2, Y\right) = Y \frac{\alpha_1}{\alpha_1 p_1 + \alpha_2 p_2}.$$

Finally, the indirect utility function is $U^* \left(p_1, p_2, Y\right) = \inf \left\{ \frac{Y}{\alpha_1 p_1 + \alpha_2 p_2}, \frac{Y}{\alpha_1 p_1 + \alpha_2 p_2} \right\}$, yielding

$$U^* \left(p_1, p_2, Y\right) = \frac{Y}{\alpha_1 p_1 + \alpha_2 p_2}.$$

We can formally show that the utility would be lower with any basket where the arguments of the inf function are not equal. Hence, any of those baskets cannot be part of the solution.

Suppose the consumer splits the income in a proportion $\delta \in [0,1]$ of good 1 and $(1-\delta)$ of good 2. Thus, the consumption of each good is obtained from $p_1 x_1 = \delta Y$ and $p_2 x_2 = (1-\delta) Y$, which determines that $x_1 = \frac{\delta}{p_1} Y$ and $x_2 = \frac{(1-\delta)}{p_1} Y$. Notice that, by using this representation of baskets, we are encompassing all the possible bundles that can be conceived. This is done by choosing different values of $\delta$.

For a given $\delta$, either $U = \frac{\delta}{\alpha_1} \frac{Y}{p_1}$ or $U = \frac{(1-\delta)}{\alpha_2} \frac{Y}{p_2}$, depending on which one is the lowest. Suppose that the first one is the lowest, then $\frac{\delta}{\alpha_1} \frac{Y}{p_1} < \frac{(1-\delta)}{\alpha_2} \frac{Y}{p_2}$. This inequality can be reexpressed as:

$\delta < (1-\delta) \frac{\alpha_1 p_1}{\alpha_2 p_2} \Rightarrow \delta \left( \frac{\alpha_1 p_1 + \alpha_2 p_2}{\alpha_2 p_2} \right) < \frac{\alpha_1 p_1}{\alpha_2 p_2} \Rightarrow \delta < \frac{\alpha_1 p_1}{\alpha_1 p_1 + \alpha_2 p_2}$.

But when the arguments of the Leontief function are equal, then the utility is $\frac{Y}{\alpha_1 p_1 + \alpha_2 p_2}$. Besides, it can be shown that $\frac{Y}{\alpha_1 p_1 + \alpha_2 p_2} > \frac{\delta}{\alpha_1} \frac{Y}{p_1}$ iff $\frac{\alpha_1 p_1}{\alpha_1 p_1 + \alpha_2 p_2} > \delta$, and the latter inequality holds since we were assuming that the first argument of the Leontief function was the minimum.

A similar proof could be provided when the second argument of the Leontief function is the minimum.

### 6.2.3 EMP

By using duality, we can easily obtain the solution of the EMP. Using the expression for the indirect utility function, we know that $E^*$ has to satisfy $U_0 = \frac{E^*(p_1,p_2,U_0)}{\alpha_1 p_1+\alpha_2 p_2}$, which determines

$$E^*(p_1, p_2, U_0) = (\alpha_1 p_1 + \alpha_2 p_2) U_0.$$

By Shepard's Lemma, which states that $\frac{\partial E^*(p_1,p_2,U_0)}{\partial p_i} = h_i^*(p_1, p_2, U_0)$, we can obtain the Hicksian demands:

$$h_1^*(p_1, U_0) = \alpha_1 U_0,$$

$$h_2^*(p_1, U_0) = \alpha_2 U_0.$$

Remember we could have also obtained the Hicksian demands by applying duality to the Marshallian demands.[1]

## 6.3 Max Function

To fix ideas, let's consider a specific example. Suppose a person is at a cafe, and she is deciding whether to consume a cup of coffee (good 1) or a cup of tea (good 2).

**Figure 6.4.** *Example of a Case Captured by the Max Function*


[1]For instance, the Marshallian for good 1 is $x_1^*(p_1, p_2, Y) = Y \frac{\alpha_1}{\alpha_1 p_1+\alpha_2 p_2}$ and, by duality, it satisfies that $h_1^*(p_1, p_2, U_0) = E^*(p_1, p_2, U_0) \frac{\alpha_1}{\alpha_1 p_1+\alpha_2 p_2}$. By using $E^*(p_1, p_2, U_0) = (\alpha_1 p_1 + \alpha_2 p_2) U_0$, then we would get the same result.

When the cashier has to take her order, we would be surprised if she orders coffee and tea at the same time. Rather, we expect that she consumes one or the other. Based on this, if we have to specify a utility function that represents this situation, it should be such that the UMP provides corner solutions and never an interior solution. In other terms, even if she is offered a diversified basket for free that includes coffee and tea, she would only derive utility by consuming one and only one of the goods. The fact that the max utility function rules out consuming both goods simultaneously is an important difference relative to the case we analyze below, given by a linear utility function.

Formally, the type of situation described can be captured by the following UMP:

$$\max_{x_1,x_2} U(x_1,x_2) = \sup\{\beta_1 x_1, \beta_2 x_2\} \text{ subject to } Y = p_1 x_1 + p_2 x_2.$$

Rcall that the sup function can reexpressed as a piecewise function, such that the utility function is equivalent to:

$$U(x_1,x_2) := \begin{cases} \beta_1 x_1 & \text{if } \beta_1 x_1 \geq \beta_2 x_2, \\ \beta_2 x_2 & \text{if } \beta_1 x_1 < \beta_2 x_2. \end{cases}$$

Let's analyze the max utility function for some specific parameter values. Suppose the consumer likes coffee more than tea, which is parametrically reflected through $\beta_1 > \beta_2$. Suppose in particular that $\beta_1 := 2$ and $\beta_2 := 1$, so that $U(x_1,x_2) = \sup\{2x_1, x_2\}$.

If she consumes one cup of coffee and none of tea, her utility would be

$$U(1,0) = \sup\{2,0\} = 2$$

Suppose now that there is a promotion where you get one cup of tea for free for every cup of coffee you consume. Then, her utility would be

$$U(1,1) = \sup\{2,1\} = 2$$

and so the utility remains the same. This occurs since she would only consume one cup of either coffee or tea, but not both together. In particular, as she likes coffee more, she would keep consuming one cup of coffee.

To provide a solution with parameters, assume that $p_1 := 1$, $p_2 := 1$, and $Y := 1$.

This allows her to buy exactly one cup of coffee or tea. As we have said, the only relevant baskets that maximize utility are those with only one good consumed. Since $U(1,0) = \sup\{2,0\} = 2$ and $U(0,1) = \sup\{0,1\} = 1$, she would end up consuming one cup of coffee.

What happens if we consider an interior solution? To see this, suppose that she splits her income, spending a proportion $\theta \in (0,1)$ on coffee and a proportion $(1-\theta)$ on tea. Since prices and income are equal to one, then the basket consumed is $(\theta, 1-\theta)$. Thus, for a given $\theta$, either $U(\theta, 1-\theta) = 2\theta$ or $U(\theta, 1-\theta) = 1-\theta$, depending on the argument of the utility function that is greater. But both values are lower than $U(1,0) = 2$ for any $\theta$, and so they cannot be a solution—the utility would be lower for any $\theta$, because she would only consume one of the goods.

One additional case to consider is a knife-edge case, where the agent is indifferent between consuming one unit of each good, i.e. between the potential corner solutions. This scenario is important, since it is when the max function differs from the case of perfect substitutes we study below.

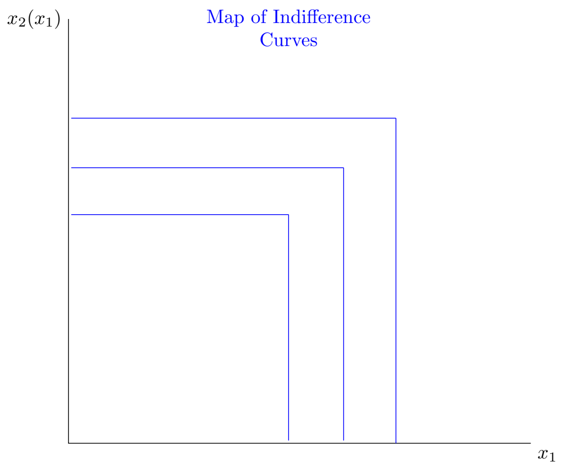
For instance, take $p_1 := 2$, $p_2 := 1$, and $Y := 1$. The maximum quantity she can afford of each good is $x_1 = 0.5$ and $x_2 = 1$. Hence, consuming only one of the goods determines:

$$U(1,0) = \sup\{2 \times 0.5, 0\} = 1 \text{ and } U(0,1) = \sup\{0,1\} = 1$$

Since both options provide the same utility, then she will either buy half a cup of coffee or one cup of tea. But, importantly, she would never consume some combination of both, because any interior solution would decrease her utility.

In summary, the conclusion is that, irrespective of the prices and income, a consumer with a sup utility function will *always* consume only one good. Although the intuition provided should be clear, you can also conclude this by looking at the shape of the indifference curves. They are illustrated in the following figure.

**Figure 6.5.** *Indifference Curves - Max Function*




**Note**: The blue lines are the indifference curves. They are concave but not strictly concave. This implies that the consumer does not like diversifying consumption.

### 6.3.1 UMP

Recall that the UMP is given by

$$\max_{x_1,x_2} U(x_1, x_2) = \sup\{\beta_1 x_1, \beta_2 x_2\} \text{ subject to } Y = p_1 x_1 + p_2 x_2.$$

As with the case of perfect complements, we cannot use the Lagrangian technique to obtain a solution—the utility function is not differentiable everywhere. Due to this, we proceed in two steps.

The first step builds on the intuition we provided before. We first calculate the utilities when only one of the goods is consumed. Then, we show that any basket with positive consumption of both goods provides less utility, and so they can be ruled out as potential solutions.

In the second step, we use the fact that, by definition, a consumption bundle is optimal if it provides the greatest utility. Therefore, we compare the utility of each corner solution, and identify the value of the parameters such that one utility is greater than the other. From this, we establish values of parameters that are consistent with one specific bundle as a solution.

**Step 1**. In case the agent only consumes good 1, the budget constraint indicates that

$$Y = p_1 x_1 + p_2 0 \Rightarrow x_1 = \frac{Y}{p_1},$$

and utility that she derives from this consumption is

$$U\left(\frac{Y}{p_1}, 0\right) = \beta_1 \frac{Y}{p_1}.$$

By the same token, if she only consumes good 2, then $x_2 = \frac{Y}{p_2}$ and so

$$U\left(0, \frac{Y}{p_2}\right) = \beta_2 \frac{Y}{p_2}.$$

Only these corner solutions need to be considered to identify the optimal solution.

Let's show formally that any basket with positive amounts of both goods provides less utility. Let's focus on the case where consuming exclusively good 1 provides a greater utility than consuming only good 2, so that $\beta_1 \frac{Y}{p_1} > \beta_2 \frac{Y}{p_2}$.This without loss of generality, because "1" is just a label for the good; if good 2 provides a greater utility, then we could relabel the problem and take it as good 1.

Suppose she considers to allocate some expenditure $Y - \delta$ on good 1 and expenditure $\delta$ on good 2. We assume that $0 < \delta < Y$, so that it covers any potential interior consumption we could think of. The utility she obtains with this income allocation is still given by the maximum utility, where she consumes only one of the goods. Consequently, depending on the value of $\delta$, the maximum utility is either $U = \beta_1 \frac{Y-\delta}{p_1}$ or $U = \beta_2 \frac{\delta}{p_2}$. But, for any value of $\delta > 0$, we have that $\beta_1 \frac{Y}{p_1} > \beta_1 \frac{Y-\delta}{p_1}$. Moreover, $\beta_2 \frac{Y}{p_2} > \beta_2 \frac{\delta}{p_2}$ for any value of $\delta < Y$, and since $\beta_1 \frac{Y}{p_1} > \beta_2 \frac{Y}{p_2}$, then $\beta_1 \frac{Y}{p_1} > \beta_2 \frac{\delta}{p_2}$. Therefore, there cannot be an interior solution, since any of them would provide a lower utility than exclusively consuming good 1.

**Step 2**. We have ruled out interior solutions. We also argued that the only two candidates for a solution are those bundles in which only one good is consumed. Next, we establish the parameters where consuming only good 1 is the solution and those where consuming good 2 is the solution.

Consuming only good 1 is optimal when it provides the greatest utility among the two corner solutions. Thus, it is optimal to consume good 1 when:

$$U\left(\frac{Y}{p_1}, 0\right) > U\left(0, \frac{Y}{p_2}\right),$$
$$\Rightarrow \beta_1 \frac{Y}{p_1} > \beta_2 \frac{Y}{p_2},$$
$$\Rightarrow \frac{p_1}{p_2} < \frac{\beta_1,}{\beta_2}$$

The interpretation of this condition is straightforward if we reexpress it as $\frac{\beta_1}{p_1} > \frac{\beta_2}{p_2}$. The term $\frac{\beta_i}{p_i}$ can be understood as the utility got per dollar spent on cups of good $i$. If the utility per dollar is greater for coffee than for tea, then she will consume coffee.

By the same token, consuming exclusively good 2 is a solution when

$$U\left(\frac{Y}{p_1},0\right) < U\left(0,\frac{Y}{p_2}\right) \Rightarrow \beta_1\frac{Y}{p_1} < \beta_2\frac{Y}{p_2}$$
$$\Rightarrow \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2}$$

It rests to determine what happens when $\frac{p_1}{p_2} = \frac{\beta_1}{\beta_2}$, in which case the utility of each corner consumption gives the same level of utility. The proof that interior consumptions are never a solution applies to this case too. Therefore, we can rule out any of those bundles as potential solutions, and they have to be a corner solution. Furthermore, since both corner solutions provide the same utility, we conclude that both are actually a solution.

Summing up, the Marshallian demands are:

$$x_1^*(p_1,p_2,Y) := \begin{cases} \frac{Y}{p_1} & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ 0 & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \left\{0,\frac{Y}{p_1}\right\} & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases} \text{ and } x_2^*(p_1,p_2,Y) := \begin{cases} 0 & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ \frac{Y}{p_2} & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \frac{Y}{p_2}\mathbf{1}_{(x_1^*=0)} & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases}$$

The function $\mathbf{1}_{(x_1^*=0)}$ is known as the *indicator function*. It takes the value 1 if the condition in brackets is met (in this case, if $x_1^* = 0$) and zero otherwise. It is just a compact way to write that $x_2^*$ would be equal to $\frac{Y}{p_2}$ when $x_1^* = 0$, and equal to 0 when $x_1^* = \frac{Y}{p_1}$.

**The Marshallian demands with a sup utility function are not functions, but correspondences**. When $\frac{p_1}{p_2} = \frac{\beta_2}{\beta_1}$, we have that $x_1^*$ could be either 0 or $\frac{Y}{p_1}$, which means that there are two potential solutions. Instead, a function requires assigning only one value of the co-domain for each value of the domain (in this case, for each value of prices and income). Correspondences refer to the case when there are multiple values.

The indirect utility function is $U^*(p_1,p_2,Y) = \sup\{\beta_1 x_1^*(p_1,p_2,Y), \beta_2 x_2^*(p_1,p_2,Y)\}$, which can be rewritten as

$$U^*(p_1,p_2,Y) = \begin{cases} \beta_1\frac{Y}{p_1} & \text{if } \frac{p_1}{p_2} \leq \frac{\beta_1}{\beta_2}, \\ \beta_2\frac{Y}{p_2} & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2}. \end{cases}$$

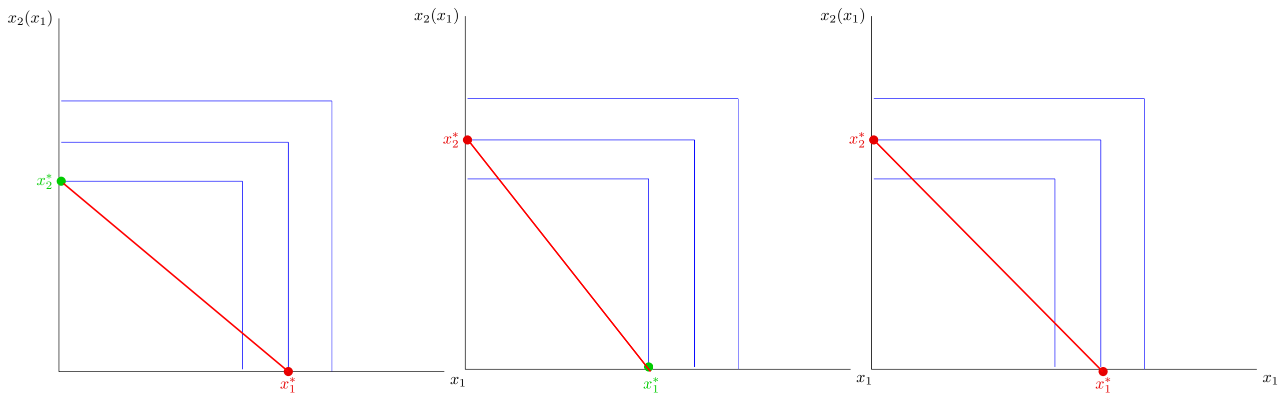
Notice that $\beta_1\frac{Y}{p_1}$ is the greatest utility when $\frac{p_1}{p_2} = \frac{\beta_2}{\beta_1}$. Instead, I could have established that $\beta_2 \frac{Y}{p_2}$ is the greatest utility. Both alternatives are valid, since $\beta_1 \frac{Y}{p_1} = \beta_2 \frac{Y}{p_2}$.

Finally, a more compact way to rewrite $U^*$ is

$$U^* (p_1, p_2, Y) = Y \sup \left\{ \frac{\beta_1}{p_1}, \frac{\beta_2}{p_2} \right\}$$

**Figure 6.6.** *Optimal Consumption - Max Function*

**(a)** Only Good 1 **(b)** Only Good 2 **(c)** Any Good




**Note**: The red line is the budget line. Blue lines are indifference curves. Moving in the north east direction, utility increases. Thus, indifference curves further away from the origin provide a greater utility. The red dot bundle provides the greatest utility and, in particular, it provides more utility than the green dot bundle.

### 6.3.2 EMP

By duality and given the indirect utility function, $U_0 = E^* (p_1, p_2, U_0) \sup \left\{ \frac{\beta_1}{p_1}, \frac{\beta_2}{p_2} \right\}$. Thus, the minimum expenditure function is

$$E^* (p_1, p_2, U_0) = \frac{U_0}{\sup \left\{ \frac{\beta_1}{p_1}, \frac{\beta_2}{p_2} \right\}}$$

The Hicksian demands are:

$$h_1^* (p_1, p_2, U_0) := \begin{cases} \frac{U_0}{\beta_1} & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ 0 & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \left\{ 0, \frac{U_0}{\beta_1} \right\} & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases} \text{ and } h_2^* (p_1, p_2, U_0) := \begin{cases} 0 & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ \frac{U_0}{\beta_2} & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \frac{U_0}{\beta_2} \mathbf{1}_{(h_1^*=0)} & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases}$$

where $\mathbf{1}_{(h_1^*=0)}$ is the indicator function taking the value 1 if the condition under brackets (i.e. $h_1^* = 0$) holds, and zero otherwise (i.e. $h_1^* \neq 0$).

To obtain the Hicksian demands, we use duality applied to the Marshallian demands. Suppose $\frac{p_1}{p_2} < \frac{\beta_1}{\beta_2}$. Then, $x_1^*(p_1, p_2, Y) = \beta_1 \frac{Y}{p_1}$ and, by duality,
$h_1^*(p_1, p_2, U_0) = \frac{E^*(p_1, p_2, U_0)}{p_1}$. Since $\frac{p_1}{p_2} < \frac{\beta_1}{\beta_2}$ implies that $\frac{\beta_1}{p_1} > \frac{\beta_2}{p_2}$, this determines $h_1^*(p_1, p_2, U_0) = \frac{U_0}{\beta_1}$.
Suppose now that $\frac{p_1}{p_2} > \frac{\beta_1}{\beta_2}$. Then, $x_1^*(p_1, p_2, Y) = 0$ and, by duality, $h_1^*(p_1, p_2, U_0) = 0$.
If $\frac{p_1}{p_2} = \frac{\beta_1}{\beta_2}$ then either $x_1^*(p_1, p_2, Y) = 0$ or $x_1^*(p_1, p_2, Y) = \frac{Y}{p_1}$. So, the Hicksian demand is either $h_1^*(p_1, p_2, U_0) = 0$ or $h_1^*(p_1, p_2, U_0) = \frac{E^*(p_1, p_2, U_0)}{p_1}$ which is just $h_1^*(p_1, p_2, U_0) = \frac{U_0}{\beta_1}$. A similar procedure can be used for good 2.

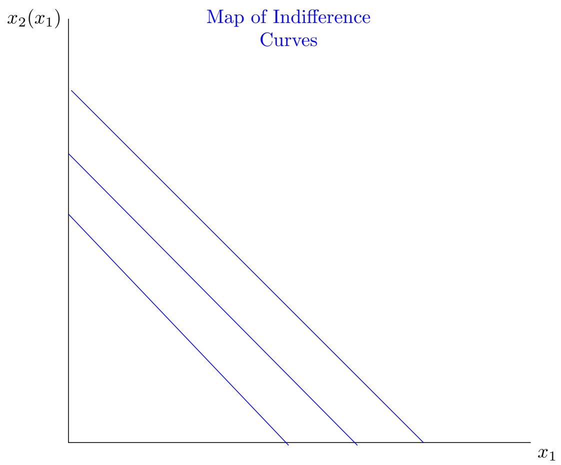
## 6.4 Linear Utility (Perfect Substitutes)

The utility for perfect substitutes is

$$U(x_1, x_2) := \beta_1 x_1 + \beta_2 x_2$$

where $\beta_1, \beta_2 > 0$. The indifference curves have the following shape:

**Figure 6.7.** *Indifference Curves - Linear Function*




**Note**: The indifference curves are convex but not strictly convex. Remember that a linear function is convex and concave at the same time.

Regarding the solution, the linear utility is akin to the sup utility. In particular, we have the same solution when $\frac{p_1}{p_2} < \frac{\beta_1}{\beta_2}$ or $\frac{p_1}{p_2} > \frac{\beta_1}{\beta_2}$, and we can arrive at them in exactly the same way. The difference between both utility functions arises when $\frac{p_1}{p_2} = \frac{\beta_2}{\beta_1}$. We proceed to explain mathematically how to find the solution when this occurs, and then explain intuitively when the linear utility should be used.

### 6.4.1 UMP

Suppose $\frac{p_1}{p_2} = \frac{\beta_1}{\beta_2}$, which implies that $\frac{\beta_1}{p_1} = \frac{\beta_2}{p_2}$. Unlike the case of the max function, any basket that satisfies the budget line is a solution under a linear utility. Thus, if $\frac{p_1}{p_2} = \frac{\beta_1}{\beta_2}$, the solution is $x_1^* \in \left[0, \frac{Y}{p_1}\right]$ and, from the budget constraint, $x_2^* = \frac{Y - p_1 x_1^*}{p_2}$.

Assume the consumer allocates expenditures $Y - \delta$ to good 1 and $\delta$ to good 2 where $0 \leq \delta \leq Y$. Notice that by considering different values of $\delta$, any feasible expenditure we could think of would be covered, including both interior and corner expenditures. Once that we have established the expenditures, we can determine the quantities consumed: since $p_1 x_1 = Y - \delta$ then $x_1 = \frac{Y-\delta}{p_1}$. By the same token, $x_2 = \frac{\delta}{p_2}$. The utility that this basket gives is

$$\begin{aligned} U\left(\frac{Y-\delta}{p_1}, \frac{\delta}{p_2}\right) &= \beta_1 \frac{(Y-\delta)}{p_1} + \beta_2 \frac{\delta}{p_2} \\ &= \frac{\beta_1}{p_1} Y - \delta \left(\frac{\beta_1}{p_1} - \frac{\beta_2}{p_2}\right) \end{aligned}$$

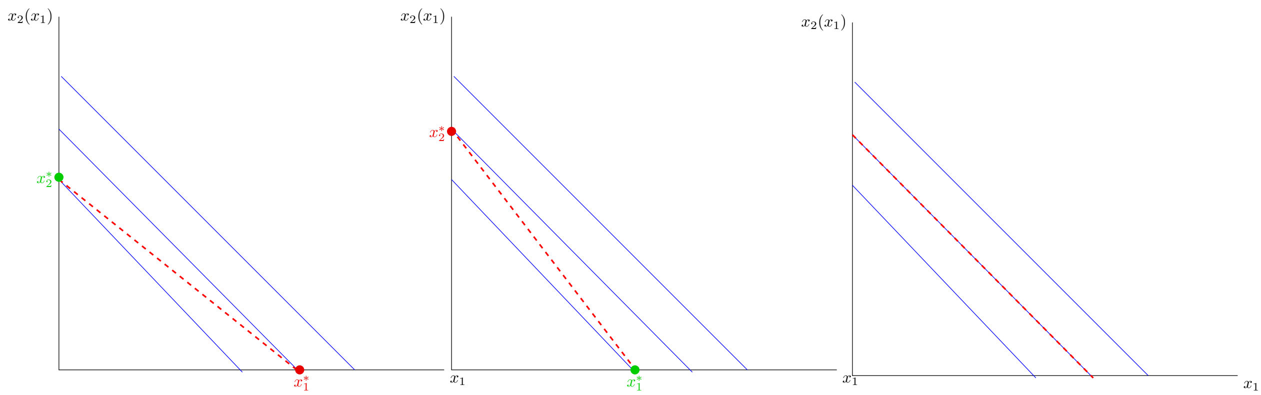
and, since $\frac{\beta_1}{p_1} = \frac{\beta_2}{p_2}$, then

$$U\left(\frac{Y-\delta}{p_1}, \frac{\delta}{p_2}\right) = \frac{\beta_1}{p_1} Y \text{ for any } 0 \leq \delta \leq Y$$

This shows that the utility level is independent of $\delta$. In other terms, irrespective of the bundle we consider (obtained by giving different values to $\delta$), the utility is always the same. Therefore, since we have covered exhaustively all the possible bundles, every basket that satisfies the budget constraint is a solution when $\frac{p_1}{p_2} = \frac{\beta_1}{\beta_2}$.

**Figure 6.8.** *Optimal Consumption - Linear Function*

**(a)** Only Good 1 **(b)** Only Good 2 **(c)** Any bundle over the budget line




**Note**: The red line is the budget line. Blue lines are indifference curves. Moving in the north east direction, utility increases, so that indifference curves further away from the origin provide a greater utility. The red dot bundle provides the greatest utility and, in particular, it provides more utility than the green dot bundle.

Overall, we have that for a linear utility the Marshallian demands are:

$$x_1^*(p_1,p_2,Y) := \begin{cases} \frac{Y}{p_1} & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ 0 & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \left[0, \frac{Y}{p_1}\right] & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases} \text{ and } x_2^*(p_1,p_2,Y) := \begin{cases} 0 & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ \frac{Y}{p_2} & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \frac{Y - p_1 x_1^*}{p_2} & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases}$$

Moreover, the indirect utility function is $U^*(p_1,p_2,Y) = \beta_1 x_1^*(p_1,p_2,Y) + \beta_2 x_2^*(p_1,p_2,Y)$. Like in the case of the sup function, this can be written as

$$U^*(p_1,p_2,Y) = Y \sup\left\{\frac{\beta_1}{p_1}, \frac{\beta_2}{p_2}\right\}.$$

### 6.4.2 EMP

To obtain the solution of the EMP, we use duality. The expenditure function is:

$$E^*(p_1,p_2,U_0) = \frac{U_0}{\sup\left\{\frac{\beta_1}{p_1}, \frac{\beta_2}{p_2}\right\}}$$

We know the indirect utility function is $U^*(p_1,p_2,Y) = Y \sup\left\{\frac{\beta_1}{p_1}, \frac{\beta_2}{p_2}\right\}$. And by duality, if $U_0 = U^*$ then $E^* = Y$. Therefore, plugging in those values in the indirect utility function, we get that $U_0 = E^* \sup\left\{\frac{\beta_1}{p_1}, \frac{\beta_2}{p_2}\right\}$ which, by isolating $E^*$, provides the result.

In turn, the Hicksian demands are

$$h_1^*(p_1,p_2,U_0) := \begin{cases} \frac{U_0}{p_1} & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ 0 & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \left[0, \frac{U_0}{\beta_1}\right] & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases} \text{ and } h_2^*(p_1,p_2,U_0) := \begin{cases} 0 & \text{if } \frac{p_1}{p_2} < \frac{\beta_1}{\beta_2} \\ \frac{U_0}{p_2} & \text{if } \frac{p_1}{p_2} > \frac{\beta_1}{\beta_2} \\ \frac{U_0}{\beta_2} - \frac{p_1}{p_2} x_1^* & \text{if } \frac{p_1}{p_2} = \frac{\beta_1}{\beta_2} \end{cases}$$

The Hicksian demands for the cases $\frac{p_1}{p_2} > \frac{\beta_1}{\beta_2}$ and $\frac{p_1}{p_2} < \frac{\beta_1}{\beta_2}$ are the same as the sup function. So, let's consider the case $\frac{\beta_1}{p_1} = \frac{\beta_2}{p_2}$.

The Marshallian demand for good 1 is $x_1^* \in \left[0, \frac{Y}{p_1}\right]$. Thus, $h_1^* \in \left[0, \frac{E^*(p_1,p_2,U_0)}{p_1}\right]$ or, what is same, $h_1^* \in \left[0, \frac{U_0}{\beta_1}\right]$ since $E^*(p_1,p_2,U_0) = \frac{p_1}{\beta_1} U_0$ at that point.[a] Regarding good 2, for a given $x_1^*$, the Marshallian demand is $x_2^* = \frac{Y - p_1 x_1^*}{p_2}$ and so $h_2^* = \frac{E^*(p_1,p_2,U_0) - p_1 x_1^*}{p_2}$. Given $E^*(p_1,p_2,U_0) = \frac{p_2}{\beta_2} U_0$ we can express it as $h_2^* = \frac{U_0}{\beta_2} - \frac{p_1}{p_2} x_1^*$.[b]

[a] Notice that, since $E^*(p_1,p_2,U_0) = \frac{p_1}{\beta_1} U_0 = \frac{p_2}{\beta_2} U_0$ when $\frac{\beta_1}{p_1} = \frac{\beta_2}{p_2}$, we could have also expressed the Hicksian demand of good 1 as $h_1^* \in \left[0, \frac{p_2}{p_1} \frac{U_0}{\beta_2}\right]$.

[b] Like in the previous footnote, since it is also true that $E^*(p_1,p_2,U_0) = \frac{p_1}{\beta_1} U_0$, we could also

have expressed it as $h_2^* = \frac{p_1}{p_2}\left[\frac{U_0}{\beta_1} - x_1^*\right]$.

### 6.4.3 Comparison with the Max Function

**Both the linear and the max utility have similar solutions to the UMP and EMP, except for the case of indifference between the corner solutions**. In that case, while a consumer with linear utility function is completely indifferent to any mixture of goods, a consumer with a max utility function would only choose one of the corner solutions.

Although it seems a pretty small difference, it makes a huge difference conceptually. Especially, in terms of the scenarios we could envision in one or the other case. For the max function, we provided the example of consuming tea or coffee. However, choosing a linear utility function to represent that example would be incorrect. If that were her preferences, she would be indifferent between drinking a cup of either tea or coffee, or simultaneously drinking tea and coffee.

So, when is the linear utility an accurate description of the situation? Let's consider the following example. A couple of years ago, Coke released a campaign called "share a Coke." The company was selling bottles of Coke with labels that included names (see Figure 13.1). The goal was to encourage customers to buy a coke with a friend's name and share it with him. The price of the bottles with either Coke's logo or names had the same price.

Suppose a consumer doing grocery shopping. Assume that she decides to buy two bottles of Coke for herself. Let the bottles of Coke with names on the label be good 1, and those with Coke's logo be good 2. Since the cokes are for herself, it is reasonable to think that she is indifferent between buying Cokes with or without names on their labels. Thus, since the prices of each type of bottle are the same, it is optimal to buy two units of good 1, two units of good 2, or even one unit of each. In other words, assuming that the consumer has a linear utility function is a reasonable assumption.

**Figure 6.9.** *Campaign Share a Coke*

**(a)** Standard Label **(b)** Names on Labels


On the other hand, if we model her behavior through a max utility function, we would be capturing a different decision process. In fact, it would be an implausible one for the example considered. To see this, imagine that she still wants to buy two bottles of coke, but there is only one bottle of each type available. With a max function, she only derives utility by consuming one of the goods. Thus, she would either buy one bottle with a name *or* one bottle with Coke's logo. But she would never buy both at the same time—with a max utility function she would only derive utility from one type of bottle.

**Remark** *One important caveat is in order. I have taken the example of Coke because it is quite intuitive. Furthermore, I considered that $p_1 = p_2$, along with a utility function having $\beta_1 = \beta_2$, since the consumer was indifferent between one good or the other. But note that the case of perfect substitutes is more general. Its relevant feature is that, even if $\beta_1 > \beta_2$, we can find prices $p_1$ and $p_2$ such that the consumer is indifferent between the corner solutions.*

## 6.5 Exercises

[1] John is a big fan of sandwiches, but he only likes plain sandwiches consisting of two slices of bread with one slice of cheddar cheese. He only enjoys that particular combination of ingredients. If, for any reason, a restaurant gives him more slices of cheese without additional bread (or more bread without cheese), he throws the cheese away. Of course, if they gave him one additional slice of cheddar joint with two slices of bread, he would prepare another sandwich and eat it.

Denote $x_1$ the slices of bread demanded and $x_2$ the slices of cheese. Suppose John has income $Y$ and the price of each good is $p_1$ and $p_2$ at his nearest supermarket.

(a) Establish a utility function that represents John's preferences. Justify your choice. Can you provide another utility function that represents his preferences?

(b) Determine his Marshallian demands and the indirect utility function.

(c) Suppose that $p_1 = p_2 = 1$ CAD. Suppose that the supermarket offers him a discount in one of the goods he consumes. The discount consists of 50 cents per unit of the good. He can decide on which good the discount applies. What good would he choose? Justify your choice.

(d) Suppose that John's income is $Y := 1000$. Starting from a situation with $p_1 = p_2 = 1$ CAD, suppose the cheese price increases to 1.50 CADs.

i. Determine the variation in each good's demand. Then, decompose the change in cheese demanded into the Slutsky substitution effect and income effect.

ii. If all your calculations are right, each effect has a peculiarity. Indicate what this is and why arises.

[2] After visiting a slaughterhouse, Janet has become vegan. However, she's a really peculiar type of vegan: she hates vegetables, except lettuce (good 1) and tomatoes

(good 2). That's all she consumes each week, and she enjoys both vegetables mixed in a salad. However, she has a stronger preference for tomatoes: she would only accept to exchange 1 kilogram of lettuce for 2 kilograms of tomatoes, irrespective of how much she's been consuming during the week.

(a) Establish a utility function that represents her utility.

(b) Suppose that each good's price is 1 CAD per kilogram. Establish Janet's Marshallian demands

(c) Suppose now that the price of tomatoes is 2 CAD per kilogram, while the price of lettuce is 1 CAD of per kilogram. Establish Janet's Marshallian demands.

**Some Answer Keys**:

1a) Leontief utility function. In fact, there are infinite possible utility functions (why?), 2a) Linear utility function.

# Lecture Note 7

# Welfare

## 7.1 Introduction

One feature of consumer theory is that any monotone transformation of a utility function still represents the same preferences. Formally, it means that utility functions are ordinal, but not cardinal: we can compare two bundles and say which one the consumer prefers, but we cannot quantify the magnitude in which she prefers a bundle. Put it differently, we can establish what the consumer prefers and hence chooses, but not the intensity in which she enjoys a particular bundle.

In this lecture, we provide two measures that quantify the intensity in which a consumer is better off. These measures overcome the issue of having an ordinal utility by using a metric based on monetary units. Throughout the presentation, we compare two scenarios. The first one is the initial situation, and we refer to it as the *status quo*. The second scenario considers that the price of one good increases relative to the status quo.

The first welfare measure we present is known as the Compensating Variation (CV). It measures the monetary compensation that allows the consumer to achieve the initial utility at the new prices. The second welfare measure is the Equivalence Variation (EV). It provides the level of income that provides the same utility as the initial situation but at the new prices. We will show that, actually, the computation of the EV and CV requires computing the minimum expenditure function of the EMP. The only difference is which utility we take as a base for the analysis.

## 7.2 Definitions of EV and CV

We denote the variables of the status quo with a prime, and the variables after the price change with a double prime. Specifically, we consider an increase in the price of good 1. Moreover, we suppose that the income and price of good 2 are the same in both situations. Formally, $p_1'' > p_1'$, and respectively denote the income and price of good by $Y$ and $p_2$.

Since we are considering a price increase, the utilities of the bundles satisfy that

$U'' < U'$. As we show below, each welfare measure we study requires establishing the minimum expenditure that would be necessary to achieve either $U''$ (the EV measure) or $U'$ (the CV measure).

### 7.2.1 Equivalent Variation (EV)

Formally, the EV is defined as

$$EV := E^*\left(p_1', p_2, U''\right) - E^*\left(p_1'', p_2, U''\right).$$

Notice that, by using duality, the minimum expenditure to achieve the indirect utility $U''$ at prices $(p_1'', p_2)$ has to satisfy $E^*\left(p_1'', p_2, U''\right) = Y$. Hence,

$$EV := E^*\left(p_1', p_2, U''\right) - Y.$$

The intuition of the EV is that if the prices had remained at $(p_1', p_2)$, the consumer would have needed less income to achieve utility $U''$. Thus, facing a reduction of income equal to $EV$ is equivalent in terms of welfare to facing the new prices and perceiving income $Y$. A consumer would claim that "facing these new prices is like if the prices had not changed, but my income had been reduced in an amount $EV$."

### 7.2.2 Compensating Variation (CV)

Mathematically, the CV is defined as

$$CV := E^*\left(p_1', p_2, U'\right) - E^*\left(p_1'', p_2, U'\right),$$

and, by using duality, we get $E^*\left(p_1', p_2, U'\right) = Y$ and so

$$CV := Y - E^*\left(p_1'', p_2, U'\right).$$

The CV provides the additional income the consumer needs at the new prices to obtain the same utility she was having before the increase in price.

**Remark** *Some authors define EV and CV differently. These definitions differ* *regarding the base utility that they use. What is important is that you understand what utility you should be plugging in into the expenditure function, depending on to the analysis you want to carry out.*

### 7.2.3 Calculating the EV and CV

The EV and CV can be computed by using the definitions we just provided. However, it is possible to compute in an alternative way, which could be easier depending on the context.

To see this, remember that the Hicksian demand of good 1 can be obtained by $h_1^*(p_1, p_2, U_0) = \frac{\partial E^*(p_1,p_2,U_0)}{\partial p_1}$ due to Shepard's Lemma. Moreover, the EV and CV are, ultimately, defined as the variations in the minimum expenditure taking a different level of utility as a base. Thus, by the fundamental theorem of the calculus,

$$EV := -\int_{p_1'}^{p_1''} h_1^*(p_1, p_2, U'')\, \mathrm{d}p_1,$$
$$CV := -\int_{p_1'}^{p_1''} h_1^*(p_1, p_2, U')\, \mathrm{d}p_1.$$

There are different ways to obtain the expression. Let's consider the case of EV since the derivations are the same for each welfare measure. The one that could be easier for you to understand is that, since $h_1^*(p_1, p_2, U_0) = \frac{\partial E^*(p_1,p_2,U_0)}{\partial p_1}$, then we can express $h_1^*(p_1, p_2, U_0)\, \mathrm{d}p_1 = \mathrm{d}E^*(p_1, p_2, U_0)$ and integrate the expression for the two prices. However, this requires that someone makes you notice that Shepard's Lemma is somehow related to EV.
One way in which we do not need a hint is just knowing that $CV := E^*(p_1', p_2, U'') - E^*(p_1'', p_2, U'')$ and then we can apply what is known as the Second Fundamental Theorem of the Calculus. Just in case you have not seen this before, the First and Second theorems are related to the conditions under which the integral and derivative are inverse operations.
In our case, it states that $E^*(p_1', p_2, U'') - E^*(p_1'', p_2, U'') = \int_{p_1''}^{p_1'} \frac{\partial E^*(p_1,p_2,U'')}{\partial p_1}\, \mathrm{d}p_1$. Hence, using Shepard's Lemma, we can reexpress the derivative within the integral by the Hicksian demand.

## 7.3 Relation between EV and CV

The two welfare measures provided do not necessarily coincide. In fact, they only do so under a quasilinear utility function, as we show below. Why do they differ? The reason is that both measures use different utility functions as a base, which in turn results in a different valuation of good 1 by the consumer.

To illustrate this, consider pronounced increases in the prices of some goods, such that a rich person ends up becoming poor. Then, the value of one additional dollar before the changes in prices (i.e., when the consumer was rich) and after (i.e., when the consumer is poor) would not be the same. Consequently, the minimum expenditure to get one more unit of utility would neither be the same, reflecting that poor people have a higher marginal utility of income.

Taking this into account, we proceed to characterize the differences between the EV and CV. In particular, we establish when one measure is greater than the other.

**Result 7.1** *For $EV$ and $CV$, we can derive two conclusions about their magnitudes:*

*[1] The* signs *$EV$ and $CV$ coincide.*

*[2] The difference between their magnitudes depends on the income effect that the good displays:*

- *if the good is normal: $CV > EV$.*
- *if the good is inferior: $CV < EV$.*
- *if the good has zero income effect: $CV = EV$.*

The case of zero income effect arises when the utility function is quasilinear and there is consumption of both goods. Then, we conclude that the EV and CV coincide under a quasilinear utility.

## 7.4 An Example

To illustrate the computation of welfare measures, let's consider a quasilinear utility function. Specifically,

$$U(x_1, x_2) := \ln\left(x_1 + \frac{1}{2}\right) + x_2.$$

We suppose that income is high enough and $p \leq 2$, which ensures that both goods are consumed in equilibrium. To simplify the calculations, we also assume that $p_2 := 1$.

Incorporating that $p_2 := 1$, the EMP is given by

$$\min_{x_1,x_2} E = p_1 x_1 + x_2 \text{ subject to } U_0 = \ln\left(x_1 + \frac{1}{2}\right) + x_2$$

which provides the following Hicksian demands and minimum expenditure function:

$$h_1^*(p_1) := \frac{1}{p_1} - \frac{1}{2},$$
$$h_2^*(p_1) := U_0 + \ln(p_1),$$
$$E^*(p_1, U_0) := 1 + U_0 + \ln(p_1) - \frac{p_1}{2}.$$

To solve the EMP for a quasilinear utility function, it is easier to plug in the constraint into the objective function. Specifically, the constraint can be rewritten as $x_2 = U_0 - \ln\left(x_1 + \frac{1}{2}\right)$, and substituting it into the expenditure function, the optimization problem is $\min_{x_1,x_2} E = p_1 x_1 + \left(U_0 - \ln\left(x_1 + \frac{1}{2}\right)\right)$. The FOC gives $p_1 - \frac{1}{x_1 + \frac{1}{2}} = 0$, which provides $h_1^*(p_1) := \frac{1}{p_1} - \frac{1}{2}$. By using the constraint equation, we can recover the Hicksian demand for good 2. Since $h_2^* = U_0 - \ln\left(\frac{1}{p_1} - \frac{1}{2} + \frac{1}{2}\right)$ and noticing that by property of the logarithm $\ln\left(\frac{1}{p_1}\right) = -\ln(p_1)$, then $h_2^*(p_1, U_0) := U_0 + \ln(p_1)$. Regarding the minimum expenditure function, it is given by $E^*(p_1, U_0) := p_1\left(\frac{1}{p_1} - \frac{1}{2}\right) + U_0 + \ln(p_1)$, or just $E^*(p_1, U_0) = 1 - \frac{p_1}{2} + U_0 + \ln p_1$.

We also need to solve the UMP, since computing welfare requires knowing the indirect utility function in each scenario. The outcomes of the UMP are

$$x_1^*(p_1) := \frac{1}{p_1} - \frac{1}{2},$$
$$x_2^*(p_1, Y) := Y - 1 + \frac{p_1}{2},$$
$$U^*(p_1, Y) := Y - 1 - \ln(p_1) + \frac{p_1}{2}.$$

Keep in mind that, for the quasilinear case, the Marshallian and Hicksian demands for good 1 are identical. Therefore, $x_1^*(p_1) = \frac{1}{p_1} - \frac{1}{2}$. By using the budget constraint, we can recover the Marshallian demand for good 2, so that $x_2^*(p_1, Y) = Y - p_1 x_1^*(p_1)$, which provides $x_2^*(p_1, Y) = Y - \left(1 - \frac{p_1}{2}\right)$. Finally, by duality and using that $E^*(p_1, U_0) := 1 + U_0 + \ln(p_1) - \frac{p_1}{2}$, we have that $Y = 1 + U^*(p_1, Y) + \ln(p_1) - \frac{p_1}{2}$, which determines that $U^*(p_1, Y) := Y - 1 - \ln(p_1) + \frac{p_1}{2}$.

Let's compare two situations: one with $p_1' := 1$ and another with $p_1'' := 2$. While $x_1' = \frac{1}{2}$, notice that $x_1'' = 0$. In fact, $p_1'' := 2$ represents the minimum price that makes the consumer demand zero quantities of good 1. Thus, welfare in this scenario quantifies the consumer's well being, relative to a scenario where she does not consume the good at all. The indirect utility function in each case is:

$$U' = Y - \frac{1}{2} \text{ and } U'' = Y - \ln(2).$$

The EV takes the second situation as the utility base and therefore,

$$\begin{aligned} EV &= E^*(p_1'', U'') - E^*(p_1', U''), \\ &= Y - \left[Y - \ln(2) + \frac{1}{2}\right] = \ln(2) - \frac{1}{2}, \end{aligned}$$

Alternatively, we can calculate the EV by integrating the Hicksian demand:

$$\begin{aligned} EV &= \int_{p_1'}^{p_1''} h_1^*(p_1)\, \mathrm{d}p_1, \\ &= \int_1^2 \left(\frac{1}{p_1} - \frac{1}{2}\right) \mathrm{d}p_1 = \left[\ln(p_1) - \frac{p_1}{2}\right]_1^2, \\ &= [\ln(2) - 1] - \left(-\frac{1}{2}\right) = \ln(2) - \frac{1}{2}. \end{aligned}$$

As for the CV, we know that it equals the EV by property of the quasilinear utility function. Nonetheless, let's compute it to show how this is done. Using the utility in the status quo as the base,

$$\begin{aligned} CV &= E^*(p_1'', U') - E^*(p_1', U'), \\ &= \left[Y - \frac{1}{2} - \ln\left(\frac{1}{2}\right)\right] - Y = \ln(2) - \frac{1}{2}. \end{aligned}$$

Alternatively, we could compute the CV by integrating the Hicksian demand. This can be done by following the same steps as we did for the EV.

## 7.5 Exercises

[1] Throughout this exercise, suppose that all monetary values are measured in USD. Laura, an Economics professor living in the States, earns $Y^{USA} := 1000$. She's a big fan of beer (good 1), which has a price $p_1^{USA} := 1$ in the US. Suppose that good 2 has $p_2^{USA} := 1$ and comprises the rest of the goods she consumes. She has received an offer to work at a university in Canada. They offer her exactly the same salary, where also $p_2^C := 1$ but the beer price is $p_1^C := 2$.

Assuming that the professor's utility function is $U(x_1, x_2) := \frac{1}{2}\ln(x_1) + \frac{1}{2}\ln(x_2)$, answer the following.

(a) The utility function reflects that the income proportion she spends on beer could be considered representative of a beer fan. What is that proportion?

(b) Since the salary in Canada is the same but prices are higher, she does not accept the offer and engages in a negotiation with the Canadian university. Someone has overheard he asked for an increase in the salary of at least $\delta$ CAD to accept the offer. The person couldn't clearly listen to what $\delta$ is. Calculate this value (hint: if you want to reduce the burden of calculations, start from $U^*$ and recover $E^*$ by duality. You can use that $U^*(p_1, Y) := \ln Y - \ln(\sqrt{p_1}) + \kappa$ where $\kappa := -\ln(2)$ is a constant).

(c) The Canadian Economics Department is constrained by the money allocated by the Dean. Thus, the salary of 1000 USD is a take-it-or-leave-it offer. The professor's family wants a fresh start and convinces her to accept the offer anyway. While living there, the professor complains that paying the Canadian beer price is like if she had never moved out from the US, but the American university had reduced her salary in $\tau$ USD. Calculate $\tau$.

**Answer Keys for some of the exercises**:

1b) $1000\left(\sqrt{2}-1\right)$ which is approx 414.21, 1c) $1000\left(1-\frac{1}{\sqrt{2}}\right)$ which is approx 292.9.

# Lecture Note 8

# Cost Minimization

## 8.1 Introduction

In this note, we focus on the cost minimization problem of a firm. Mathematically, the problem is equivalent to the EMP of consumer theory, and so its treatment will be relatively brief. Chiefly, the focus will be on the mathematical concepts that differ from consumer theory. In particular, we present production functions that are homogeneous, and define the concepts of returns to scale and economies of scale. We conclude by showing a tractable way to introduce increasing returns to scale, through the existence of fixed costs.

## 8.2 Cost Minimization Problem (CMP)

Suppose a firm producing a good with a technology that uses two factors, labor and capital. We refer to a production function or technology indistinctly, and define it as a function $(l, k) \mapsto f(l, k)$, where $l \in \mathbb{R}_+$ and $k \in \mathbb{R}_+$ are respectively the amount of labor and capital used by the firm.

We suppose that $f$ satisfies several properties. First, $f$ is supposed to be increasing, so that $\frac{\partial f(l,k)}{\partial l} > 0$ and $\frac{\partial f(l,k)}{\partial k} > 0$. Moreover, $f$ is assumed strictly concave[1], implying that each factor has either constant or decreasing marginal returns. Formally, $\frac{\partial^2 f(l,k)}{\partial l^2} < 0$, $\frac{\partial^2 f(l,k)}{\partial k^2} < 0$, and $\det J_f > 0$, where $J_f$ is the Jacobian of the production function.

These assumptions are similar to those defining a well-behaved utility function, and are not necessarily satisfied wwith standard functional forms. Their goal is to ensure that there is a unique and interior equilibrium.

> The function $f$ represents production plans that are technologically efficient, rather than just possible. This means that, even when the firm has different technologies available, $f$ provides the most efficient way to use inputs $(l, k)$. For example, suppose that technique 1 produces $f^1(1, 1) = 4$ and technique 2 $f^2(1, 1) = 3$. The production function $f$ discards technique 2 since it is less efficient at the point where $l = 1$ and $k = 1$. Thus, $f(1, 1) = 4$, since technique 1 gives the maximum output, given the possible technologies.

[1] Just in case you are curious, in fact, we could simply assume quasiconcavity. Below, we are going to assume that functions display nonincreasing returns to scale and, in that case, quasiconcavity and concavity are equivalent.

Notice that efficiency is only necessary to minimize costs, but not sufficient. For instance, to determine whether $(l = 2, k = 1)$ or $(l = 1, k = 2)$ minimizes costs, we need to additionally know the price of the inputs. Efficiency only establishes that $f(2,1)$ and $f(1,2)$ are chosen among the technologically efficient techniques that are available to the firm.

The CMP identifies the combinations of $(l,k)$ that minimize the cost of producing $\overline{q}$ units of the good, where $\overline{q}$ is treated as a parameter. To understand the role of the CMP, we can think of a firm as solving two different optimization problems. First, it has to choose the optimal quantities to deliver to the market.[2] This is usually expressed as a profit-maximization problem. Second, it has to decide how to produce those quantities that maximize profits. In this decision process, the CMP constitutes the second step of the analysis, and we relegate the analysis of the first step to subsequent lectures.

Formally, the CMP is given by

$$\min_{l,j} C = lw + rk$$

$$\text{subject to } f(l,k) = \overline{q},$$

where $w$ and $r$ are the wages and remuneration to the capital, respectively. These variables are exogenously given, reflecting that firms do not have monopsony power—the firm is not so big that it can influence the price at which it buys its inputs.

**Remark** *Notice that by appropriately relabeling variables and functions,* ***the CMP is exactly the same as the EMP of consumer theory****. The function $f$ plays the same role as $U$, $\overline{q}$ as $U_0$, and $C$ as $E$.*

The assumptions on $f$ ensure that the CMP is well-behaved, and hence we can use the Lagrange technique to solve it. The Lagrangian is

$$\mathscr{L}(l,r,\lambda;w,r,\overline{q}) := lw + rk + \lambda\left[\overline{q} - f(k,l)\right].$$

[2]Alternatively, a firm could choose the optimal price, which indirectly defines the quantities to deliver according to its demand.

and the FOCs are:

$$\text{FOC:} \begin{cases} \frac{\partial \mathscr{L}(\cdot)}{\partial l} = w - \lambda f_l'(l,k) = 0 \\ \frac{\partial \mathscr{L}(\cdot)}{\partial x_2} = r - \lambda f_k'(l,k) = 0 \\ \frac{\partial \mathscr{L}(\cdot)}{\partial \lambda} = \overline{q} - f(l,k) = 0 \end{cases}$$

The FOCs determines optimal choices $l^* = l(w,r,\overline{q})$ and $k^* = k(w,r,\overline{q})$, along with a minimum cost function given by

$$C^*(w,r,\overline{q}) := wl(w,r,\overline{q}) + rk(w,r,\overline{q}).$$

To characterize the solution, we use the first two equations from the FOC, which determine that

$$\frac{w}{r} = \frac{f_l'(l,k)}{f_k'(l,k)}. \tag{8.1}$$

Equation (8.1) establishes that the optimal input choices depend on $\frac{w}{r}$, rather than $(w,r)$. This is the same as in consumer theory, where only the relative prices matter for choices, the absolute prices do not. Likewise, we can obtain a similar tangent condition as in consumer theory, between the slope of the indifference curves and the budget constraint. The only difference is that now we instead use isoquants (combinations of inputs that produce the same level of output) and the cost function. In firm theory, the expression $\frac{f_l'(l,k)}{f_k'(l,k)}$ is usually referred to as the "technical rate of substitution", so that (8.1) indicates that the TRS equals the relative input prices.

We can also provided an alternative interpretation of the optimality condition, (8.1). To do this, we can multiply and divide the left-hand side of (8.1) by the price of the good, $p$. Rearranging the terms, (8.1) is equivalent to

$$\frac{p\, f_k'(l,k)}{r} = \frac{p\, f_l'(l,k)}{w}. \tag{8.2}$$

The term $p\, f_l'(l,k)$ indicates the revenue garnered by hiring one more worker: it provides an additional output $f_l'(l,k)$, with a market value of $p\, f_l'(l,k)$. Since this term is divided by the wage, the right-hand side provides the revenue-cost relation of hiring one additional worker. A similar interpretation can be established for the left-hand side, but in terms of capital.

Consequently, (8.2) shows that the optimal input choices need to equalize the ratio revenue-cost for each input. On the contrary, suppose that there differences between these rations, such that, for instance, $\frac{p f_k'(l,k)}{r} < \frac{p f_l'(l,k)}{w}$. Then, it is optimal to buy less capital and hire more workers instead, since it would allow the firm to save costs.

### 8.2.1 Marginal Costs

Using the Envelope Theorem, we can derive some additional results. They provide in particular an interpretation of the Lagrange multiplier, $\lambda$.

**Procedure to apply the Envelope Theorem**

Step 1. Construct the Lagrangian: $\mathscr{L}(l,k,\lambda;w,r,\overline{q}) := lw + rk + \lambda[\overline{q} - f(l,k)]$

Step 2. Take derivatives of the Lagrangian with respect to each parameter of interest, without embedding the optimal solutions:

- $\frac{\partial \mathscr{L}(l,k,\lambda;w,r,\overline{q})}{\partial \overline{q}} = \lambda$.
- $\frac{\partial \mathscr{L}(l,k,\lambda;w,r,\overline{q})}{\partial w} = l$ and $\frac{\partial \mathscr{L}(l,k,\lambda;w,r,\overline{q})}{\partial r} = k$.

Step 3. Evaluate each derivative at the optimal values to obtain $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}}$ and $\frac{\partial C^*(w,r,\overline{q})}{\partial w}$:

- $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} = \lambda^*(w,r,\overline{q})$.
- $\frac{\partial C^*(w,r,\overline{q})}{\partial w} = l^*(w,r,\overline{q})$ and $\frac{\partial C^*(w,r,\overline{q})}{\partial r} = k^*(w,r,\overline{q})$.

We can derive two conclusions from these results. First, if we know the minimum cost function, we can obtain the optimal demand of factors: they are given by the partial derivatives of $C^*$ with respect to the input prices. Furthermore, $\lambda^*$ is the **marginal cost**. It says how the minimum cost varies when the firm produces one additional unit of the good.

**Remark** *Strictly speaking, the marginal cost is neither the cost of producing one additional unit of the good, nor the cost of producing the last unit of the good. The expression* $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}}$ *indicates the* total impact on *costs by producing one additional unit of the good. Thus, it takes into account the impact on all the units produced.*

For future references, we denote marginal costs as $MC(w,r,\overline{q}) := \lambda^*(w,r,\overline{q})$, and the average costs by $AC(w,r,\overline{q}) := \frac{C^*(w,r,\overline{q})}{\overline{q}}$.

### 8.2.2 Comparative Statics

We analyze how variations in the parameter affect a firm's optimal choices of inputs. This is done by performing comparative statics for the CMP. We state the results without any formal proof. The derivations are identical to those of the EMP.

**Result 8.1** *The comparative statics of the CMP provide the following results:*

- $\frac{\partial l^*(w,r,\overline{q})}{\partial w} < 0$ and $\frac{\partial k^*(w,r,\overline{q})}{\partial r} < 0$.
- $\frac{\partial l^*(w,r,k)}{\partial r} > 0$ and $\frac{\partial k^*(w,r,\overline{q})}{\partial w} > 0$ (this is only for the two inputs case, with an ambiguous result for three or more inputs).
- $\frac{\partial l^*(w,r,\overline{q})}{\partial \overline{q}} \gtreqqless 0$ and $\frac{\partial k^*(w,r,\overline{q})}{\partial \overline{q}} \gtreqqless 0$.

## 8.3 Homogeneous Technologies

We study some technologies that are widely used in the literature. Our focus is on the additional properties that they satisfy, since specific functional forms put additional structure to the CMP. With this aim, we start by defining homogeneous functions.

**Definition 8.1:** *A function* $f : X_1 \times X_2 \to \mathbb{R}$ *is* ***homogeneous of degree h*** *with* $h \in \mathbb{R}$ *if, for any* $\alpha > 0$ *and* $(x_1, x_2) \in X_1 \times X_2$*, then* $f(\alpha x_1, \alpha x_2) = \alpha^h f(x_1, x_2)$*. This is denoted* $f \in H^h$*.*

Commonly, the degree of homogeneity satisfies$h \in \mathbb{Z}$ (i.e. it is an integer). Also, homogeneity of a function allows for $h$ taking a negative value. Notice $f$ need not be differentiable to define a homogeneous function. This is important when we analyze cases like the Leontief production function, which is not differentiable everywhere (in fact, the Leontief function is homogeneous of degree one).

Example

The Cobb Douglas production function, $f(l, k) := l^{\beta_l} k^{\beta_k}$, is homogeneous of degree $\beta_l + \beta_k$. This follows since, given $\alpha > 0$,

$$\begin{aligned} f(\alpha l, \alpha k) &= (\alpha l)^{\beta_l} (\alpha k)^{\beta_k}, \\ &= \alpha^{\beta_l + \beta_k} l^{\beta_l} k^{\beta_k}, \\ &= \alpha^{\beta_l + \beta_k} f(l, k), \end{aligned}$$

and so $h := \beta_l + \beta_k$ by definition of homogeneous functions.

To provide some intuition for this result, we can use the characterization of homogeneous functions for the one-variable case. Given a function of one variable, $g : \mathbb{R}_+ \to \mathbb{R}_+$, it can be proven that $g$ is homogeneous of degree $\alpha$ iff it is a power function, i.e. $g(x) := Ax^\alpha$ for some $A \in \mathbb{R}$. Although the result cannot be generalized to multi-dimensional domains, it helps us understand that there is relation between the concept of homogeneous functions and power functions.

### 8.3.1 Further Results (OPTIONAL)

Next, I add some concepts that are related to this topic, but we will be studied in this course. I state them just in case you stumble upon them in other courses.

It is customary to work with functions that are not homogeneous, but rather a monotone transformation of a homogeneous function. When this occurs, several properties holding for homogeneous functions are preserved. For example, optimal solutions are the same, thus inheriting the properties of the solutions for homogeneous functions (this is not true for the value function).

**Definition 8.2:** *Let the production function $f$ be strictly increasing. Then, we say that $f$ is* ***homothetic*** *when $f$ is a monotone transformation of a homogeneous function.*

Example

The Cobb Douglas production function expressed in logs is $f(l,k) := \beta_l \ln(l) + \beta_k \ln(k)$. It is not homogeneous of any degree, but it is a monotone transformation of a homogeneous function and hence homothetic.

Homogeneous functions that are differentiable satisfy an additional property known as Euler's theorem. In fact, this theorem holds as an "if and only if", thereby providing an alternative equivalent definition of homogeneous functions.

**Theorem 8.3.1: Euler's Theorem**

*Let $f : \mathbb{R}^2_+ \to \mathbb{R}$ with $f \in \mathcal{C}^1$. Then, $f \in H^h$ iff for any $x_1, x_2 > 0$, $\frac{\partial f(x_1,x_2)}{\partial x_1} x_1 + \frac{\partial f(x_1,x_2)}{\partial x_2} x_2 = h\, f(x_1, x_2)$.*

Finally, the following property can sometimes come in handy when you work with homogeneous functions.

**Theorem 8.3.2**

*If $f \in H^h$ then each partial derivative is homogeneous of degree $h-1$. Formally, $\frac{\partial f(x_1,x_2)}{\partial x_i} \in H^{h-1}$ for $i = 1, 2$.*

## 8.4 Returns to Scale (RS) and Economies of Scale (ES)

Given some technology $f$, let $f(\alpha; l, k) := f(\alpha l, \alpha k)$ for $\alpha > 0$. This function gives information about the effect on total output from increasing *all* inputs in a proportion $\alpha$. For instance, starting from input choices $(\bar{l}, \bar{k})$, a value $\alpha = 2$ means we are doubling the amount of inputs relative to $(\bar{l}, \bar{k})$.

**Definition 8.3:** *Define the* ***Elasticity of Production*** *at a point* $(\bar{l}, \bar{k})$ *by* $EP(\bar{l}, \bar{k}) := \left.\frac{\mathrm{d} \ln f(\alpha; \bar{l}, \bar{k})}{\mathrm{d} \ln \alpha}\right|_{\alpha=1}$. *We say that there are:*

[1] *Increasing RS (IRS) at* $(\bar{l}, \bar{k})$ *if* $EP(\bar{l}, \bar{k}) > 1$,

[2] *decreasing RS (DRS) at* $(\bar{l}, \bar{k})$ *if* $EP(\bar{l}, \bar{k}) < 1$,

[3] *constant RS (CRS) at* $(\bar{l}, \bar{k})$ *if* $EP(\bar{l}, \bar{k}) = 1$.

I have added a bar above in $(\bar{l}, \bar{k})$ to emphasize that $EP$ is a function and depends on the combination of inputs used. Additionally, the fact that it depends on $(\bar{l}, \bar{k})$ implies that $EP$ is a *local* measure of scale. We will see, nonetheless, that the value of $EP$ for homogeneous functions is the same for any $(\bar{l}, \bar{k})$.

Let's analyze in particular the case with $EP > 1$ and $EP = 1$. CRS, i.e. when $E = 1$, will be studied in more detail below. $EP > 1$ describes a technology such that varying the inputs $(\bar{l}, \bar{k})$ in $\alpha\%$ results in an increase in production higher than $\alpha\%$. Put it differently, there is a more than proportional variation in output relative to the variation in inputs.

There is an additional concept that is intimately related with $EP$, which is known as the **Elasticity of Scale**. This is defined in terms of the minimum cost function.

**Definition 8.4:** *Define* ***Elasticity of Scale*** *at a point* $(w, r, \bar{q})$ *by* $ES(w, r, \bar{q}) := \frac{\partial \ln C^*(w, r, \bar{q})}{\partial \ln \bar{q}}$. *We say there are:*

[1] *economies of scale (ES) at* $(w, r, \bar{q})$ *if* $ES(w, r, \bar{q}) < 1$,

[2] *diseconomies of scale at* $(w, r, \bar{q})$ *if* $ES(w, r, \bar{q}) > 1$.

The concept of ES can be alternatively characterized by its relation with costs.

**Result 8.2** $ES(w,r,\overline{q}) = \frac{MC(w,r,\overline{q})}{AC(w,r,\overline{q})}$.

By definition, $ES(w,r,\overline{q}) := \frac{\partial \ln C^*(w,r,\overline{q})}{\partial \ln \overline{q}}$ and, like any other elasticity, $\frac{\partial \ln C^*(w,r,\overline{q})}{\partial \ln \overline{q}} = \frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} \frac{\overline{q}}{C^*(w,r,\overline{q})}$ or just $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} \frac{1}{\frac{C^*(w,r,\overline{q})}{\overline{q}}}$ and the result follows.

The result implies that ES emerge when $\frac{MC(w,r,\overline{q})}{AC(w,r,\overline{q})} < 1$, and so when marginal costs are lower than the average costs. We conclude this part with an alternative characterization of ES. It states that ES arise when average costs are decreasing.

**Result 8.3** $\operatorname{sgn}\left(\frac{\partial AC(w,r,\overline{q})}{\partial \overline{q}}\right) = \operatorname{sgn}(ES - 1)$ *where* sgn *is the sign function.*

By definition, $AC(w,r,\overline{q}) := \frac{C^*(w,r,\overline{q})}{\overline{q}}$ and so $\frac{\partial AC(w,r,\overline{q})}{\partial \overline{q}} = \frac{MC(w,r,\overline{q})\overline{q} - C^*(w,r,\overline{q})}{(\overline{q})^2}$ and so $\frac{\partial AC(w,r,\overline{q})}{\partial \overline{q}} > 0$ iff $MC(w,r,\overline{q})\overline{q} - C^*(w,r,\overline{q}) > 0$ or, dividing by $\overline{q}$, iff $MC(w,r,\overline{q}) > \frac{C^*(w,r,\overline{q})}{\overline{q}} = AC(w,r,\overline{q})$. Since $ES(w,r,\overline{q}) = \frac{MC(w,r,\overline{q})}{AC(w,r,\overline{q})}$, the result follows.
$\frac{\partial AC(w,r,\overline{q})}{\partial \overline{q}}$
$ES(w,r,\overline{q}) := \frac{\partial \ln C^*(w,r,\overline{q})}{\partial \ln \overline{q}}$ and, like any other elasticity, $\frac{\partial \ln C^*(w,r,\overline{q})}{\partial \ln \overline{q}} = \frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} \frac{\overline{q}}{C^*(w,r,\overline{q})}$ or just $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} \frac{1}{\frac{C^*(w,r,\overline{q})}{\overline{q}}}$ and the result follows.

### 8.4.1 CRS and ES under Homogeneous Functions

The characterization of RS and ES is straightforward when the production function satisfies homogeneity: both concepts are completely identified by the function's degree of homogeneity.

**Result 8.4** *If the production function satisfies $f \in H^h$, then $EP\left(\overline{l},\overline{k}\right) = h$ for any $\left(\overline{l},\overline{k}\right)$.*

By definition $EP\left(\overline{l},\overline{k}\right) := \left.\frac{\mathrm{d} \ln f(\alpha;\overline{l},\overline{k})}{\mathrm{d} \ln \alpha}\right|_{\alpha=1}$, and it can be expressed as $\frac{\mathrm{d} f(\alpha l,\alpha k)}{\mathrm{d}\alpha} \frac{\alpha}{f(\alpha;\overline{l},\overline{k})}$ like any other elasticity. Moreover, by the fact that $f \in H^h$ then $f(\alpha l,\alpha k) = \alpha^h f(l,k)$ and so $\frac{\mathrm{d} f(\alpha l,\alpha k)}{\mathrm{d}\alpha} = h\alpha^{h-1} f(l,k)$. Therefore, $EP\left(\alpha;\overline{l},\overline{k}\right) = h\alpha^{h-1} f(l,k) \frac{\alpha}{f(\alpha l,\alpha k)} \Rightarrow EP\left(\alpha;\overline{l},\overline{k}\right) = h\alpha^h f(l,k) \frac{1}{f(\alpha l,\alpha k)} \Rightarrow EP\left(\alpha;\overline{l},\overline{k}\right) = h$.

**Result 8.5** *Suppose the production function satisfies $f \in H^h$. Then, $EP\left(\bar{l},\bar{k}\right) = h = \frac{1}{ES(w,r,\bar{q})}$ for any $\left(\bar{l},\bar{k}\right)$.*

There is an equivalence between $EP$ and $ES$ when $f \in H^h$ and we evaluate $EP$ at the optimal values.
$\frac{\mathrm{d}f(\alpha l,\alpha k)}{\mathrm{d}\alpha}\frac{\alpha}{f(\alpha;\bar{l},\bar{k})} = \left[\frac{\partial f(\alpha l,\alpha k)}{\partial(\alpha l)}l + \frac{\partial f(\alpha l,\alpha k)}{\partial(\alpha k)}k\right]\frac{\alpha}{f(\alpha;\bar{l},\bar{k})}$ and evaluating at $\alpha = 1$, we obtain that $\left.\frac{\mathrm{d}\ln f(\alpha l,\alpha k)}{\mathrm{d}\ln\alpha}\right\rfloor_{\alpha=1} = \left[\frac{\partial f(l,k)}{\partial l}l + \frac{\partial f(l,k)}{\partial k}k\right]\frac{1}{f(l,k)}$
Also, at the optimal inputs, the FOCs give $\frac{\partial f(l^*,k^*)}{\partial l}\lambda^* = w$ and $\frac{\partial f(l^*,k^*)}{\partial k}\lambda^* = r$. So,
$\left.\frac{\mathrm{d}\ln f(\alpha l,\alpha k)}{\mathrm{d}\ln\alpha}\right\rfloor_{\alpha=1} = \left[\frac{w}{\lambda^*}l^* + \frac{r}{\lambda^*}k^*\right]\frac{1}{f(l,k)}$
$\Rightarrow \left.\frac{\mathrm{d}\ln f(\alpha l,\alpha k)}{\mathrm{d}\ln\alpha}\right\rfloor_{\alpha=1} = \frac{wl^*+rk^*}{\lambda^* f(l,k)}$
since we have that $\lambda^*$ is the marginal cost and $\frac{wl^*+rk^*}{f(l^*,k^*)}$ the average cost. Thus, since $f(l^*,k^*) = \bar{q}$, then
$\Rightarrow \left.\frac{\mathrm{d}\ln f(\alpha l,\alpha k)}{\mathrm{d}\ln\alpha}\right\rfloor_{\alpha=1} = \frac{AC(w,r,\bar{q})}{MC(w,r,\bar{q})} = \frac{1}{ES(w,r,\bar{q})}$.

Using these results, we conclude that $ES = 1$ when $h = 1$, which in turn implies that the technology exhibits CRS. This is why researchers usually refer to a unitary elasticity of scale and CRS as if they were synonyms.

## 8.5 Increasing Returns to Scale (IRS)

We have stated that any technology exhibits IRS if $ES > 1$. For the particular case of homogeneous technologies, this occurs when the degree of homogeneity is greater than one. A Cobb Douglas with $\beta_l + \beta_k > 1$ constitutes an example. However, depending on the type of analysis we perform, assuming this type of technology could make undesirable results emerge.

For example, consider a firm with a Cobb Douglas exhibiting IRS. Suppose that it is deciding its production and does so with the goal of maximizing profits. A Cobb Douglas with IRS implies that the more a firm produces, the lower its marginal costs. Hence, the firm would always have incentives to increase its production, determining that the only possible solution is to produce infinite output. When we translate this type of technology to the characterization of an industry, it would entail that only one firm would operate in the market.

However, a technology with decreasing marginal costs is only one type of IRS technology. Another case is when a firm has fixed costs of production. Fixed costs are, by definition, costs that are independent of the quantity produced. An example of this is the rent paid by a clothing store to sell shirts—this cost is independent of the level of production, and the store cannot make sales if it does not incur it.

To formalize this, suppose a technology that uses labor as the only production factor. Suppose that the firm requires hiring a fixed number of workers $\delta$ to produce positive quantities. After this, one unit of labor produces one unit of output. The production function of this case is

$$f(l) := \begin{cases} l-\delta & \text{if } l \geq \delta \\ 0 & \text{otherwise} \end{cases}.$$

The technology reflects that any quantity of labor $l < \delta$ results in zero quantities, since the firm cannot cover the minimum labor required to produce. However, one unit of labor produces one unit of output as soon as $l > \delta$, which is reflected in $\frac{\mathrm{d}f(l)}{\mathrm{d}l} = 1$ if $l \geq \delta$.

The function $f$ is not homogeneous, and so we need to use the definition of IRS to check that the property holds. With this goal, we first show that $f$ is indeed not homogeneous.

The definition of homogeneity is $f(\alpha l) = \alpha f(l)$ for *any* $\alpha > 0$. This means that if we can find an $\alpha$ where this does not hold, then $f$ is not homogeneous. To see this, note that

$$f(\alpha l) = \begin{cases} \alpha l-\delta & \text{if } \alpha l \geq \delta \\ 0 & \text{otherwise} \end{cases}.$$

Take $l > \delta$ and choose $\alpha$ such that $\alpha l > \delta$. Then,

$$f(\alpha l) = \alpha l - \delta \neq \alpha(l-\delta) = \alpha f(l),$$

which proves our claim.

To show that there are IRS, we can show that $ES > 1$.[3] The CMP for $\overline{q} > 0$ requires that $l > \delta$, because otherwise the firm cannot produce positive quantities. Thus, the

[3] Although we have stated the relation for homogeneous technologies, this is actually true as long as the input choices are the solution to the CMP.

CMP is

$$\min_{l} C = wl$$
$$\text{subject to } \overline{q} = l - \delta,$$

which shows that the CMP is trivial when there is only one production factor—its constraint identifies the solution:

$$l^* (\overline{q}) = \delta + \overline{q},$$

and so the minimum cost for $\overline{q} > 0$ is

$$C^* (w, \overline{q}) := w (\delta + \overline{q}).$$

From this, we obtain that $MC(w) := w$ and $AC(w) = \frac{w(\delta+\overline{q})}{\overline{q}}$. And since $\frac{MC(w)}{AC(w)} = \frac{w\overline{q}}{w\delta+w\overline{q}} < 1$, then $MC < AC$ and so there are ES. We could have also obtained the same conclusion by using that $\frac{\partial AC(w)}{\partial \overline{q}} = -\frac{w\delta}{(\overline{q})^2} < 0$ and checking that average costs are decreasing.

Alternatively, we could have shown that there are IRS through the elasticity of production. Let $\alpha > 0$. Then,

$$f(l) := \begin{cases} \alpha l - \delta & \text{if } \alpha l \geq \delta \\ 0 & \text{otherwise} \end{cases},$$

with for $l \geq \delta$ implies

$$\frac{\partial f(\alpha l)}{\partial \alpha} \frac{\alpha}{f(\alpha l)} = l \frac{\alpha}{\alpha l - \delta} = \frac{1}{1 - \frac{\delta}{l\alpha}} > 1,$$

and so there are IRS.

## 8.6 Exercises

[1] Consider the baseline case of IRS with fixed costs, analyzed in the lecture note. A simplifying assumption I made was that one unit of labor produces one unit of output. Let's incorporate a parameter $\varphi$ that captures the productivity of workers. This parameter affects how easily they can produce the good. Nonetheless, $\varphi$ does not affect the labor necessary to produce positive quantities, which is still $\delta$ (i.e., it does not affect the fixed cost). Thus, the production function is

$$f(l) := \begin{cases} \varphi l - \delta & \text{if } l \geq \delta \\ 0 & \text{otherwise} \end{cases} .$$

(a) Show that my claim of $\varphi$ capturing productivity is in fact true. How can $\varphi$ be specifically interpreted? (that is, "read" the expression you have used to show that $\varphi$ is indeed a productivity term).

(b) Write down and solve the cost minimization problem for $\overline{q} > 0$.

(c) Do increases in productivity reduce total costs, marginal costs, and average costs, as we would expect?

(d) Show that the technology exhibits CRS when $\delta \to 0$. Interpret.

[2] Bart and Milhouse want to gather some money to watch the new Krusty's movie. With this goal, they start selling lemonades in the street by using Marge's recipe. This combines one lemon and a quarter liter of mineral water. The cost of one lemon and one liter of water is the same and equals 1 dollar.

(a) Establish the production function describing the technology for glasses of lemonade. Show it displays CRS.

(b) Bart and Milhouse treat the business quite seriously and behave as cost minimizers. What is the total cost of producing $\overline{q}$ glasses?

(c) What is the proportion of expenditure in lemons and water relative to total cost? What are these proportions if the prices of lemons increase to 2 dollars?

(d) Suppose you're a customer. Like any customer, you don't have information on the production technology they use. In particular, you don't know that they use Marge's recipe for the lemonades. While you're talking to Milhouse, he reveals that the expenditure proportion is the one computed in c). In addition, he tells you that those proportions are independent of how many glasses they produce and sell. Those are the only pieces of information you have about the technology they use.

i. Would you have concluded that the production function is necessarily the same as in a)? If your answer is no, mention one production function also consistent with the information at your disposal.

ii. Keep assuming what Milhouse told you about the cost shares. Last week, you bought one glass and noticed that the lemonade was watered down. Since the price of lemons has risen, you suspect that the mischievous Bart has used more water and less lemon per glass. In fact, you have noticed that every time the price of the lemons rises, the lemonade has more water than usual.

1. What kinds of production functions we saw in class are consistent with this information?

2. Suppose that you come back the following week, and complain to Milhouse that last week's lemonade had too much water. Milhouse shows you the receipts of mineral water and lemons bought, which indicate that each input's cost share has not changed. Is this enough argument to show that you're wrong? Explain your answer by using a production function.

[3] (***I'll solve this one in class***) Nike has been one of the pioneers in outsourcing production to the developing world, establishing factories all around Asia. However, things are changing nowadays: wages in Asia have been rising (although they

remain low). Due to this, Nike has started to look at cheaper production alternatives. The following excerpt is from an article of the Financial Times, touching upon the subject.

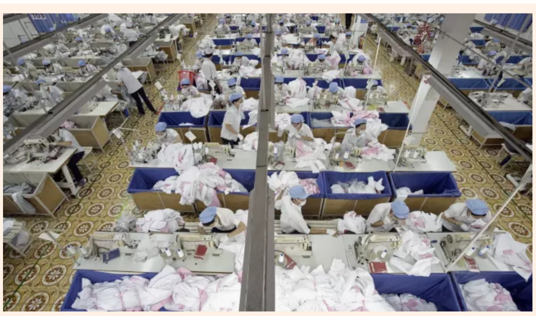
**Nike's focus on robotics threatens Asia's low-cost workforce**

"*The very-low labour costs in Asia are no longer that low unless you go to Africa or somewhere else . . . The pressure has been mounting for a long time to either move to a super low-cost place or automate more... That has come to a point where people are more seriously looking to automation.*"

By the title and the excerpt, we can envision two production modes.

**Figure 8.1.** *Production Modes*

**(a)** Mode 1



**(b)** Mode 2

The mode 1 (picture 8.1a) corresponds to a typical Nike's factory in Asia (in this case, Vietnam), where low-skilled workers manually produce shoes. This is the current predominant technology. The mode 2 (picture 8.1b) is the typical factory that Nike aims to have in the next years. It is capital intensive and needs of high-skilled workers.

In this exercise, we want to model Nike's production function for shoes. There are three factors: capital, low-skilled workers, and high-skilled workers. Let $K$, $L$ and $H$ be the units of each factor employed, with prices $r$, $w_L$ and $w_H$. We first model each mode of production separately, and then combine them into one production function.

(a) Mode 1 requires only low-skilled workers, with 10 workers producing one pair of shoes. Write down a function that captures this technique.

(b) Mode 2 consists of high-skilled workers designing and operating machines. The production of one pair of shoes requires one machine and two workers (one worker that designs the machine, and another that operates it). Write down a function that captures this technique.

(c) Establish Nike's production factory for shoes assuming that Nike can potentially produce with mode 1 or mode 2 technology.

(d) According to the article, Nike still uses the first mode of production since automation is an expensive alternative these days. Determine what condition $w_L$ has to satisfy to be consistent with Nike using the first mode to minimize costs.

# Lecture Note 9

# Constant Returns To Scale

## 9.1 Introduction

In these notes, we present typical production functions used in the literature. Our focus is on the implications that these technologies have when we analyze the CMP, with special emphasis on technologies exhibiting CRS. We study specific functional forms, including perfect complements, perfect substitutes, and the Cobb Douglas. We do this briefly, since the mathematical derivations are the same as in consumer theory.

We then introduce one new functional form not covered in consumer theory, known as the constant elasticity of substitution (CES) production function. This constitutes a generalization of the other production functions, since it converges to the cases of perfect complements, perfect substitutes, and Cobb Douglas, depending on the parametric version we use. Although you do not have to know how the solution of the CES is obtained (the algebra is quite messy), I want you to learn its properties and when its use is appropriate.

## 9.2 CRS with Homogeneous Functions

Consider a production function that is homogeneous of some degree. CRS arises for these types of technologies when they are in particular homogeneous of degree one (aka linear homogeneity). Formally, $f$ satisfies that $f(\alpha l, \alpha k) = \alpha f(l, k)$ for any $\alpha > 0$. Intuitively, it means that the output increases in a proportion $\alpha$ when all the inputs increase in a proportion $\alpha$.

> When we deal with a utility function, any monotonic transformation of it represents the same preferences. Consumer's ranking of products is all what we need to know and, so, the values attached to utility are meaningless.
> On the other hand, when we deal with production functions, quantities and costs have an explicit unit of measure. As a result, monotone transformations do not represent the same production function. This is why we defined CRS production functions as functions which are homogeneous of degree one instead of allowing for any monotone transformation and, hence, defining them as homothetic functions.

There are several properties that the solution to CMP satisfies when the technology exhibits CRS. The first property is related to the optimal demand of factors.

**1. Input demands are linear in $\overline{q}$.** Let each factor demand for one unit of product be $a_l^*(w,r) := l(w,r,1)$ and $a_k^*(w,r) := k(w,r,1)$.

Then, $l^*(w,k,\overline{q}) = \overline{q}a_l^*(w,r)$ and $k^*(w,k,\overline{q}) = \overline{q}a_k^*(w,r)$.

(OPTIONAL) The proof consists in rewriting the optimization problem in terms of one unit of product. This includes rewriting both the objective function and the constraint.
The optimization problem is

$$\min_{l,k} C = wl + rk \text{ subject to } f(l,k) = \overline{q}$$

We use that since the production function has CRS, then $f(\alpha l, \alpha k) = \alpha f(l,k)$ for any $\alpha > 0$. One trick used when we have a linearly homogeneous function is defining $\alpha$ in such a way that it expresses the arguments of the function in a specific way.
Take in particular, $\alpha := \frac{1}{\overline{q}}$ so that $f\left(\frac{l}{\overline{q}}, \frac{k}{\overline{q}}\right) = \frac{f(l,k)}{\overline{q}}$. The constraint $f(l,k) = \overline{q}$ is equivalent to $\frac{f(l,k)}{\overline{q}} = 1$ and so it can be reexpressed as $f\left(\frac{l}{\overline{q}}, \frac{k}{\overline{q}}\right) = 1$.
Moreover, we know that if we apply a monotone transformation to the objective function, we would still obtain the same solution. Keep in mind that the objective function would change its value, but this does not affect the optimization problem. Thus, by dividing by $\overline{q}$ we can express the total cost as an average cost: $\frac{C}{\overline{q}} = w\frac{l}{\overline{q}} + r\frac{k}{\overline{q}}$.
Then, the optimization problem becomes

$$\min_{l,k} \frac{C}{\overline{q}} = w\frac{l}{\overline{q}} + r\frac{k}{\overline{q}} \text{ subject to } f\left(\frac{l}{\overline{q}}, \frac{k}{\overline{q}}\right) = 1$$

Define $a_l := \frac{l}{\overline{q}}$, $a_k := \frac{k}{\overline{q}}$ and $c := \frac{C}{\overline{q}}$. Then, the optimization problem can be even expressed as,

$$\min_{a_l,a_k} c = wa_l + ra_k \text{ subject to } f(a_l,a_k) = 1. \tag{9.1}$$

Notice $c$ was originally defined as the average cost. But, given how we wrote the optimization problem, it is also the solution to the problem when $\overline{q} = 1$. If we forget about the definitions of each variable, (9.1) is just a change of variables with the constraint defined for $\overline{q} = 1$.
The result implies that the optimization problem can be solved by finding the choice of inputs to produce $\overline{q} = 1$, which we denote $a_l^*(w,r)$ and $a_k^*(w,r)$, and then recovering optimal input choices for any $\overline{q}$ by using the definition of the $a$'s. This deermines that $l^*(w,r,\overline{q}) = \overline{q}a_l^*(w,r)$ and $k^*(w,r,\overline{q}) = \overline{q}a_k^*(w,r)$.

The property states that factor demands are proportional to $\overline{q}$. Due to this, each can be expressed as the factor necessary to produce one unit of the good times the number of goods to produce. Thus, the optimal demands of factors can be fully characterized by simply knowing the factor demands for one unit of the good.

The second property that they satisfy is related to the value function (i.e. the minimum cost function).

**2. Minimum costs are linear in $\overline{q}$.** Let $c^*(w,r) := C^*(w,r,1)$ be the minimum cost to produce one unit of the good.

Then, $C^*(w,r,\overline{q}) = \overline{q}c^*(w,r)$.

By definition, the minimum cost function is $C^*(w,r,\overline{q}) = wl^*(w,r,\overline{q}) + rk^*(w,r,\overline{q})$. Also, we have shown that the optimal inputs demands are $l^*(w,k,\overline{q}) = \overline{q}a_l^*(w,r)$ and $k^*(w,k,\overline{q}) = \overline{q}a_k^*(w,r)$. Thus,
$C^*(w,r,\overline{q}) = w\left[\overline{q}a_l^*(w,r)\right] + r\left[\overline{q}a_k^*(w,r)\right]$
$\Rightarrow C^*(w,r,\overline{q}) = \overline{q}\left[wa_l^*(w,r) + ra_k^*(w,r)\right]$
$\Rightarrow C^*(w,r,\overline{q}) = \overline{q}c^*(w,r)$ where $c^*(w,r) := C^*(w,r,1)$.

Just like with the optimal demand of factors, the result states that the minimum cost function is fully characterized by the costs along one isoquant. Thus, given the costs when $\overline{q} = 1$, denoted by $c^*(w,r)$, the total cost to produce $\overline{q}$ units is $\overline{q}c^*(w,r)$.

Summing up, Property 1 and 2 imply that, *when the technology displays CRS, the solution for one unit of the product completely characterizes both the factors demands and the minimum costs, for any level of output*—the relation of each is always proportional. The next result follows as a corollary of these two properties, and it simply follows by the proportional relation with $\overline{q}$.[1]

**3. Optimal marginal costs, average costs and unitary costs are equal.** Formally, it means that $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} = \frac{C^*(w,r,\overline{q})}{\overline{q}} = c^*(w,r)$.

By taking the derivative of $C^*(w,r,\overline{q}) = \overline{q}c^*(w,r)$ with respect to $\overline{q}$, we get $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} = c^*(w,r)$. Then, by using that $C^*(w,r,\overline{q}) = \overline{q}c^*(w,r)$ we have that $\frac{C^*(w,r,\overline{q})}{q} = c^*(w,r)$.

Expressed in words, marginal costs and average costs are equal when a technology displays CRS, and both in turn equal the costs of producing one unit of the good.

The last property we present states that, when the technology exhibits CRS, *the expenditure share of an input is only affected by the input prices, but not by the level of production.* This comes in handy when a researcher works empirically, where we usually do not have information about quantities and prices of each input used. Instead, we observe the total expenditure on each input.

**4. Expenditure shares of factors are independent of the scale of production.** Let the expenditure share be defined as the expenditure relative to total costs. Formally

[1] In fact, any linear function has the property that its elasticity is unitary. In this case, notice that we are saying that $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} = \frac{C^*(w,r,\overline{q})}{\overline{q}}$. Hence, $\frac{\partial C^*(w,r,\overline{q})}{\partial \overline{q}} \frac{\overline{q}}{C^*(w,r,\overline{q})} = 1$ and the LHS is just the definition of an elasticity. This can be even seen by $\frac{\partial \ln C^*(w,r,\overline{q})}{\partial \ln \overline{q}} = 1$.

$s_l^*(w,r) := \frac{wl^*(w,r,\overline{q})}{C^*(w,r,\overline{q})} = \frac{wa_l^*(w,r)}{c^*(w,r)}$. If there are CRS, then $s_k^*(w,r) := \frac{rk^*(w,r,\overline{q})}{C^*(w,r,\overline{q})} = \frac{ra_k^*(w,r)}{c^*(w,r)}$.

Let's prove the result for labor. The proof for capital is analogous. The optimal labor demand is $l^*(w,r,\overline{q}) = \overline{q}a_l^*(w,r)$. Multiplying both sides by $w$,

$wl^*(w,r,\overline{q}) = \overline{q}wa_l^*(w,r)$,

and dividing both sides by $C^*(w,r,\overline{q})$,

$\frac{wl^*(w,r,\overline{q})}{C^*(w,r,\overline{q})} = \frac{\overline{q}wa_l^*(w,r)}{C^*(w,r,\overline{q})}$,

and replacing in the RHS for $C^*(w,r,\overline{q}) = \overline{q}c^*(w,r)$,

$\frac{wl^*(w,r,\overline{q})}{C^*(w,r,\overline{q})} = \frac{wa_l^*(w,r)}{c^*(w,r)}$.

An implication of this result is that, irrespective of whether a firm produces a little or a lot, the expenditure share of each factor is always the same when the factor prices do not change. This can simplify the problem analyzed, but unfortunately it also rules out some common scenarios. For instance, it cannot reflect that big firms usually spend proportionally more on capital or high-skilled workers, relative to small firms.

## 9.3 Homogeneous Technologies with CRS

Next, we investigate some specific functional forms that satisfy CRS, including the Cobb Douglas, Leontief, and a linear production function. To provide a solution for each case, we build on the solution used in the EMP of consumer theory. Recall that the EMP of consumer theory and the CMP in firm theory are equivalent. Due to this, the focus will be primarily on the interpretations of these functions in the context of firm theory.

The analysis concludes by studying one more functional form, not covered in consumer theory: the CES production function. For some values of parameters, the CES collapses to the previous three cases, thereby constituting a parsimonious technology representation.

### 9.3.1 Cobb Douglas

Suppose that the production function is a Cobb Douglas

$$\min_{l,k} C = wl + rk$$

$$\text{subject to } \overline{q} = l^{\beta_l} k^{\beta_k},$$

where $\beta_l + \beta_k = 1$ to ensure that the technology exhibits CRS.

The Cobb Douglas $f(l,k) := l^{\beta_l} k^{\beta_k}$ satisfies $f \in H^{\beta_l+\beta_k}$ and so it is a homogeneous function. It defines IRS, DRS or CRS depending if $\beta_l + \beta_k > 1$, $< 1$ or $= 1$, respectively.

CRS require that the production function is homogeneous of degree one. Hence, we cannot use the Cobb Douglas specification with a logarithmic transformation. This is in contrast to what we could do with utility functions.
Formally, the Cobb Douglas specification has to be $f(k,l) = k^{\beta_k} l^{\beta_l}$ which is homogeneous of degree one if $\beta_k + \beta_l = 1$. On the contrary, $f(k,l) = \beta_k \ln k + \beta_l \ln l$ does not satisfy the property of CRS.
Some people specify the Cobb Douglas production by $f(k,l) = \exp(\beta_k \ln k + \beta_l \ln l)$. This $f$ is the same Cobb Douglas given by $k^{\beta_k} l^{\beta_l}$. The expression is obtained by applying logs and exps to $k^{\beta_k} l^{\beta_l}$ which cancel out. Thus, it is not a monotone transformation but a way to rewrite $k^{\beta_k} l^{\beta_l}$.

Sometimes, the Cobb Douglas production function includes a parameter $A > 0$ such that $f(k,l) = A l^{\beta_l} k^{\beta_k}$. This parameter is known as Hicks-neutral technological improvements. It reflects variations in productivity that do not affect the optimal input choices.

By solving the optimization problem, we obtain

$$l^*(w,r,\overline{q}) = \overline{q} \left( \frac{\beta_l}{\beta_k} \frac{r}{w} \right)^{\beta_k},$$

$$k^*(w,r,\overline{q}) = \overline{q} \left( \frac{\beta_k}{\beta_l} \frac{w}{r} \right)^{\beta_l}.$$

The Lagrangian is:
$\mathscr{L} := wl + rk + \lambda \left[ \overline{q} - l^{\beta_l} k^{\beta_k} \right]$
and the FOCs are:
$\mathscr{L}'_l = \beta_l (l)^{\beta_l - 1} k^{\beta_k} - \lambda w = 0$
$\mathscr{L}'_k = \beta_k (k)^{\beta_k - 1} l^{\beta_l} - \lambda r = 0$
$\mathscr{L}'_\lambda = \overline{q} - l^{\beta_l} k^{\beta_k} = 0$
Like for any solution that involves a Cobb Douglas, we divide the first two equations so that $\frac{\beta_l}{\beta_k} \frac{k}{l} = \frac{w}{r}$. From this, we obtain an expression for $k$ as a function of $l$: $k = \frac{\beta_k}{\beta_l} \frac{w}{r} l$.
Plugging in this expression into $\mathscr{L}'_\lambda = 0$:
$\overline{q} - l^{\beta_l} k^{\beta_k} = 0 \Rightarrow \overline{q} - l^{\beta_l} \left( \frac{\beta_k}{\beta_l} \frac{w}{r} l \right)^{\beta_k} = 0 \Rightarrow \overline{q} - l^{\beta_l + \beta_k} \left( \frac{\beta_k}{\beta_l} \frac{w}{r} \right)^{\beta_k} = 0$

and using that $\beta_l + \beta_k = 1$, then we can determine that $l^* (w, r, \overline{q}) = \overline{q} \left( \frac{\beta_l}{\beta_k} \frac{r}{w} \right)^{\beta_k}$.
Likewise, we use that $k = \frac{\beta_k}{\beta_l} \frac{w}{r} l$ and so $k^* (w, r, \overline{q}) = \left( \frac{\beta_k}{\beta_l} \frac{w}{r} \right) l^* (w, r, \overline{q})$. This determines that $k^* (w, r, \overline{q}) = \overline{q} \left( \frac{\beta_k}{\beta_l} \frac{w}{r} \right)^{\beta_l}$

Using the notation we used in Property 1 for the general case of CRS, the demand of factors for one unit of production is $a_l^* (w, r) := \left( \frac{\beta_l}{\beta_k} \frac{r}{w} \right)^{\beta_k}$ and $a_k^* (w, r) := \overline{q} \left( \frac{\beta_k}{\beta_l} \frac{w}{r} \right)^{\beta_l}$. Likewise, the minimum cost function is

$$C^* (w, r, \overline{q}) = \overline{q} \left( \frac{w}{\beta_l} \right)^{\beta_l} \left( \frac{r}{\beta_k} \right)^{\beta_k}$$

By definition, $E^* (p_1, p_2, U_0) = p_1 h_1^* (p_1, p_2, U_0) + p_2 h_2^* (p_1, p_2, U_0)$. Hence,
$E^* (p_1, p_2, U_0) = p_1 U_0 \left( \frac{\alpha_1}{\alpha_2} \frac{p_2}{p_1} \right)^{\alpha_2} + p_2 U_0 \left( \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} \right)^{\alpha_1}$
$\Rightarrow E^* (p_1, p_2, U_0) = U_0 \left[ p_1 \left( \frac{\alpha_1}{\alpha_2} \frac{p_2}{p_1} \right)^{\alpha_2} + p_2 \left( \frac{\alpha_2}{\alpha_1} \frac{p_1}{p_2} \right)^{\alpha_1} \right]$
$\Rightarrow E^* (p_1, p_2, U_0) = U_0 \left[ (p_1)^{1-\alpha_2} \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} (p_2)^{\alpha_2} + (p_2)^{1-\alpha_1} \left( \frac{\alpha_2}{\alpha_1} \right)^{\alpha_1} (p_1)^{\alpha_1} \right]$
By using that $\alpha_1 + \alpha_2 = 1$, then $\alpha_2 = 1 - \alpha_1$ and $\alpha_1 = 1 - \alpha_2$. Therefore,
$E^* (p_1, p_2, U_0) = U_0 \left[ (p_1)^{\alpha_1} \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} (p_2)^{\alpha_2} + (p_2)^{\alpha_2} \left( \frac{\alpha_2}{\alpha_1} \right)^{\alpha_1} (p_1)^{\alpha_1} \right]$
$\Rightarrow E^* (p_1, p_2, U_0) = U_0 (p_1)^{\alpha_1} (p_2)^{\alpha_2} \left[ \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} + \left( \frac{\alpha_2}{\alpha_1} \right)^{\alpha_1} \right]$
Finally, using that $\left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} + \left( \frac{\alpha_2}{\alpha_1} \right)^{\alpha_1} = \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} + \left( \frac{\alpha_2}{\alpha_1} \right)^{1-\alpha_2}$ we can reexpress the RHS $\left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} + \left( \frac{\alpha_2}{\alpha_1} \right) \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} \Rightarrow \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} \left( 1 + \frac{\alpha_2}{\alpha_1} \right) \Rightarrow \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} \left( \frac{\alpha_1 + \alpha_2}{\alpha_1} \right) \Rightarrow (\alpha_1)^{\alpha_2 - 1} (\alpha_2)^{-\alpha_2}$ which is just $\left( \frac{1}{\alpha_1} \right)^{\alpha_1} \left( \frac{1}{\alpha_2} \right)^{\alpha_2}$.
Thus, $E^* (p_1, p_2, U_0) = U_0 (p_1)^{\alpha_1} (p_2)^{\alpha_2} \left[ \left( \frac{\alpha_1}{\alpha_2} \right)^{\alpha_2} + \left( \frac{\alpha_2}{\alpha_1} \right)^{\alpha_1} \right]$ becomes
$E^* (p_1, p_2, U_0) = U_0 (p_1)^{\alpha_1} (p_2)^{\alpha_2} \left( \frac{1}{\alpha_1} \right)^{\alpha_1} \left( \frac{1}{\alpha_2} \right)^{\alpha_2}$ which gives the result.

Using the notation of Property 1 again, this means that the unit cost is

$$c^* (w, r) := \left( \frac{w}{\beta_l} \right)^{\beta_l} \left( \frac{r}{\beta_k} \right)^{\beta_k}$$

Notice that simultaneously applying a log and exp transformation to $c^*$ does not affect $c^*$. However, it results results in expression commonly used in academic articles, which is

$$c^* (w, r) := \exp \left[ \beta_l \ln \left( \frac{w}{\beta_l} \right) + \beta_k \ln \left( \frac{r}{\beta_k} \right) \right].$$

Finally, the share of labor and capital expenditures relative to total cost are defined as $s_l^* (w, r, \overline{q}) := \frac{w l^* (w, r, \overline{q})}{C^* (w, r, \overline{q})}$ and $s_k^* (w, r, \overline{q}) := \frac{r k^* (w, r, \overline{q})}{C^* (w, r, \overline{q})}$. In the case of Cobb Douglas, they are equal to

$$s_l^*(w,r,\overline{q}) = \beta_l,$$

$$s_k^*(w,r,\overline{q}) = \beta_k.$$

This implies that the expenditure shares for each input are constant under a Cobb Douglas.

Notice that the share of input expenditure gives us another way to express each optimal input demand since $l^*(w,r,\overline{q}) = \frac{s_l^*(w,r,\overline{q})C^*(w,r,\overline{q})}{w}$ and $k^*(w,r,\overline{q}) = \frac{s_k^*(w,r,\overline{q})C^*(w,r,\overline{q})}{r}$. In principle, the terms on the RHS could be obtained.

Keep in mind that the shares $s_l^*$ and $s_k^*$ never depend on $\overline{q}$ under any CRS technology. As we indicated above, this implies that big and small firms have the same cost shares. However, the Cobb Douglas additionally implies that the factor shares neither depend on $(w,r)$. This adds more structure to problem relative to technologies with CRS, ruling out some scenarios that we can observe in reality. For instance, it entails that the cost shares of labor and capital are the same in a country with cheaper labor like China as in a country like the USA.

### 9.3.2 Leontief Function (Perfect Complements)

Just like in consumer theory, the Leontief function describes technologies with inputs that are perfect complements. The optimization problem is then

$$\min_{l,k} \text{C} = wl + rk$$

$$\text{subject to } \overline{q} = \inf\left\{\frac{l}{a_l}, \frac{k}{a_k}\right\}.$$

The production function exhibits CRS since $f(\alpha l, \alpha k) = \inf\left\{\frac{\alpha l}{a_l}, \frac{\alpha k}{a_k}\right\}$ and we have that $\frac{\alpha l}{a_l} \leqslant \frac{\alpha k}{a_k}$ iff $\frac{l}{a_l} \leqslant \frac{k}{a_k}$. Hence, $\inf\left\{\frac{\alpha l}{a_l}, \frac{\alpha k}{a_k}\right\} = \alpha \inf\left\{\frac{l}{a_l}, \frac{k}{a_k}\right\}$ which equals $\alpha f(l,k)$.

The Leontief production function has a similar interpretation as in consumer theory. In particular, $a_l$ and $a_k$ are the requirements of each factor to produce one unit of the good. This implies that a firm obtains one unit of the good when *both* $l = a_l$ and $k = a_k$.

What happens if the firm is having inputs $l = a_l$ and $k = a_k$, and suddenly increases the amount of one factor in isolation? The minimum would still be one, and so the quantities produced would remain the same—there is no possibility of substitution between factors, since both are essential. Both factors need to increase simultaneously and in a proportion $a_l$ and $a_k$ to increase production.

One example of a Leontief production function is a recipe to prepare some dessert. Imagine you want to prepare a brownie cake and have found on the internet that you need 0.5 kgs of cocoa powder, 2 eggs and 0.25 liters of milk. Denote $C$, $E$ and $M$ the quantity you are going to buy of each input. The requirements to produce one brownie cake are $a_C := 0.5$, $a_E := 2$ and $a_M := 0.25$. Thus, the production function is

$$f(C, E, M) := \inf \left\{ \frac{C}{0.5}, \frac{E}{2}, \frac{M}{0.25} \right\}$$

Notice that if you buy $C = 0.5$, $E = 2$ and $M = 0.25$, then you can produce one brownie cake. If you decide to buy more of any input in isolation, for example you buy three eggs, then the additional egg is not going to increase the number of brownie cakes you would get. You would still be getting one brownie cake.

Following the intuition provided in consumer theory, an optimal solution needs to satisfy that the arguments in $\inf \left\{ \frac{l}{a_l}, \frac{k}{a_k} \right\}$ are equal, so that $\frac{l}{a_l} = \frac{k}{a_k}$. If we instead increase one of the inputs, then the product would remain the same but the firm would have higher costs. Thus, any pair of factors with $\frac{l}{a_l} \neq \frac{k}{a_k}$ cannot be part of a solution to the CMP.

Since both arguments have to be equal, the constraint can be written as $\overline{q} = \inf \left\{ \frac{l}{a_l}, \left( \frac{a_k l}{a_l} \right) \frac{1}{a_k} \right\}$, giving as solution

$$l^* (\overline{q}) = a_l \overline{q},$$

$$k^* (\overline{q}) = a_k \overline{q},$$

and so

$$C^* (w, r, \overline{q}) = \overline{q} (w a_l + r a_k).$$

Notice that, consistent with the idea that there is no substitution between factors, changes in wages or in the remuneration to capital do not change the demand for factors. This arises since factors are perfect complements, so that the firm cannot increase its production by just substituting a more expensive factor for another cheaper one.

#### 9.3.2.1 Some Examples Covered by a Leontief Technology

We provide two scenarios where using a Leontief technology is appropriate. This requires identifying situations where inputs have to be combined in a specific proportion.

The first example is intuitive and really simple: a call center. Providing service to a customer needs at least one person with one phone (presumably, it also requires a computer). Thus, $a_l = a_k = 1$, which reflects that a person without a phone or a phone without a person cannot provide service to a customer.

A second example occurs when production takes place in stages. This entails that inputs are transformed during the production process, until the final product is created. For example, the production of cell phones requires building all all the parts and then assemble them. This two-stage process can be captured by a Leontief technology.

To see this, suppose that the production of cell phones requires producing an $n$ numbers inputs, where the unitary requirement of input $j$ is $a_j$ for $j = 1, 2, ..., n$. After this, the inputs are assembled to produce one cell phone by using $a_l$ workers. Then, letting $i_j$ and $l$ respectively denote input $j$ and labor, we can represent the production function by

$$\inf \left\{ \frac{i_1}{a_1}, \frac{i_2}{a_2}, ..., \frac{i_n}{a_n}, \frac{l}{a_l} \right\}.$$

Summing up, we use a Leontief technology to describe the production process of a firm when two conditions are met: each input is essential and there is no possibility of substitution between them.

### 9.3.3 Perfect Substitutes

For the following production function, we change the type of factors under consideration. The goal is to show that production theory is flexible enough to capture different phenomena, depending on what we want to model.

One area where production functions are relevant is income distribution. Technologies determines the remuneration of each factor, including wages, which is the main determinant of income for a vast part of the population. By thinking of how different policies and shocks affect the production process, we can then infer the impact on income distribution.

Some decades ago (actually, more like a century ago), models were using capital and labor as the production factors. The goal was to analyze the distribution of income between entrepreneurs (i.e. owners of capital) and workers. The distinction was suitable, since labor was a more or less homogeneous factor. Thus, tensions regarding income distribution were mainly between workers and owners.

Lately, differences between workers have become more pronounced, making differences between wages of workers become more disparate. Papers have started to recognize this, analyzing the phenomenon by assuming low-skilled (LS) and high-skilled workers (HS) as production factors. Next, we provide an example based on this classification.

#### 9.3.3.1 The CMP

Suppose there are two techniques to produce a good. To get one unit of output, the first technology only requires $a_H$ units of HS workers (no LS workers), while the second one only $a_L$ units of LS workers (no HS workers).[2] Denote the HS and LS workers hired by the firm by $H$ and $L$, respectively. Then, the firm's technology can be captured through

$$f(L,H) := \frac{L}{a_L} + \frac{H}{a_H},$$

where the production function exhibits CRS, since $f(\alpha L, \alpha H) = \frac{\alpha L}{a_L} + \frac{\alpha H}{a_H} = \alpha\left(\frac{L}{a_L} + \frac{H}{a_H}\right) = \alpha f(L,H)$.

As in the case of the Leontief production function, we have expressed the function by defining parameters $a_H$ and $a_L$ that define the necessary units of inputs to produce one unit of the good. It is also common to specify the production function by:

$$f(H,L) := \varphi_H H + \varphi_L L$$

where $\varphi_H := \frac{1}{a_H}$ and $\varphi_L := \frac{1}{a_L}$. In this case, $\varphi_i$ indicates the marginal productivity of a typical worker having skills $i$. This means, that $\varphi_i$ is the units of output that one unit of labor with skill $i$ would produce. For example, if you need 2 units of $H$ to produce one unit of the good, so that $a_H := 2$, you are also saying that one unit of $H$

[2] I units of the factor as the measure unit since it depends how we measure them. It could be hours or number of people.

produces $\varphi_H := \frac{1}{2}$ of output.

Let $w_H$ and $w_L$ be the wages of HS and LS respectively. The CMP is

$$\min_{L,H} C = w_L L + w_H H$$
$$\text{subject to } \overline{q} = \frac{H}{a_H} + \frac{L}{a_L}$$

There are two different ways to produce the $\overline{q}$ units of the good: either by only hiring HS workers or by only hiring LS workers. Since the firm's goal is to minimize cost, it will choose the less expensive technique. If both methods entail the same cost, then any combination is a solution to the CMP.

To obtain the solution, we proceed as in consumer theory with perfect substitutes. We assume that the solution is at a corner. Then, we obtain the conditions to make each a solution, by using that the solution has to provide the minimum cost.

Suppose $L^* = 0$. Using the constraint equation, $H^* = \overline{q}a_H$ and hence $C^* = \overline{q}w_H a_H$. Similarly, if $H^* = 0$, then $L^* = \overline{q}a_L$ and $C^* = \overline{q}w_L a_L$.

The combination $L^* = 0$ and $H^* = \overline{q}a_H$ is the unique solution when it represents the lowest cost. Thus, it has to satisfy that $\overline{q}w_H a_H < \overline{q}w_L a_L$, or, what is same, $\frac{w_H}{w_L} < \frac{a_L}{a_H}$.

By the same token, the combination $H^* = 0$ and $L^* = \overline{q}a_L$ is a solution when $\frac{w_H}{w_L} > \frac{a_L}{a_H}$.

In case that $\frac{w_H}{w_L} = \frac{a_L}{a_H}$, both combinations are a solution. But, also, any combination that produces in total $\overline{q}$ is so. Thus, given a value $L^* \in [0, \overline{q}a_L]$ $H^*$ is pinned down by the fact that production equals $\overline{q}$, so that, $H^* = \left(\overline{q} - \frac{L^*}{a_L}\right) a_H$.

Summing up, the solution is

$$L^*(w_H, w_L, \overline{q}) := \begin{cases} \overline{q}a_L & \text{if } \frac{w_H}{w_L} > \frac{a_L}{a_H} \\ 0 & \text{if } \frac{w_H}{w_L} < \frac{a_L}{a_H} \\ [0, \overline{q}a_L] & \text{if } \frac{w_H}{w_L} = \frac{a_L}{a_H} \end{cases},$$

$$H^*(w_H, w_L, \overline{q}) := \begin{cases} 0 & \text{if } \frac{w_H}{w_L} > \frac{a_L}{a_H} \\ \overline{q}a_H & \text{if } \frac{w_H}{w_L} < \frac{a_L}{a_H} \\ \left(\overline{q} - \frac{L^*}{a_L}\right) a_H & \text{if } \frac{w_H}{w_L} = \frac{a_L}{a_H} \end{cases},$$

and the minimum cost function is

$$C^*(w_H, w_L, \overline{q}) := \begin{cases} \overline{q}a_L w_L & \text{if } \frac{w_H}{w_L} > \frac{a_L}{a_H} \\ \overline{q}a_H w_H & \text{if } \frac{w_H}{w_L} < \frac{a_L}{a_H} \\ \overline{q}a_H w_H & \text{if } \frac{w_H}{w_L} = \frac{a_L}{a_H} \end{cases}.$$

Notice that $w_H a_H = w_L a_L$ when $\frac{w_H}{w_L} = \frac{a_L}{a_H}$, and so saying that the minimum cost is either $\overline{q} w_H a_H$ or $\overline{q} w_L a_L$ is equivalent. Furthermore, we can write the minimum cost in a more compact way:

$$C^* (w_H, w_L, \overline{q}) := \overline{q} \inf \{w_L a_L, w_H a_H\}.$$

Although it is not necessary for the solution, we should suppose that $a_L > a_H$ to be consistent with our interpretation of workers of different skills. This means that producing one unit of output with LS requires more LS workers than if that unit of output is produced with HS workers. In that way, we capture that HS workers are more productive than LS workers.

Would a firm choose producing with LS workers if HS workers are more productive? For this to occur, LS workers need to have a low wage relative to HS. Formally, $\frac{w_L}{w_H}$ has to be sufficiently low. This type of phenomenon could arise in the apparel industry. North American firms might produce in Asia or Mexico, where labor is cheaper, even though their workers are potentially less productive than the ones in USA or Canada (see the case of Nike in the Problem Set).

### 9.3.4 CES (Constant Elasticity of Substitution) Production Function

The CES is a production function belongs to the group of well-behaved technologies, such as the Cobb Douglas. Nonetheless, unlike this one, its expenditure shares are not independent of the factors prices.

**Remark** *For the CES case, I want you to focus on the properties of the solution to the CMP, rather than how to solve it. The algebra is a little bit messy, without providing much insight.*

We keep using LS and HS workers as the production factors. The CMP is

$$\min_{L,H} C = w_L L + w_H H$$
$$\text{subject to } \overline{q} = \left( (\varphi_L)^{\frac{1}{\gamma}} (L)^{\frac{\gamma-1}{\gamma}} + (\varphi_H)^{\frac{1}{\gamma}} (H)^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}},$$

where the parameters $\varphi_L$ and $\varphi_H$ reflect each factor's productivity.

There are two important properties that the CES has. First, it displays CRS. Second, it converges to the three production functions we have covered, depending on the value of $\gamma$:

- if $\gamma \to 0$, the CES converges to the Leontief production function.
- if $\gamma \to 1$, the CES converges to the Cobb Douglas production function.
- if $\gamma \to \infty$, the CES converges to the linear production function.

The term "CES" reflects that the degree of substitution between inputs is constant and given by the parameter $\gamma$. It is such that the higher $\gamma$ is, the greater the substitutability between inputs, which is equivalent to a lower the complementarity.

The case with $\gamma \to 0$ represents perfect complelementary. Likewise, inputs are imperfect complements when $\gamma \in (0, 1)$, which means that they are complements, but not required in a fixed proportion. Finally, inputs are imperfect substitutes when $\gamma > 1$, and perfect substitutes in the limit case where $\gamma \to \infty$.

Since we can interpret a lower complementarity as a higher degree of substitutability, we can say that $\gamma$ controls the substitutability. This is important from an empirical point of view. Sometimes we do not want to directly *assume* that inputs are substitutes or complements. This is implicitly done when we suppose a Leontief or a linear production function. Rather, we want the data to tell us how inputs behave, and the CES allows us to do this by inferring the information through estimating $\gamma$.

Another advantage of the CES is that it is a well-behaved production function. Therefore, it can be solved by using the Lagrange technique. The solution is given by:

$$c(w_L, w_H) := \left(\sum_{l\in\mathcal{L}} \varphi_l w_l^{1-\gamma}\right)^{\frac{1}{1-\gamma}},$$

$$C(\overline{q}, w_L, w_H) := \overline{q}c(\mathbf{w}),$$

$$L^*(\overline{q}, w_L, w_H) := \overline{q}\varphi_L\left(\frac{w_L}{c(w_L, w_H)}\right)^{-\gamma} = \overline{q}\frac{\varphi_L (w_L)^{-\gamma}}{\left[\varphi_L (w_L)^{1-\gamma} + \varphi_H (w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}},$$

$$H^*(\overline{q}, w_L, w_H) := \overline{q}\varphi_H\left(\frac{w_H}{c(w_L, w_H)}\right)^{-\gamma} = \overline{q}\frac{\varphi_H (w_H)^{-\gamma}}{\left[\varphi_L (w_L)^{1-\gamma} + \varphi_H (w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}.$$

(OPTIONAL) The derivation is optional since it requires a lot of algebra and there is nothing really insightful in it. I want you to know the results and implications of the CES production function.

The Lagrangian is:

$\mathscr{L} := w_L L + w_H H + \lambda\left[\overline{q} - \left((\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}} + (\varphi_H)^{\frac{1}{\gamma}}(H)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}}\right]$

and the FOCs are:

$\mathscr{L}'_l = w_L - \lambda\left((\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}} + (\varphi_H)^{\frac{1}{\gamma}}(H)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}-1}(\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}-1} = 0$

$\mathscr{L}'_k = w_H - \lambda\left((\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}} + (\varphi_H)^{\frac{1}{\gamma}}(H)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}-1}(\varphi_H)^{\frac{1}{\gamma}}(H)^{\frac{\gamma-1}{\gamma}-1} = 0$

$\mathscr{L}'_\lambda = \overline{q} - \left((\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}} + (\varphi_H)^{\frac{1}{\gamma}}(H)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}} = 0$

Notice that the constraint can be expressed as

$\mathscr{L}'_\lambda = \overline{q} - \left(\left(\frac{L}{(\varphi_L)^{\frac{1}{\gamma-1}}}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{H}{(\varphi_H)^{\frac{1}{\gamma-1}}}\right)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}} = 0$

Divide the first two equations, we obtain that:

$\frac{w_L}{w_H} = \frac{(\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}-1}}{(\varphi_H)^{\frac{1}{\gamma}}(H)^{\frac{\gamma-1}{\gamma}-1}}$ which, using that $\frac{\gamma-1}{\gamma} - 1 = \frac{-1}{\gamma}$, is equal to $\frac{w_L}{w_H} = \frac{\left(\frac{H}{\varphi_H}\right)^{\frac{1}{\gamma}}}{\left(\frac{L}{\varphi_L}\right)^{\frac{1}{\gamma}}}$. From this, we obtain a relation between $L$ and $H$:

$\Rightarrow \left(\frac{w_L}{w_H}\right)^{\gamma} = \frac{H/\varphi_H}{L/\varphi_L}$

note that $(\varphi_L)^{\frac{1}{\gamma}}(L)^{\frac{\gamma-1}{\gamma}} \Rightarrow \frac{\varphi_L}{(\varphi_L)^{1-\frac{1}{\gamma}}}(L)^{\frac{\gamma-1}{\gamma}} \Rightarrow \varphi_L\left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}}$ so

$\overline{q} - \left(\varphi_L\left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}} + \varphi_H\left(\frac{H}{\varphi_H}\right)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}} = 0 \Rightarrow (\overline{q})^{\frac{\gamma-1}{\gamma}} = \varphi_L\left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}} + \varphi_H\left(\left(\frac{w_L}{w_H}\right)^{\gamma}\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}}$

$\Rightarrow (\overline{q})^{\frac{\gamma-1}{\gamma}} = \varphi_L\left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}} + \varphi_H\left(\frac{w_L}{w_H}\right)^{\gamma-1}\left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}} \Rightarrow (\overline{q})^{\frac{\gamma-1}{\gamma}} = \left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}}\left[\varphi_L + \varphi_H\left(\frac{w_H}{w_L}\right)^{1-\gamma}\right]$

$\Rightarrow (\overline{q})^{\frac{\gamma-1}{\gamma}} = \left(\frac{L}{\varphi_L}\right)^{\frac{\gamma-1}{\gamma}}\left[\frac{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}{(w_L)^{1-\gamma}}\right] \Rightarrow \overline{q} = \left(\frac{L}{\varphi_L}\right)\left[\frac{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}{(w_L)^{1-\gamma}}\right]^{\frac{\gamma}{\gamma-1}}$

and isolating $L$

$L^*(\overline{q}, w_L, w_H) = \overline{q}\frac{\varphi_L(w_L)^{-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}$

By symmetry of the problem, by just relabeling the variables, we can also obtain $H^*$:

$H^*(\overline{q}, w_L, w_H) = \overline{q}\frac{\varphi_H(w_H)^{-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}$

To determine the minimum cost, first let's determine the total expenditure on each factor:

$w_L L^*(\overline{q}, w_L, w_H) = \overline{q}w_L\left[\frac{\varphi_L(w_L)^{-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}\right] \Rightarrow w_L L^*(\overline{q}, w_L, w_H) = \overline{q}\frac{\varphi_L(w_L)^{1-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}$

From this we determine the minimum cost:

$C^*(\overline{q}, w_L, w_H) := w_L L^*(\overline{q}, w_L, w_H) + w_H H^*(\overline{q}, w_L, w_H)$

$\Rightarrow C^{*}(\overline{q}, w_L, w_H) := \overline{q}\left[\frac{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}\right]$

From this we also determine that $c^{*}(\overline{q}, w_L, w_H) := \frac{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^{\frac{\gamma}{\gamma-1}}}$

In turn, the cost shares are:

$$s_L^{*}(w_L, w_H) := \frac{\varphi_L(w_L)^{1-\gamma}}{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}},$$
$$s_H^{*}(w_L, w_H) := \frac{\varphi_H(w_H)^{1-\gamma}}{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}.$$

Since the cost share is defined by $s_L^{*}(\overline{q}, w_L, w_H) := \frac{w_L L^{*}(\overline{q}, w_L, w_H)}{C^{*}(\overline{q}, w_L, w_H)}$, using the results we got for $w_L L^{*}(\overline{q}, w_L, w_H)$ and $C^{*}(\overline{q}, w_L, w_H)$, we establish that $s_L^{*}(w_L, w_H) = \frac{\varphi_L(w_L)^{1-\gamma}}{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}$.

Notice that now, unlike what happens with the Cobb Douglas, the costs shares are not constant with a CES. Specifically, although cost shares still do not depend on the quantities produced by the property of CRS, they do vary when the wages of each group change. In particular, for LS workers (similar result for HS workers):

$$\frac{\partial s_L^{*}}{\partial w_L} = \frac{(1-\gamma)}{w_L} s_L^{*}\left(1-s_L^{*}\right), \text{ which is } < 0 \text{ if } \gamma > 1,$$
$$\frac{\partial s_L^{*}}{\partial w_H} = \frac{(1-\gamma)}{w_H} s_L^{*}\left(1-s_L^{*}\right), \text{ which is } > 0 \text{ if } \gamma > 1$$

(OPTIONAL) Taking the derivative:

$\frac{\partial s_L^{*}}{\partial w_L} = \frac{(1-\gamma)\varphi_L(w_L)^{-\gamma}\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]-\varphi_L(w_L)^{1-\gamma}(1-\gamma)\varphi_L(w_L)^{-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^2}$

$\Rightarrow \frac{\partial s_L^{*}}{\partial w_L} = (1-\gamma)\frac{\varphi_L(w_L)^{-\gamma}\varphi_H(w_H)^{1-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^2} < 0.$

This can be reexpressed by:

$\frac{\partial s_L^{*}}{\partial w_L} = (1-\gamma)\frac{w_L}{w_L}\frac{\varphi_L(w_L)^{-\gamma}}{\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}}\frac{\varphi_H(w_H)^{1-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]}$

$\Rightarrow \frac{\partial s_L^{*}}{\partial w_L} = \frac{(1-\gamma)}{w_L} s_L^{*} s_H^{*}$, and since $s_H^{*} = 1 - s_L^{*}$, the result follows.

Also, by the same token, $\frac{\partial s_L^{*}}{\partial w_H} = (1-\gamma)\frac{\varphi_L(w_L)^{1-\gamma}\varphi_H(w_H)^{-\gamma}}{\left[\varphi_L(w_L)^{1-\gamma}+\varphi_H(w_H)^{1-\gamma}\right]^2}$ and so $\frac{\partial s_L^{*}}{\partial w_H} = \frac{(1-\gamma)}{w_H} s_L^{*}\left(1-s_L^{*}\right)$.

Keeping in mind that the case of $\gamma > 1$ represents perfect substitutes, we conclude that increases on the wages of LS workers or decreases in the wages of HS workers decrease the cost share of LS workers. Thus, substitutability is not only reflected in the LS workers hired, but also in terms of their cost share.

# Lecture Note 10

# What Makes a Firm Successful?

## 10.1 Introduction to Monopoly

In this lecture note, we develop a formal framework to identify what makes a firm successful. Broadly speaking, firms can be leaders in their markets due to advantages on the supply or the demand side. More specifically, one option is that firms can produce goods more efficiently, resulting in lower costs than the rest of the firms. This translates into a lower price, thereby gaining a better position in the market. Alternatively, firms can produce goods that are appealing for consumers, thus allowing them to operate in the market successfully.

In this note, we analyze these two sources that make a firm succeed. The goal is twofold:

[1] to understand how firms make choices, according to the nature of a firm's comparative advantage. The decisions to be analyzed encompass quantities, prices, and markups.

[2] to learn how to build a model.

Roughly speaking, to build a model we begin by establishing some assumptions about the environment. Then, we specify the agents' goals, what they know, and how they make choices to pursue their objectives. Finally, we derive some predictions, and in this note we do this by focusing on a **comparative statics** analysis. This methodology consists in varying the exogenous conditions of the model, and investigate how the endogenous variables react to it.

In this note, we proceed as follows. We start by presenting the baseline model, where we describe a firm. Then, we analyze how firms make pricing and quantity decisions. After this, we analyze how variations in efficiency and appeal impact a firm. Finally, we translate this information into different strategies that a firm can deploy to succeed.

## 10.2 The Baseline Model of Monopoly

Since our aim is to understand how firms make decisions, we keep the model as simple as possible. In particular, we analyze one industry in isolation and suppose it comprises only one firm.

Taking into account that we will consider an industry with only one firm, it is important to be clear about how a monopoly is defined. We usually tend to associate the term monopoly with the existence of one firm in the market. However, this is a narrow definition.[1] A **monopoly** is defined as any firm that has some scope to increase its price without losing its demand completely. Technically, we say that a monopoly does not face an infinitely elastic demand.

Although a firm operating alone in a market quite likely satisfies this condition, it is not the only possibility. For example, a firm active in an industry with multiple competitors could charge a price higher than its marginal cost if its product is sufficiently differentiated. A typical example is Apple with the iPhone in the cell phone industry.

In this sense, the results in these notes should be understood in a broad sense. They are not only relevant for a monopoly, but actually aim to identify the decision process of any firm that has some market power.

### 10.2.1 Setup

There is a single firm in the industry supplying one good. This firm has constant marginal costs $c$ and no fixed costs. Formally, the cost function is $C(q) := cq$, where $q$ is the quantity produced.

The total demand for the good is a function $q(p;\alpha) \in \mathcal{C}^2$, where $p \in \left[\underline{p}, \overline{p}\right]$ is the price and $\alpha \in \mathbb{R}_+$ a parameter. The features of the good are exogenously given, and we

[1] In fact, the existence of only one firm in the market is not sufficient to have a monopoly. If the barriers to entry are really low, the tacit competition that the firm faces could turn the industry competitive. The same caveat applies to the case of an oligopoly with a small number of firms. For instance, a competitive outcome would emerge when a duopoly competes in prices, in case goods are homogeneous and both firms have the same marginal costs.

assume that $\frac{\partial q(p;\alpha)}{\partial p} < 0$, thereby ruling out Giffen goods.

We will refer to $\alpha$ as **appeal**. It reflects that consumers make consumption decisions by comparing prices, but also other non-price aspects of the goods (e.g., quality, after-sale services, etc.) Below, we will make assumptions consistent with this interpretation of $\alpha$ (for instance, that a greater $\alpha$ increases demand).

Notice that by requiring $\alpha \in \mathbb{R}_+$, we are implicitly assuming that all the tangible and intangible features of the good can be reflected through a single measure. This is to keep the model parsimonious.

### 10.2.2 A Digression: Elasticities

To get information on how prices affect demand, the firm could analyze the derivative of the demand function with respect to prices. This provides information *expressed in units* of each variable. However, we will see that a firm's decisions ultimately depend on the derivatives expressed in *percentage* variations. In other terms, it is the concept of elasticity that matters for the firm.

The **price elasticity** of the demand is defined as $\varepsilon_p(p;\alpha) := -\frac{\partial \ln q(p;\alpha)}{\partial \ln p}$. Recall that there are three mathematical equivalent ways to expressing an elasticity, either

[1] $-\frac{\partial q(p;\alpha)/q(p;\alpha)}{\partial p/p}$,

[2] $-\frac{\partial q(p;\alpha)}{\partial p}\frac{p}{q(p;\alpha)}$ or

[3] $-\frac{\partial \ln q(p;\alpha)}{\partial \ln p}$.

The first one is the more intuitive. It shows that $\varepsilon_p$ is the percentage variation in quantities when there is a 1% variation in price. The elasticity has been multiplied by a negative sign to express it in absolute terms: since $\frac{\partial q(p;\alpha)}{\partial p} < 0$, then $-\frac{\partial q(p;\alpha)}{\partial p}\frac{p}{q(p;\alpha)} > 0$.

Although the first definition of elasticity provides a clear-cut interpretation, it is usually the second and the third definitions that are used for calculations. In particular, we will make extensive use of the logarithmic definition. The key to understanding expressions like $-\frac{\partial \ln q(p;\alpha)}{\partial \ln p}$ is treating the numerator and denominator like differentials.

For example let's take a function $f(q) := \ln q$. Differentiating $f$ we obtain $\mathrm{d}f = f'_q \,\mathrm{d}q$, which implies that $\mathrm{d}f = \frac{\mathrm{d}q}{q}$ since $f'_q = \frac{1}{q}$. Using that $\mathrm{d}f = \mathrm{d}\ln q$, we conclude that $\mathrm{d}\ln q = \frac{\mathrm{d}q}{q}$. This explains why the numerator in the first definition of elasticity can be expressed as either $\frac{\mathrm{d}q(p;\alpha)}{q(p;\alpha)}$ or $\mathrm{d}\ln q(p;\alpha)$.

### 10.2.3 Basic Assumptions

In this part, we state some basic assumptions about the relation of the quantity demanded with $p$ and $\alpha$. We will do it in terms of elasticities. In a later subsection, we add assumptions that are necessary to get unambiguous comparative statics.

Regarding price, we have already stated that $\frac{\partial q(p;\alpha)}{\partial p} < 0$, and so $\varepsilon_p(p;\alpha) > 0$ for all $(p;\alpha)$. The price elasticity provides information regarding the sensitivity of the quantity demanded to price increases. Intuitively, it quantifies the the firm's trade-off regarding revenues, between a higher price and a lower demand.

Some terminology is in order. We say that the **demand is inelastic** when $\varepsilon_p(p;\alpha) < 1$, whereas $\varepsilon_p(p;\alpha) > 1$ defines an **elastic demand**. For the baseline model, we assume that $\varepsilon_p(p;\alpha) > 1$ for all $(p;\alpha)$. As a corollary, we rule out demands that are inelastic for all prices.[2]

We have referred to $\alpha$ as appeal, but have not established assumptions that give meaning to the label. There are two salient features that demand appeal should fulfill to receive such denomination. First, it should be such that **a greater appeal results in an increase in demand**. This is captured by $\frac{\partial q(p;\alpha)}{\partial \alpha} > 0$, or $\varepsilon_\alpha := \frac{\partial \ln q(p;\alpha)}{\partial \ln \alpha} > 0$ in elasticity terms. It means that greater values of $\alpha$ increase the quantity demanded, keeping the price fixed. Notice we remain agnostic about how appeal increases demand—it could be either because old consumers demand more quantities or because the firm starts selling to new customers.

The second feature is given by the relation of appeal with price. We could conceive that a more attractive product not only increases the total quantities sold, but also affects

[2] In one of subsections, we outline the pricing decision of a firm when it faces an inelastic demand. We do this as a digression.

the sensitivity of the aggregate demand to price increases. For example, if a cell phone is faster, has more memory, better resolution, longer warranty, etc., it is reasonable to assume that a higher price reduces aggregate demand, but less so relative to a cell phone without those characteristics.

This feature can be captured by a greater value of $\alpha$ reducing $\varepsilon_p(p;\alpha)$. Expressed in words, it means that **when a good has more appeal, the demand becomes less price elastic** (equivalently, more price inelastic). Formally, $\frac{\partial \varepsilon_p(p;\alpha)}{\partial \alpha} \leq 0$, allowing for the possibility that appeal has a zero effect on the price elasticity. A zero effect could occur for, example, if the firm improves its distribution channels to reach new consumers that have the same valuation as the old ones. In that case, the quantities demanded would increase, but the sensitivity of the aggregate demand to prices would not change–the average sensitivity of the new consumers would be the same as the old ones.

**Assumption 10.2.1.** *Summing up, we have assumed that:*

- $\varepsilon_p > 1$
- $\varepsilon_\alpha(p;\alpha) > 0$ *and* $\frac{\partial \varepsilon_p(p;\alpha)}{\partial \alpha} \leq 0$.

We will need further assumptions in relation with $\frac{\partial \varepsilon_p(p;\alpha)}{\partial p}$. However, these assumptions will be added later, since they require deriving additional results.

### 10.2.4 The Optimization Problem

The firm's optimization problem is maximizing its profit by choosing price. Although the firm also has to choose its supply, this is completely determined by the demand function $q(p;\alpha)$, once the prices have been chosen—given the price chosen, the firm supplies a quantity that equals its demand.

**Remark** *When we ignore strategic interactions, the firm could alternatively choose quantities and let the price be determined by the condition of supply equal demand. In that case, rather than using the direct demand $q(p;\alpha)$, we would make use of the inverse demand $p(q;\alpha)$. In the absence of strategic interactions, nonetheless,*

**both optimization problems provide the same characterization of prices and quantities**. *It is only when strategic interactions are incorporated that we need to distinguish between both, giving rise to the Cournot and Bertrand models.*

Formally, the optimization problem is

$$\max_{p\in\left[\underline{p},\overline{p}\right]} \pi\left(p;\alpha,c\right) := q\left(p;\alpha\right)\left(p-c\right).$$

There are different assumptions guaranteeing that this is a well-behaved problem, in the sense that a solution exists, is unique, and interior. When we did a math review in the first classes of the course, I provided different conditions that we could use to ensure this. However, consistent with how the literature tackles exercises of comparative statics these days, we proceed in a different way. Specifically, we will state only assumptions that are necessary to get unambiguous results, and ***assume*** that a solution exists, is unique, and interior. To justify why we proceed in this fashion, some remarks are in order.

In recent decades, comparative statics has had some revival. One message of this new literature is that many assumptions usually made to perform comparative statics are in fact not necessary.[3] For this reason, it has been common since then only to establish a minimal set of assumptions to get unambiguous comparative statics. The goal is to make a clear distinction between assumptions made to obtain a well-behaved problem and those related to the comparative statics analysis.

I will proceed under the assumption that a solution exists, is unique, and interior, so you have a grasp of how papers present the results these days. Nevertheless, the conditions for this in the problem at hand can be characterized in a relatively simple way.

Just in case you are curious, I outline assumptions to obtain a well-defined problem. Since the price domain is compact and $\pi$ is continuous, an optimal price exists. Moreover, by using Inada conditions and assuming that profits are strictly quasiconcave, then a solution is interior and unique.

However, keep in mind that these are only sufficient conditions. For instance, it is not uncommon to have profits functions that are not strictly quasiconcave but the solution is unique anyway. Furthermore, we can sometimes predict how a parameter affects a solution, even if the solution is not unique.

There are many details that we need to take into account to tackle these issues, But a lot of them are unrelated or not necessary to provide an answer to what makes a firm successful.

[3]For instance, it has been shown that we can predict how a parameter affects a solution, even when the solution is not unique.

We characterize the solution by the FOC:

$$p^* = \frac{\varepsilon_p(p^*;\alpha)}{\varepsilon_p(p^*;\alpha) - 1} c. \tag{PRI}$$

The FOC is
$\frac{\mathrm{d}\pi}{\mathrm{d}p} = \frac{\partial Q(p;\alpha)}{\partial p}(p-c) + Q(p,\alpha) = 0 \Rightarrow \frac{\partial Q(p;\alpha)}{\partial p}\frac{1}{Q(p,\alpha)} = -\frac{1}{(p-c)}$
By multiplying both sides by $p$, then $-\frac{\partial Q(p;\alpha)}{\partial p}\frac{p}{Q(p,\alpha)} = \frac{p}{(p-c)}$.
Since $\varepsilon(p;\alpha) := -\frac{\partial Q(p;\alpha)}{\partial p}\frac{p}{Q(p,\alpha)}$ then $\varepsilon(p;\alpha) = \frac{p}{(p-c)}$ which by working it out becomes $p = \frac{\varepsilon(p;\alpha)}{\varepsilon(p;\alpha)-1}c$.

Since we have assumed that $\varepsilon_p(p;\alpha) > 1$ for all $(p;\alpha)$, then (PRI) determines that $p^* > c$. Notice that (PRI) provides only an implicit characterization of the optimal prices $p^*$, since the price elasticity also depends on the price. We denote the implicit value $p^*$ satisfying equation (PRI) by $p^*(\alpha, c)$.

Once that optimal prices are pinned down, optimal profits are

$$\pi^*(\alpha, c) := Q[p^*(\alpha, c), \alpha][p^*(\alpha, c) - c].$$

### 10.2.5 A Digression: The Inelastic Demand Case

There exists a solution when $\varepsilon(p;\alpha) < 1$ for any $p$, but it is not the one given by (PRI). Instead, the solution would lie on the boundary. The case is particularly interesting, since it justifies government regulation of markets.

Demand is inelastic when the firm increases the price of the good, and consumers decrease their consumption less than proportionally. This can be observed clearly in the extreme case where $\varepsilon(p;\alpha) = 0$, such that consumers still keep demanding the same quantity following a price increase. Examples potentially covered by this are necessity goods (electricity, water) or critical medicine (v.gr. oncological medication, insulin). Regarding the former type of good, they are usually provided by one firm, given the economies of scale that characterize the industry. If the government would not intervene at all and a monopoly supplied these goods, we could expect that the firm would charge such an extremely high price that not everyone could afford the good.

Let's first provide an intuition of the solution. Then, show formally that increases in price make the firm garner higher profits. Profits are given by the difference between revenues $R(p;\alpha) := pq(p;\alpha)$ and costs $C(p;\alpha) := cq(p;\alpha)$. A firm can increase its revenue through either a higher price or greater quantities. However, there is a trade-off: ceteris paribus, greater prices result in lower quantities demanded.

However, when the demand is inelastic, a 1% increase in price decreases the quantity demanded less than 1%. Thus, overall, increases in prices end up determining higher revenues. At the same time, the decrease in quantities reduces the costs: since the firm has to produce less, the cost $C(p;\alpha)$ becomes lower. All this provides the intuition that the best strategy for a monopoly when it sells an inelastic product is charging a price as high as possible.

Let's now show this formally. First, (PRI) cannot be a maximum, and it is in fact the solution that minimizes profits. This can be easily seen, since (PRI) entails that the price is lower than its cost. Thus, profits would be necessarily negative, and it would be better not to sell at all. From this, we conclude that the FOC cannot provide us with a solution. Since the FOC is a necessary condition for an interior solution (and since it can be shown that a solution exists), the solution must lie at the boundary.

How can we detect a corner solution? One clue is when an increase or decrease of the price entails greater profits for any $p$. In other terms, a corner solution arises when the first derivative is monotone. When this occurs, the best choice is to increase or decrease the control variable as much as possible.

To establish a direct link between this approach and the fact that $\varepsilon_p(p;\alpha) < 1$, it will be easier to analyze the behavior of $\frac{\partial \ln \pi(p,\alpha)}{\partial \ln p}$. Notice that this is without loss of generality, since $\frac{\partial \ln \pi(p,\alpha)}{\partial \ln p} > 0$ iff $\frac{\partial \pi(p,\alpha)}{\partial p} > 0$. Formally, $\frac{\partial \ln \pi(p,\alpha)}{\partial \ln p} = \frac{\partial \pi(p,\alpha)}{\partial p}\frac{p}{\pi(p,\alpha)}$, where profits and prices are positive.

We will show that $\frac{\partial \ln \pi(p,\alpha)}{\partial \ln p} > 0$ for any $p > c$. This requires showing that the sign holds, irrespective of the price at which we are evaluating the derivative (that is why we ask for the condition to hold for any $p > c$).

Applying logs to the profit function, $\ln \pi (p; \alpha) = \ln Q (p; \alpha) + \ln (p - c)$, we get:

$$\frac{\partial \ln \pi (p, \alpha)}{\partial \ln p} = \underbrace{\frac{\partial \ln Q (p, \alpha)}{\partial \ln p}}_{= -\varepsilon_p (p; \alpha)} + \frac{\partial \ln (p - c)}{\partial \ln p}.$$

Notice that $\frac{\partial \ln(p-c)}{\partial \ln p} = \frac{\partial \ln(p-c)}{\partial p} p$ because $\mathrm{d} \ln p = \frac{1}{p} \mathrm{d}p$. Then, $\frac{\partial \ln(p-c)}{\partial \ln p} = \frac{p}{p-c}$, and so

$$\frac{\partial \ln \pi (p, \alpha)}{\partial \ln p} = - \underbrace{\varepsilon_p (p; \alpha)}_{<1} + \underbrace{\frac{p}{p - c}}_{>1} > 0.$$

The inequality of each term follows because, $\varepsilon_p (p; \alpha) < 1$ by assumption, and firm always sells its good to a price $p > c$. Thus, $\frac{p}{p-c} > 1$. Moreover, since $\frac{\partial \ln \pi(p,\alpha)}{\partial \ln p} > 0$ for any $p > c$, then$p^* (\alpha, c) := \overline{p}$ for any $(\alpha, c)$ if $p \in [\underline{p}, \overline{p}]$. The firm would then choose $\overline{p}$, and so the firm would sell a quantity close to zero if $\overline{p}$ is a big number.

### 10.2.6 About Markups

Let's go back to the baseline case where $\varepsilon_p (p; \alpha) > 1$ for any $(p; \alpha)$. The optimal price is implicitly characterized by equation (PRI).

We introduce an important concept for the analysis, which is the **markup**. It is denoted by $\mu$ and defined by:

$$\mu (p; \alpha) := \frac{\varepsilon_p (p; \alpha)}{\varepsilon_p (p; \alpha) - 1}.$$

The name follows because $\mu$ evaluated at $p^*$ implies that (PRI) is

$$p^* = \mu (p^*; \alpha) c \tag{PRI-1}$$

and so $\frac{p^*}{c} = \mu (p^*; \alpha)$.

Markups give information about the revenue from selling one unit of the good (i.e. $p$) relative to the cost of producing that unit (i.e. $c$). Put it simple, it provides a ratio of revenue over cost per unit sold.

The first conclusion we can obtain using the concept of markups is regarding consumer welfare. In perfect competition, prices equal marginal costs, and so $\mu = \frac{p}{c} = 1$. For this reason, markups provide information about the monopoly power that a firm has, relative to the ideal case of perfect competition. This also reveals why consumers do not benefit from a monopoly: a firm prefers to restrict quantities with the goal of benefiting from higher prices and savings in production costs. The level of markup provides information about the extent to which this strategy is followed.

## 10.3 Comparative Statics (CS)

A CS analysis identifies the impact on endogenous variables due to exogenous changes in the parameters. In our model, the choice decision is the price, and there are two parameters $c$ and $\alpha$. Once we determine how prices are affected by these parameters, we also identify the effect on other endogenous variables, like profits, quantities, and markups.

For now, let's focus on how $\alpha$ and $c$ affect the optimal price $p^*(\alpha, c)$. We will perform the CS analysis by varying one parameter at a time. Our goal is to establish the signs of $\frac{\partial p^*(\alpha,c)}{\partial c}$ and $\frac{\partial p^*(\alpha,c)}{\partial \alpha}$.

### 10.3.1 Some Additional Assumptions

Before doing a comparative static analysis, it is necessary to add some assumptions to obtain unambiguous results. These assumptions are related to the impact of prices on price elasticity and markups.

We begin by showing that the effect of prices on the price elasticity has the same sign as the negative effect of prices on markups. To observe this, we know that $\mu(p;\alpha) := \frac{\varepsilon_p(p;\alpha)}{\varepsilon_p(p;\alpha)-1}$. Taking $\mu$ as a function of $\varepsilon_p$, the relation between these two variables is:

$$\frac{\partial \mu(\varepsilon_p)}{\partial \varepsilon_p} = \frac{-1}{(\varepsilon_p - 1)^2}.$$

From this, we conclude that the markup increases if the price elasticity decreases. As a corollary, if a parameter changes and decreases the price elasticity, the markup would increase. This implies that the impact on markups depends on whether the change in a parameter makes the demand more inelastic (higher markup) or more elastic (lower markups). Based on this, the assumptions we make about how $\varepsilon_p$ is impacted when $p$ or $\alpha$ varies completely determine how markups are affected.

So far, we have only supposed that $\frac{\partial \varepsilon_p(p;\alpha)}{\partial \alpha} \leq 0$, and therefore:

$$\frac{\partial \mu \left(p^*;\alpha\right)}{\partial \alpha} = \frac{-1}{\left(\varepsilon_p - 1\right)^2} \frac{\partial \varepsilon_p \left(p^*;\alpha\right)}{\partial \alpha} \geq 0.$$

Thus, increases in the good's appeal determine a higher markup, because it makes the demand more inelastic.

Now, let's consider how variations in prices affect the price elasticity and, hence, markups. Formally,

$$\frac{\partial \mu \left(p^*;\alpha\right)}{\partial p} = \frac{-1}{\left(\varepsilon_p - 1\right)^2} \frac{\partial \varepsilon_p \left(p^*;\alpha\right)}{\partial p},$$

which determines that there is a negative relation between $\frac{\partial \mu(p^*;\alpha)}{\partial p}$ and $\frac{\partial \varepsilon_p(p^*;\alpha)}{\partial p}$. It is not obvious what the sign of $\frac{\partial \varepsilon_p(p^*;\alpha)}{\partial p}$ should be. Consistent with the results we want to get below, we suppose that

$$\frac{\partial \varepsilon_p \left(p^*;\alpha\right)}{\partial p} < 0,$$

which implies that

$$\frac{\partial \mu \left(p^*;\alpha\right)}{\partial p} > 0,$$

and, consequently, firms that charge a higher price set a higher markup or, equivalently, firms that charge a low price have a lower markup. One way to interpret this result is to think about income-constrained consumers. Presumably, richer people are less sensitive to price increases. Moreover, an increase in price makes a part of the poor people not being able to afford it. Due to this, price increases can determine that rich people get a greater weight in aggregate demand. Thus, the price elasticity of the aggregate demand would be lower, determining that price increases allow the firm to raise its markup.

The second assumption we make is that, even though $\frac{\partial \mu(p^*;\alpha)}{\partial p} > 0$, the following holds:

$$1 - \frac{\partial \ln \mu \left(p;\alpha\right)}{\partial \ln p} > 0. \tag{10.1}$$

The assumption $1-\frac{\partial \ln \mu(p;\alpha)}{\partial \ln p}>0$ can be rewritten in terms of the price elasticity by asking for the following inquality: $\varepsilon_p(p;\alpha)-1>-\frac{\partial \ln \varepsilon_p(p;\alpha)}{\partial \ln p}$. To show that $1-\frac{\partial \ln \mu(p;\alpha)}{\partial \ln p}>0 \Leftrightarrow \varepsilon_p(p;\alpha)-1>-\frac{\partial \ln \varepsilon_p(p;\alpha)}{\partial \ln p}$, let's begin by showing that $\frac{\partial \ln \mu(p;\alpha)}{\partial \ln p}=-\frac{\partial \ln \varepsilon_p(p;\alpha)}{\partial \ln p}\frac{1}{\varepsilon_p(p;\alpha)-1}$. This follows because:
$\frac{\partial \mu(p^*;\alpha)}{\partial p}=\frac{-1}{(\varepsilon_p-1)^2}\frac{\partial \varepsilon_p(p^*;\alpha)}{\partial p}$ which by multiplying both sides by $\frac{p^*}{\mu(p^*;\alpha)}$ becomes $\frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}=\frac{p^*}{\mu(p^*;\alpha)}\frac{-1}{(\varepsilon_p-1)^2}\frac{\partial \varepsilon_p(p^*;\alpha)}{\partial p}$.
By using the definition of $\mu(p^*;\alpha)$, then $\frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}=\frac{p^*}{\varepsilon_p(p^*;\alpha)}\frac{-1}{(\varepsilon_p-1)}\frac{\partial \varepsilon_p(p^*;\alpha)}{\partial p}$ and so the result follows.
Once that we know this, then $1-\frac{\partial \ln \mu(p;\alpha)}{\partial \ln p}>0 \Leftrightarrow 1+\frac{\partial \ln \varepsilon_p(p;\alpha)}{\partial \ln p}\frac{1}{\varepsilon_p(p;\alpha)-1}>0$ or, just, $\varepsilon_p(p;\alpha)-1>-\frac{\partial \ln \varepsilon_p(p;\alpha)}{\partial \ln p}$.

What is the justification for Assumption (10.1)? It is necessary to get results in line with the taxonomy of firms' strategies we want. This is consistent with the approach we have used so far, where we only add assumptions to have unambiguous effects in the comparative statics analysis.

However, the assumption could also be justified by showing its relevance on other grounds. In particular, we could use the argument that Assumption (10.1) evaluated at the optimal price $p^*$ ensures the second-order condition and uniqueness of the equilibrium.

**Summary of the Asssumptions**

- $\varepsilon_p(p;\alpha)>1$ for any $(p;\alpha)$ (elastic demand at any point)
- $\alpha$ is appeal:
  - $\varepsilon_\alpha(p;\alpha)>0$ (increases of $\alpha$ boost demand)
  - $\frac{\partial \varepsilon_p(p;\alpha)}{\partial \alpha}\leq 0$ (greater $\alpha$ makes demand more inelastic/less elastic)
- $\frac{\partial \varepsilon_p(p^*;\alpha)}{\partial p}<0$, which implies $\frac{\partial \ln \mu(p;\alpha)}{\partial \ln p}>0$ (for definite CS)
- $1-\frac{\partial \ln \mu(p;\alpha)}{\partial \ln p}>0$ (for definite CS)

### 10.3.2 Variations in $c$

There are two equivalent ways to perform CS. Both proceed in a similar fashion. They require differentiating the equilibrium conditions, which in our case are given by the FOC (PRI-1). The methods differ regarding whether we use the solution in the FOC as a *function* of the parameters or as a *value*. Let's state each method and explain this more clearly.

**Method 1 of CS**

Step 1. Start from the equilibrium conditions of the model. Evaluate them at the optimal values.

Step 2. Differentiate the equilibrium conditions, allowing for changes in the endogenous variables and the parameter of interest.

Step 3. Work out the expressions to obtain the derivative of each endogenous variable with respect to the parameter.

We know that the FOC is given by $p^* = \mu(p^*;\alpha)\,c$, which implicitly provides the solution $p^*(\alpha, c)$. Step 1 in method 1 uses $p^*$ (a value) rather than $p^*(\alpha, c)$ (a function).

Let's first consider the case where the parameter of interest is $c$. Our goal is to identify the impact on the endogenous variable $p^*$.Differentiating equation (PRI-1) with $\mathrm{d}p^* \neq 0$ and $\mathrm{d}c \neq 0$,

$$\frac{\partial p^*(\alpha, c)}{\partial c} = \frac{\mu(p^*;\alpha)}{1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}} > 0, \tag{PRICE-$c$}$$

where we have used Assumption (10.1) to determine that $\frac{\partial p^*(\alpha,c)}{\partial c} > 0$. From this we conclude that more efficient firms (lower $c$) charge lower prices.

The FOC is $p^* = \mu(p^*;\alpha)\,c$ and differentiating it with $\mathrm{d}p^* \neq 0$ and $\mathrm{d}c \neq 0$:

$$\left[1 - \frac{\partial \mu(p^*;\alpha)}{\partial p} c\right] \mathrm{d}p^* = \mu(p^*;\alpha)\,\mathrm{d}c$$

which implies that $\frac{\partial p^*(\alpha,c)}{\partial c} = \frac{\mu(p^*;\alpha)}{1 - \frac{\partial \mu(p^*;\alpha)}{\partial p} c}$.

Let's now show that $\frac{\partial \mu(p^*;\alpha)}{\partial p} c = \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}$. Given $\frac{\partial \mu(p^*;\alpha)}{\partial p} c$, we know $p^* = \mu(p^*;\alpha)\,c$ by (PRI-1), and so we can substitute $c$ for $\frac{p^*}{\mu(p^*;\alpha)}$. This determines that $\frac{\partial \mu(p^*;\alpha)}{\partial p} c = \frac{\partial \mu(p^*;\alpha)}{\partial p} \frac{p^*}{\mu(p^*;\alpha)}$. Then, since $\frac{\partial \mu(p^*;\alpha)}{\partial p} \frac{p^*}{\mu(p^*;\alpha)} = \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}$, we get $\frac{\partial p^*(\alpha,c)}{\partial c} = \frac{\mu(p^*;\alpha)}{1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}}$, which is positive since Assumption (10.1) is $1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p} > 0$.

The second method requires treating the optimal solution as a function of the parameters, rather than a specific value.

**Method 2 of CS**

Step 1. Start from the equilibrium conditions of the model. Evaluate them at the optimal solutions expressed as functions of the parameters.

Step 2. Take derivatives of the equilibrium conditions with respect to the parameter of interest.

Step 3. Work out the expressions to get the derivative of each endogenous variable with respect to the parameter.

The FOC evaluated at the optimal solution as a function of the parameters becomes $p^*(\alpha, c) = \mu[p^*(\alpha, c); \alpha]c$. This equation is now only a function of $\alpha$ and $c$. To identify the result, we need to take the derivative of this expression with respect to the parameter of interest. Methods 1 and 2 provide the same result, which is given by equation (PRICE-$c$).

The FOC is $p^* = \mu(p^*; \alpha)c$, and hence $p^*(\alpha, c) = \mu[p^*(\alpha, c); \alpha]c$ by treating $p^*$ as a function. Taking the derivative with respect to $c$ then,
$\frac{\partial p^*(\alpha,c)}{\partial c} = \left[\frac{\partial \mu[p^*(\alpha,c);\alpha]}{\partial p}c\right]\frac{\partial p^*(\alpha,c)}{\partial c} + \mu[p^*(\alpha, c); \alpha]$, which gives
$\frac{\partial p^*(\alpha,c)}{\partial c} = \frac{\mu[p^*(\alpha,c);\alpha]}{1-\frac{\partial \mu[p^*(\alpha,c);\alpha]}{\partial p}c}$. In the derivation with the method 1 of CS we have shown that $\frac{\partial \mu[p^*(\alpha,c);\alpha]}{\partial p}c = \frac{\partial \ln \mu[p^*(\alpha,c);\alpha]}{\partial \ln p}$, and so the result follows.

Once we have determined the effect of variations in $c$ on prices, we can identify the effect of $c$ on quantities and markups. Optimal quantities are given by $q^*[p(\alpha, c); \alpha]$. Thus,

$$\frac{\mathrm{d}q^*[p^*(\alpha, c); \alpha]}{\mathrm{d}c} = \underbrace{\frac{\partial q(p^*; \alpha)}{\partial p}}_{-}\underbrace{\frac{\partial p^*(\alpha, c)}{\partial c}}_{+} < 0. \tag{QUANT-$c$}$$

The result is intuitive. Less efficient firms (i.e. firms with greater marginal costs) set a higher price, and consequently sell less.

As far as markups, optimal markups are $\mu[p^*(\alpha, c); \alpha]$, and so

$$\frac{\mathrm{d}\mu^*[p^*(\alpha, c); \alpha]}{\mathrm{d}c} = \underbrace{\frac{\partial \mu(p^*; \alpha)}{\partial p}}_{+}\underbrace{\frac{\partial p^*(\alpha, c)}{\partial c}}_{+} > 0. \tag{MK-$c$}$$

As a corollary, more efficient firms charge a lower markup.

Overall, we have determined **that less productive firms charge higher prices, sell less, and charge higher markups**. Equivalently, more productive firms charge lower prices, sell more quantity, and charge lower markups.

### 10.3.3 Variations in $\alpha$

Let's consider variations in $\alpha$ and perform comparative statics by using Method 1. Differentiating equation (PRI-1) with $\mathrm{d}p^* \neq 0$ and $\mathrm{d}\alpha \neq 0$,

$$\frac{\partial \ln p^*(\alpha, c)}{\partial \ln \alpha} = \frac{\frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln \alpha}}{1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}} \geq 0. \tag{PRICE-$\alpha$}$$

Since $p^*, \alpha > 0$, then $\frac{\partial \ln p^*(\alpha,c)}{\partial \ln \alpha} \geq 0$ iff $\frac{\partial p^*(\alpha,c)}{\partial \alpha} \geq 0$. Thus, a greater appeal make a firm charge a weakly higher price.

We use the method 1 of CS. The FOC is $p^* = \mu(p^*;\alpha)\, c$, and differentiating it under $\mathrm{d}p^* \neq 0$ and $\mathrm{d}\alpha \neq 0$:

$$\left[1 - \frac{\partial \mu(p^*;\alpha)}{\partial p} c\right] \mathrm{d}p^* = \frac{\partial \mu(p^*;\alpha)}{\partial \alpha} c\, \mathrm{d}\alpha$$

which implies that $\frac{\partial p^*(\alpha,c)}{\partial \alpha} = \frac{\frac{\partial \mu(p^*;\alpha)}{\partial \alpha} c}{1 - \frac{\partial \mu(p^*;\alpha)}{\partial p} c}$. Regarding the denominator, we have already shown in the derivation of (PRICE-$c$) that $\frac{\partial \mu(p^*;\alpha)}{\partial p} c = \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}$. As for the numerator, using that $c = \frac{p^*}{\mu(p^*;\alpha)}$, then $\frac{\partial \mu(p^*;\alpha)}{\partial \alpha} c = \frac{\partial \mu(p^*;\alpha)}{\partial \alpha} \frac{p^*}{\mu(p^*;\alpha)}$, which equals $\frac{\partial \ln \mu(p^*;\alpha)}{\partial \alpha} p^*$.

All these results determine that

$$\frac{\partial p^*(\alpha, c)}{\partial \alpha} = \frac{\frac{\partial \ln \mu(p^*;\alpha)}{\partial \alpha} p^*}{1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}}$$

Multiplying both sides by $\alpha$, then $\frac{\partial p^*(\alpha,c)}{\partial \alpha} \alpha = \frac{\frac{\partial \ln \mu(p^*;\alpha)}{\partial \alpha} \alpha p^*}{1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}}$, and dividing both sides by $p^*$, then $\frac{\partial p^*(\alpha,c)}{\partial \alpha} \frac{\alpha}{p^*} = \frac{\frac{\partial \ln \mu(p^*;\alpha)}{\partial \alpha} \alpha}{1 - \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln p}}$. Since $\frac{\partial \ln \mu(p^*;\alpha)}{\partial \alpha} \alpha = \frac{\partial \ln \mu(p^*;\alpha)}{\partial \ln \alpha}$ and $\frac{\partial p^*(\alpha,c)}{\partial \alpha} \frac{\alpha}{p^*} = \frac{\partial \ln p^*(\alpha,c)}{\partial \ln \alpha}$, the result follows.

To understand why prices are increasing in appeal, keep in mind that when a firm sells a product with more appeal, it faces a more inelastic demand. Hence, the firm can increase its price, and yet the quantities sold are not heavily affected. Thus, it is optimal to charge a higher price.

Regarding optimal quantities $q\left[p^{*}\left(\alpha,c\right);\alpha\right]$:

$$\frac{\mathrm{d}q\left[p^{*}\left(\alpha,c\right);\alpha\right]}{\mathrm{d}\alpha}=\underbrace{\frac{\partial q\left(p^{*};\alpha\right)}{\partial\alpha}}_{+}+\underbrace{\frac{\partial q\left(p^{*};\alpha\right)}{\partial p}}_{-}\underbrace{\frac{\partial p\left(\alpha,c\right)}{\partial\alpha}}_{+\text{ or }0}\gtreqless 0, \tag{QUANT-$\alpha$}$$

and so the effect has an ambiguous sign.

> **Remark** *Observe the distinction between total and partial derivatives. Our assumptions refer to* partial *derivatives. But, in equilibrium, all the variables are changing at the same time. Thus, when we analyze the sign of* $\frac{\mathrm{d}q[p(\alpha,c);\alpha]}{\mathrm{d}\alpha}$*, quantities are affected directly by a greater appeal, but also indirectly by the effect that appeal has on prices.*

The ambiguous effect of appeal on quantities occurs because two opposing effects are working simultaneously. First, there is a positive direct effect, where more appeal increases the demand for the good. However, there is also a negative indirect effect in case appeal impacts the price elasticity. If appeal in particular turns the demand more inelastic, the firm would have incentives to increase its price, thus reducing its demand. Overall, depending on which effect dominates, demand can increase or decrease. Notice that if appeal does not affect the price elasticity, the quantity demanded would necessarily be greater.

For future references, we distinguish between cases depending on the total effect of appeal on quantities:

$$\text{Case I of (QUANT-}\alpha\text{): } \frac{\mathrm{d}q\left[p^{*}\left(\alpha,c\right);\alpha\right]}{\mathrm{d}\alpha}>0,$$

$$\text{Case II of (QUANT-}\alpha\text{): } \frac{\mathrm{d}q\left[p^{*}\left(\alpha,c\right);\alpha\right]}{\mathrm{d}\alpha}<0.$$

Concerning the effects on markups:

$$\frac{\mathrm{d}\mu\left[p^{*}\left(\alpha,c\right);\alpha\right]}{\mathrm{d}\alpha}=\underbrace{\frac{\partial\mu\left(p^{*};\alpha\right)}{\partial\alpha}}_{+}+\underbrace{\frac{\partial\mu\left(p^{*};\alpha\right)}{\partial p}}_{+}\underbrace{\frac{\partial p^{*}\left(\alpha,c\right)}{\partial\alpha}}_{+}>0. \tag{MK $-\ \alpha$}$$

Intuitively, the result arises since there is a one-to-one relation between the sign of $\mu$ and of $\varepsilon_p$, where a decrease in the price elasticity results in greater markups. And a greater appeal always turns the demand more inelastic, due to its direct effect and its indirect effect through $p^*$ (increases in prices make the demand more inelastic.

In elasticity terms, it can be even shown that $\frac{\mathrm{d}\mu[p^*(\alpha,c);\alpha]}{\mathrm{d}\alpha} = \frac{\partial \ln p^*(\alpha,c)}{\partial \ln \alpha}$. To see this, optimal prices imply that $\mu\left[p^*(\alpha,c);\alpha\right] = \frac{p^*(\alpha,c)}{c}$. where $\mu^*(\alpha,c) :=: \mu\left[p^*(\alpha,c);\alpha\right]$. Notice that $\frac{\partial \mu^*(\alpha,c)}{\partial \ln \alpha} = \frac{\mathrm{d}\mu[p^*(\alpha,c);\alpha]}{\mathrm{d}\ln \alpha}$. And since $\ln \mu^*(\alpha,c) = \ln p^*(\alpha,c) + \ln c$, then $\frac{\partial \ln \mu^*(\alpha,c)}{\partial \ln \alpha} = \frac{\partial \ln p^*(\alpha,c)}{\partial \ln \alpha}$.
Try to understand the difference between the total and partial derivatives. The term $\frac{\partial \mu^*(\alpha,c)}{\partial \ln \alpha}$ is only function of $\alpha$, and it captures the direct effect of $\alpha$ on $\mu$ as well as the indirect effect of $\alpha$ on prices.

## 10.4 What Makes A Firm Successful?

Remember that our analysis had the ultimate goal of identifying what makes a firm successful. In terms of the model, this means that the firm garners a high profit. To accomplish this, we distinguish between the different types of successful firms we can conceive. These types are established according to the specific vector $(\alpha, c)$ a firm has. The parameter $\alpha$ provides information about how popular the good is and the price the consumers are willing to pay for it. Regarding the parameter $c$, it indicates the firm's efficiency to produce the good, and hence its cost.

We begin by showing that, indeed, greater appeal or more efficiency make a firm garner a higher profit. To show this formally, recall that a firm's optimal profit is

$$\pi^*(\alpha,c) := Q\left[p^*(\alpha,c),\alpha\right]\left[p^*(\alpha,c) - c\right]$$

If we want to know how $\pi^*$ varies when one of the parameters ($\alpha$ or $c$) changes, we can apply the Envelope Theorem. This theorem allows us to compute the impact on the value function due to change in a parameter. Thus,

$$\frac{\partial \pi^*(\alpha,c)}{\partial \alpha} = \frac{\partial Q(p^*;\alpha)}{\partial \alpha} > 0,$$
$$\frac{\partial \pi^*(\alpha,c)}{\partial c} = -Q(p^*;\alpha) < 0.$$

The result states that successful firms have high $\alpha$ and/or a low $c$.

Notice we have treated $c$ and $\alpha$ as parameters, even when a firm's appeal and productivity is partly decided by the company. Actually, our results apply to this case too. Treating $c$ and $\alpha$ as parameters can be understood as the final outcome of a well-defined model. For instance, $c$ and $\alpha$ could be the outcome in a model where firms have different abilities to differentiate their product or reduce their costs.

Formally, we can respectively define these skills by $\varphi^{\alpha}$ and $\varphi^{c}$. In a model like this, each firm having a value $(\varphi^{c},\varphi^{\alpha})$ makes choices $c^*(\varphi^{c},\varphi^{\alpha})$ and $\alpha^*(\varphi^{c},\varphi^{\alpha})$. Given heterogeneity of firms in industry, different combinations $(c^*,\alpha^*)$ will arise in equilibrium. By treating $(c,\alpha)$ as parameters and asking how their variations affect a firm's decisions, we are in fact asking how firms with a greater $\varphi^{c}$ (lower $c$) or a greater $\varphi^{\alpha}$ (higher $\alpha$) make their choices.

According to the pair $(\alpha, c)$, firms will behave differently in the market. According to Michael Porter, a famous academic specialized in business, there are three strategies that firms can pursue to be successful. He refers to them as "Generic Competitive Strategies" and comprise:

[1] **Overall cost leadership**

*Examples*: Walmart and Costco (retailers), RyanAir and EasyJet (airlines), Ikea (furniture), H&M (apparel).

[2] **Differentiation**

*Examples*: Nike and Adidas (sport clothes) Coke and Pepsi (carbonated beverages), Duracell and Energizer (batteries) Bayer and Pfizer (pharmaceutical products), Apple (computers).

[3] **Focus**

*Examples*: Ferrari, BMW and Mercedes Benz (cars), Louis Vuitton and Gucci (apparel), Dom Pérignon (champagne), TAG Heuer (clocks).

Next, we show how each category can be reflected by appropriate choices of $\alpha$ and $c$.

### 10.4.1 What are the Strategies that a Successful Firm Follows?

We have shown that firms with a lower $c$ or greater $\alpha$ have higher profits. Moreover, the analysis of the model has established the following results:

**Summary of the Results**

**Variations in $c$**

- $\frac{\partial \pi^*(\alpha,c)}{\partial c} < 0$

- $\frac{\partial p^*(\alpha,c)}{\partial c} > 0$
- $\frac{\mathrm{d}q^*[p^*(\alpha,c);\alpha]}{\mathrm{d}c} < 0$
- $\frac{\mathrm{d}\mu^*[p^*(\alpha,c);\alpha]}{\mathrm{d}c} > 0$

**Variations in $\alpha$**

- $\frac{\partial \pi^*(\alpha,c)}{\partial \alpha} > 0$
- $\frac{\partial p^*(\alpha,c)}{\partial \alpha} > 0$
- $\frac{\mathrm{d}q[p^*(\alpha,c);\alpha]}{\mathrm{d}\alpha} \gtreqless 0$
  - Case I: $\frac{\mathrm{d}q[p^*(\alpha,c);\alpha]}{\mathrm{d}\alpha} > 0$
  - Case II: $\frac{\mathrm{d}q[p^*(\alpha,c);\alpha]}{\mathrm{d}\alpha} < 0$
- $\frac{\mathrm{d}\mu[p^*(\alpha,c);\alpha]}{\mathrm{d}\alpha} > 0$

With this information, we can now establish Porter's taxonomy:

[1] **Overall Cost Leadership**: it comprises firms with a lower $c$. Relative to other firms in the industry, they have a high $q^*$, a low $p^*$, and a low $\mu^*$.

[2] **Differentiation**: it comprises firms with a high $\alpha$ and satisfying Case I of (QUANT-$\alpha$). Relative to other firms in the industry, they have a high $q^*$, a high $p^*$, and a high $\mu^*$

[3] **Focus**: it comprises firms with a high $\alpha$ and satisfying Case II of (QUANT-$\alpha$). Relative to other firms in the industry, they have a low $q^*$, a high $p^*$, and a high $\mu^*$.

The profit function enables us to show how these strategies are reflected in a firm's features. There are two ways in which we can reexpress optimal profits. First,

$$\pi^*(\alpha, c) := \frac{R[p^*(\alpha, c), \alpha]}{\varepsilon_p[p^*(\alpha, c), \alpha]}. \tag{PROF1}$$

Let's indicate optimal variables without arguments and with a * as a superscript. Optimal profits are $\pi^*(\alpha, c) := Q^*(p^* - c)$. By the FOC, $p^* = \frac{\varepsilon^*}{\varepsilon^*-1}c$ and so by subtracting $c = \frac{\varepsilon^*-1}{\varepsilon^*}p^*$. Thus, optimal profits are $\pi^*(\alpha, c) := Q^*\left(p^* - \frac{\varepsilon^*-1}{\varepsilon^*}p^*\right)$ or, just $\pi^*(\alpha, c) := Q^* p^*\left(1 - \frac{\varepsilon^*-1}{\varepsilon^*}\right)$ which determines the result.

Moreover, by the mere definition of a profit function, we can divide and multiply by $c$ and obtain:

$$\pi\left(p^{*};\alpha,c\right)=\underbrace{cQ\left(p^{*};\alpha\right)}_{=:(1)}\underbrace{\left[\mu\left(p^{*};\alpha\right)-1\right]}_{=:(2)}, \qquad \text{(PROF2)}$$

where we have used that $\frac{p^{*}}{c}=\mu\left(p^{*};\alpha\right)$

Equation (PROF2) indicates that one way to garner high profits is by charging a lower markup (low $\mu$, and so a small term (2)) and having a great scale of production (high $Q$, and so a big term (1)). A firm that deploys this strategy is identified as being massive and cheap. For this to occur, two conditions have to be met: the firm's efficiency has to be substantially high and consumers have to be price sensitive. The last condition follows by (PROF1), which shows that consumers need to have a quite elastic demand. Basically, these aspects determine that a low price attracts a significant part of the consumers, and that a firm is capable of setting this low price. Overall, the firm would get high revenues, since the pronounced quantity sold more than compensates for the low price charged. This strategy corresponds to a cost leadership strategy in terms of Porter's taxonomy.

At the other extreme, (PROF2) reveals that a firm can get high profits by having low sales (and hence a small term (1)), but charging high markups for each unit sold (high $\mu$, and so a big term (2)). This type of firm focuses on niche markets, where customers have a high willingness to pay for some distinctive features of the good. In terms of equation (PROF1), the strategy requires increasing a good's appeal to make $\varepsilon_p$ decrease significantly. Thus, these firms can have lower total revenues relative to other firms in the industry. Nonetheless, they could have substantially greater profits due to a high price and savings in production costs by low production. This is the strategy "focus" in Porter's taxonomy.

Finally, we can conceive firms with high profits by balancing their total sales and the prices/markups charged. These firms set high prices relative to other firms, but do not focus on a niche market. Rather, they try to be wide in terms of the consumers reached. The strategy requires increasing the good's appeal, but without doing it to such an extent that only a few consumers can afford it. In terms of Porter's taxonomy, this corresponds to the differentiation case.

## 10.5 Exercises

[1] Let us focus on a firm that is successful due to demand appeal, rather than the cost side (i.e. efficiency). We'll choose a specific functional form for demand, and consider two different parameters ($A$ and $\sigma$) that represent appeal: both boost demand, but one does not affect the price elasticity.

Consider that demand is $q(p) := Ap^{-\sigma}$, with parameters $A > 0$ and $\sigma > 1$. The firm produces with a technology exhibiting CRS, with unitary cost $c > 0$.

(a) Establish the price elasticity of demand (hint: use the log definition). Does it depend on $A$?

(b) Determine the firm's optimal price and quantities.

(c) Show that increases in $A$ and decreases in $\sigma$ increase the firm's profit (hint: use the Envelope Theorem. It'll also help work in logs).

(d) Suppose that the firm is considering two types of investments. Investment 1 increases $A$, while Investment 2 decreases $\sigma$. To keep matters simple, suppose that these investments do not entail additional costs to be implemented. In question c), you determined that both types of investments increase the firm's profit. Now, we establish why this occurs. Higher profits can arise because a firm is selling more, charging a higher price, or both. Establish which channel operates with Investment 1 and 2.

(e) Suppose that you're working on your thesis, and have to identify the investment type used by a particular firm in two scenarios. Determine for the following situations if you would capture the situation by Investment 1 or 2.

i. The firm has established new channels of product distribution, and also enhanced the existent ones. This makes the firm's product more widely available, thereby allowing it to reach more people.

ii. The firm has overhauled the features of its product. This has not only boosted the firm's sales, but also increased a customer's willingness to pay for the product.

[2] Let's consider the monopoly model of the lecture note, but modified in one respect: increasing a good's appeal entails greater marginal costs. This makes "appeal" be more in line with a quality definition. For instance, a cell phone becomes more appealing if it has a faster processor, has more gigabytes of RAM, is made with more resistant materials, etc. But those features also increase the cost per unit produced.

To capture this, we add marginal costs $c(\alpha)$ with $c'(\alpha) > 0$ to the baseline model.

(a) Characterize the optimal prices, $p^*[c(\alpha), \alpha]$.

(b) Show that $\frac{\partial \ln p^*[c(\alpha),\alpha]}{\partial \ln \alpha} = \frac{\frac{\partial \ln \mu^*}{\partial \ln \alpha} + \frac{\mathrm{d} \ln c(\alpha)}{\mathrm{d} \ln \alpha}}{1 - \frac{\partial \ln \mu^*}{\partial \ln p}}$. Interpret the result and compare it with the baseline case. Intuitively, why does its sign do not change with respect to the baseline case?

(c) Does a higher $\alpha$ have the same qualitative effect on markups and quantities as in the baseline model? Check this and explain your result.

(d) Given all your answers, consider Porter's taxonomy for successful firms due to the demand side ("differentiation" and "focus"). Is this still valid? Justify your answer.

**Answer Keys for Some of the Exercises**:

1) In terms of the baseline model, the exercise is just providing a functional form to capture two types of appeal parameters. Both satisfy $\frac{\partial q}{\partial \alpha} > 0$, but one is such that $\frac{\partial \varepsilon_p}{\partial \alpha} = 0$ and the other satisfies $\frac{\partial \varepsilon_p}{\partial \alpha} < 0$. The aim of the exercise is that you can identify which case corresponds to $A$ and which to $\sigma$.

2) *Qualitatively*, nothing changes.

# Lecture Note 11

# Multiproduct Firms

## 11.1 Roadmap

In the previous lecture note, we started our analysis of a firm's decisions. Since the aim was identifying what makes a firm successful, we made some simplifying assumptions. In particular, we assumed that the firm under analysis was supplying only one good.

In this and the next lecture note, we delve into different strategies that the firm could pursue to increase its profit. In this note in particular, we analyze the case of multiproduct firms. This requires us to study how firms make price decisions when they incorporate the interdependence between the different goods. We consider the possibility that goods are either substitutes or complements, whose implications are different.

## 11.2 Multiproduct Firms

We keep considering the case of one firm in a specific industry in isolation. The firm makes price decisions relative to two goods, labeled 1 and 2. Aside from the existence of two goods, the setup is similar to that used in previous lecture notes. In particular, we suppose the firm has a production technology with constant marginal costs, $C(q_i) := c_i q_i$, and the price elasticity of $i$ satisfies $\varepsilon_i > 1$, where good $i = 1, 2$.

Goods 1 and 2 display **one-way complementarity/substitution**. This occurs when their demands are $q_1(p_1, p_2)$ and $q_2(p_2)$ respsectively, so that good 1's demand 1 is influenced by the price of good 2, but the consumption of good 2 only depends on its own price. We say that **good 2 is a substitute for good 1** if $\frac{\partial q_1(p_1,p_2)}{\partial p_2} > 0$, whereas **good 2 is a complement for good 1** if $\frac{\partial q_1(p_1,p_2)}{\partial p_2} < 0$. We can equivalently define this relation through the **cross price elasticity** of good 1 with respect to good 2, which is formally defined by $\varepsilon_{12}(p_1, p_2) := \frac{\partial \ln q_1(p_1,p_2)}{\partial \ln p_2}$ and satisfies

$$\operatorname{sgn}\left[\varepsilon_{12}(p_1, p_2)\right] = \operatorname{sgn}\left[\frac{\partial q_1(p_1, p_2)}{\partial p_2}\right].$$

Thus, good 2 is a substitute for good 1 when $\varepsilon_{12} > 0$, while good 2 is a complement for good 1 when $\varepsilon_{12} < 0$.

For now, we will solve the firm's optimization problem without making any assumption about the relation between goods—$\varepsilon_{12}$ could be positive or negative. The optimization problem is given by

$$\max_{p_1,p_2} \pi(p_1,p_2) = q_1(p_1,p_2)(p_1 - c_1) + q_2(p_2)(p_2 - c_2)$$

where we suppose that $p_i \in [0,\infty)$ for each good $i$. For future references, we define $\pi_i(p_1,p_2) := q_i(p_i - c_i)$ and $R_i := p_i q_i$ as the profit and revenue of good $i$, respectively.

**Remark** *The assumption that $p_i \in [0,\infty)$ for each $i$ implies that negative prices are not allowed. If it is optimal to set a negative price for a range of the parameters, we suppose that the firm sets a price equal to zero. In the context of multiproduct firms, negative prices should not be ruled out, since it is possible that the price of a good may be lower than its marginal cost, as we will see below.*

This model also could capture the decision of whether to introduce a specific good, by interpreting $p_i = \infty$ as the decision of not selling good $i$. At an infinite price, the consumption would necessarily be zero, which is equivalent to saying that the firm does not want to sell the good.

Considering good 1's optimal price in isolation, the only difference with the case of a single-product firm is that $q_1$ now depends on one more variable in addition to $p_1$, which is $p_2$. Nonetheless, the mathematical structure is the same, since $\pi_2$ does not depend on $p_1$. Due to this, the optimal price of good 1 satisfies the following equation:

$$p_1^* = \frac{\varepsilon_1(p_1^*,p_2)}{\varepsilon_1(p_1^*,p_2) - 1} c_1 \tag{11.1}$$

where we denote the implicit solution of $p_1^*$ for some value of $p_2$ by $p_1^*(p_2)$.

The FOC is
$\frac{\partial \pi(p_1,p_2)}{\partial p_1} = \frac{\partial \pi_1(p_1,p_2)}{\partial p_1} = \frac{\partial q_1(p_1,p_2)}{\partial p_1}(p_1 - c_1) + q_1(p_1,p_2) = 0$
which is exactly the same FOC as in the case of a single-product firm, treating $\pi_{11}$ as the profits. Therefore, the same pricing rule holds, with the difference that now the price elasticity also depends on the price of good 2.

Unlike good 1, the firm chooses good 2's price taking into account the effect that $p_2$ has on good 1's demand. Depending on whether good 1 is a substitute or a complement for the good 2, an increase in price 1 might increase or decrease good 2's demand. This is reflected in the FOC for good 2's price, which is:

$$p_2^* = \frac{\varepsilon_2 (p_2^*)}{\varepsilon_2 (p_2^*) - 1} c_2 + \rho_{12} (p_2^*) \tag{P2}$$

where

$$\rho_{12} (p_2^*) := \varepsilon_{12} [p_1 (p_2^*), p_2^*] \, \pi_1 [p_1 (p_2^*), p_2^*] \frac{1}{q_2 (p_2^*) (\varepsilon_2 (p_2^*) - 1)}. \tag{$\rho_{12}$-b}$$

The term $\rho_{12}$ can be reexpressed by rewriting the optimal profit obtained through good 1, given by $\pi_1 = \frac{R_1}{\varepsilon_1}$:

$$\rho_{12} (p_2^*) := \varepsilon_{12} [p_1 (p_2^*), p_2^*] \frac{R_1 [p_1 (p_2^*), p_2^*]}{\varepsilon_1 [p_1 (p_2^*), p_2^*]} \frac{1}{q_2 (p_2^*) (\varepsilon_2 (p_2^*) - 1)} \tag{$\rho_{12}$-a}$$

We skip the arguments of each function to keep the notation as simple as possible. The FOC is given by:

$$\frac{\partial \pi}{\partial p_2} = \underbrace{\frac{\partial q_2}{\partial p_2} (p_2 - c_2) + q_2}_{\text{same as in the single-product case}} + \underbrace{\frac{\partial q_1}{\partial p_2} (p_1 - c)}_{\text{cross effect}} = 0$$

Relative to the case of a single-product firm, there is an additional term given by the cross-effect $\frac{\partial q_1}{\partial p_2} (p_1 - c)$. This captures how the pricing of good 2 impacts good 1.

Now, we work out the expression to make it comparable to equation (11.1).

$\frac{\partial q_2}{\partial p_2} (p_2 - c_2) + q_2 + \frac{\partial q_1}{\partial p_2} (p_1 - c) = 0.$

Multiplying both sides by $-\frac{p_2}{q_2}$, and multiplying and dividing the last term by $q_1$:

$-\frac{\partial q_2}{\partial p_2} \frac{p_2}{q_2} (p_2 - c_2) - \frac{p_2}{q_2} q_2 - \frac{p_2}{q_2} \frac{q_1}{q_1} \frac{\partial q_1}{\partial p_2} (p_1 - c) = 0$

$\Rightarrow p_2 (\varepsilon_2 - 1) - \varepsilon_2 c_2 - \frac{q_1}{q_2} \varepsilon_{12} (p_1 - c) = 0$

Thus, $p_2 = \frac{\varepsilon_2}{\varepsilon_2 - 1} c_2 + \frac{\varepsilon_{12}}{q_2 (\varepsilon_2 - 1)} \pi_1$.

To obtain the second expression, we use (11.1) to reexpress $\pi_1$. From (11.1), we obtain that $p_1 - c = \frac{1}{\varepsilon_1 - 1} c_1$ and also $\frac{p_1}{\varepsilon_1} = \frac{1}{\varepsilon_1 - 1} c_1$. Therefore, $q_1 (p_1 - c) = \frac{q_1 p_1}{\varepsilon_1} = \frac{R_1}{\varepsilon_1}$ since $\pi_1 = q_1 (p_1 - c)$, and the result follows.

With the derivation of the optimal prices, we can now consider the cases of substitutes and complements. This requires analyzing the pricing of good 2, which takes into account the influence of good 2's price on good 1's demand.

To do this, we define a benchmark scenario to compare our results. Let $p_2^{SP}$ be the price of good 2 when goods are independent, i.e. when $\varepsilon_{12} = 0$ and hence $\rho_{12} = 0$. The superscript SP stands for "single product" and reflects that the pricing is the same as a single-product firm when goods are independent. Formally, $p_2^{SP}$ is given by

$$p_2^{SP} = \frac{\varepsilon_2 \left(p_2^{SP}\right)}{\varepsilon_2 (p_2^{SP}) - 1} c_2$$

### 11.2.1 Substitute Goods

Let's suppose that good 2 is a substitute for good 1. Thus, $\frac{\partial q_1(p_1,p_2)}{\partial p_2} > 0$ or, equivalently, $\varepsilon_{12}(p_1, p_2) > 0$. By simple inspection of equation (P2), we can determine that $\rho_{12} > 0$, thereby obtaining the following result:

> **Result 11.1** *Suppose that good 2 is a substitute for good 1. Then, $p_2^* > p_2^{SP}$, so that the optimal price is higher than that set by a single-product firm.*

In the case of substitutes, the term $\rho_{12}$ captures what is known in the literature as the *cannibalization effect*. The name follows because when a firm introduces some good in the market, it steals demand from its other goods. This provides the firm with incentives to increase the price of good 2 (relative to the case where there is no relation between goods), diminishing the cannibalization effect of good 2 on good. Notice that a greater price of good 2 is equivalent to lower sales of good 2. Consequently, the firm produces less of good 2 relative to a single-product firm, due to the cannibalization effect.

The result begs the question of why a firm would launch a product that will compete with itself. We can provide different answers. First, by assuming that the solution is interior as we did.[1] we are implicitly assuming that selling both goods is more profitable than just selling good 2 and avoiding the cannibalization effect. This requires that the loss in good 1's profits 1 due to the cannibalization by good 2 is more than compensated with the total increase in profits from good 2's sales.

Going beyond the setup under consideration, launching a new good could be rationalized in another way: the good steals demand from the company's own products, but also from rival firms. Furthermore, cannibalization is sometimes unavoidable. This is related to the theory of product cycles. After some time, products start to become obsolete or less trendy. In this context, a firm's best strategy is to constantly innovate and launch new products, even though this means stealing demand from its own products.

[1] Keep in mind that if we had a corner solution where the firm sets an infinite price of good 1, this would be equivalent to not introducing the product in the market.

Notice that the higher the cannibalization effect, the more incentives the firm have to increase the price of good 2. Given this, it is worth analyzing when the cannibalization effect is bigger. This question can be answered by inspecting what $\rho_{12}$ depends on through equation ($\rho_{12}$-b).

- **The greater $\varepsilon_{12}$, the greater $\rho_{12}$.** The explanation for this result is trivial: the greater the effect that the price of good 2 has on the demand of good 1, the greater the cannibalization effect. Thus, the firm has more incentives to increase the price of good 2.

- **The greater $\pi_1$, the greater $\rho_{12}$.** Keep in mind that $\pi_1$ refers to the profit stemming exclusively from good 1, not the firm's total profit. The result entails that if good 1 is one of the firm's star products (that is, a good providing high profits), the firm has more incentives to increase the price of good 2. A corollary of this is that good 2's quantities produced would be lower too. Using that $\pi_1 = \frac{R_1}{\varepsilon_1}$ as in equation ($\rho_{12}$-a), we can also conclude that the cannibalization effect is bigger when the revenues due good 1 are greater or good 1 is more price inelastic.

### 11.2.2 Complementary Goods

Let's now suppose that good 2 is a complement for good 1. Thus, $\frac{\partial q_1(p_1,p_2)}{\partial p_2} < 0$ or, equivalently, $\varepsilon_{12}(p_1,p_2) < 0$. Simple inspection of (P2) allows us to conclude that $\rho_{12} < 0$, thereby obtaining the following result.

> **Result 11.2** *Suppose that good 2 is a complement for good 1. Then, $p_2^* < p_2^{SP}$,so that the optimal price is lower than that set by a single-product firm.*

Intuitively, $\rho_{12} < 0$ captures that increases in good 2's price would reduce not only the own good's demand, but also that of good 1. Complementarity of goods means that greater consumption of one good increases the utility of consuming the other good. The result can be clearly appreciated when demands derive from a Leontief utility function, such that goods are perfect complements. In the case of one-way complementarity, this means that the firm has incentives to decrease the price of good 2, with the aim of boosting the demand for good 1.

Just like we did in the case of substitutes, let's analyze when a firm has more pronounced incentives to decrease the price of good 2.

- **The greater the absolute value of $\varepsilon_{12}$, the greater the absolute value of $\rho_{12}$.** Intuitively, the greater the effect of good 2's price on good 1's demand, the greater the incentives of the firm to decrease good 2's price.

- **The greater $\pi_1$, the greater the absolute value of $\rho_{12}$.** The firm has incentives to decrease the price of good 2 when good 1 gives high profits. In fact, if good 1 were the company's star product, it is possible that *the price of good 2 is set at a level lower than its marginal cost*. A simple example of this is observed in coffee shops, where sugar packets are free. Through the lens of the model presented, this can be rationalized as a strategy to boost coffee consumption, which is the good providing profits to a cafe.

## 11.3 Some Applications

Next, I illustrate strategies deployed by multiproduct firms in real life. For multiproduct firms selling substitute goods, I consider Apple launching the iPhone, despite affecting the sales of the iPod. For complementary goods, I consider Sony and the Playstation as an example.

### 11.3.1 The iPhone and iPod as Substitutes

Apple entered the cell phone industry in 2007, when it launched its first iPhone. At that moment, the iPod (the mp3 player of Apple) was accounting for almost 50% of its total revenues, charging markups of around 40% for this product. On the contrary, the iPhone only represented around 2.5% of its total revenues at the end of 2007.

The iPhone includes similar features to the iPod. In particular, it can be used as an mp3 player, making it a substitute for the iPod. Apple knew that the introduction of the iPhone would then reduce the sales of the iPod, which in our terminology means that the iPhone would partly cannibalize the iPod's sales. Nonetheless, Steve Jobs did this by arguing that "if you don't cannibalize yourself, someone else will."

Even when it was inevitable to reduce the iPod's sales considerably, Apple's expectation was to increase its overall profit. In other words, it expected that the sales of the iPhone would more than compensate for any loss due to cannibalization. Indeed, Apple's strategy was highly successful. These days, the iPod only represents around 2% of Apple's sales, while the iPhone has become the company's star product. In fact, it is possible to say that the iPhone allowed Apple to become one of the most profitable firms in the industry.

Through the lens of the model, notice that there is a one-way substitution relation between the iPod and the iPhone—the iPod is not a substitute for the iPhone, but the iPhone is a substitute for the iPod. Nowadays, the pricing strategy of the iPhone is not highly influenced by the cannibalization effect on the iPod. Mp3 players have lost attraction for consumers (low $\rho_{12}$), and the profits garnered through the iPod are too low to have a pronounced influence (low $\pi_1$).

### 11.3.2 The Playstation and Games as Complements

Sony has consistently sold the Playstation console at a price lower than its marginal cost. To provide specific numbers to this strategy, we focus on the Playstation 3, launched in 2006. Building the console at that time had a cost of around 805 USD, while its retail price was only 599 USD. After some time, Sony improved its efficiency in production, thereby reducing its costs. Nonetheless, the strategy was still deployed: by late 2009, the Playstation 3 was sold for 299 USD, even though its cost was 336 USD.[2]

[2] Nowadays, there is evidence that Sony is not selling the new generations of Playstation to a loss. Nonetheless, the margins of profits per unit sold are really slim. For instance, in 2013, manufacturing each Playstation 4 cost around 381 USD, while it was sold to a price of around 400 USD. There is evidence that Microsoft deploys a similar pricing strategy for its Xbox. For more on the Economics of

The strategy is chiefly explained because Sony's profits regarding the Playstation come from the games sold and the online subscriptions. Specifically, Sony charges a licensing fee for the use of its console to third-party game developers, and also performs in-house production of games, although to a lower extent.

The model we have used can justify Sony's strategy. To illustrate this, think of Playstation as good 2 and games as good 1. A console price lower than its marginal costs is optimal for a firm when two conditions are met: $\rho_{12}$ is negative (good 1 has to be a complement of good 2) and has a big value (great value of $\rho_{12}$ in absolute terms).

We can be confident that this was indeed the case, by noticing the following. First, since having the console is a necessary condition to play any game, the price of the Playstation acts like an entry fee that Sony charges customers to allow them to play games. Thus, the console (good 2) is a complement for the video games (good 1), which implies that $\varepsilon_{12} < 0$ and hence$\rho_{12} < 0$. Moreover, there is evidence that the profits accrued by Sony through games is substantial, which is captured in the model through a high $\pi_1$. Both effects act in the same direction, determining a great absolute value of of $\rho_{12}$.[3]

gaming consoles, you can read this article.

[3]Unlike our baseline model with one-way complementarity, we might think of the Playstation and the games exhibiting a two-way complementarity. This arises if consumers analyze the price of games when they decide which console to buy.

# Lecture Note 12

# Second-Degree Price Discrimination

## 12.1 Roadmap

In this lecture note, we study how firms choose their good's features, by focusing on decisions on quality. Importantly, we use this topic to illustrate the tools of a more general subject of the microeconomic literature, known as ***adverse selection*** (aka ***hidden knowledge***).

We present a model in which there are several types of consumers, according to their willingness to pay for quality. Firms do not know the specific type of each consumer, and even if they do so, they cannot force consumers to buy a specific version of the product. However, firms do know the distribution of preferences in the population, allowing them to design a screening device: the goods are presented in different versions, where each of them is defined as a quality-price pair. The firm designs each version of the good with the goal that consumers self-select and choose the variety created for their type.

This scenario is compared with a baseline situation, where firms have full information about consumers and can force them to choose a specific version of the good. Two results emerge. First, the firm has incentives to downgrade the quality of the version designed for the consumers with the lowest willingness to pay. Furthermore, there is ***no distortion at the top***: the quality-price pair for the consumers with the highest willingness to pay is the same as under full information. Our analysis concludes by providing an application of this strategy, through the so-called damaged goods.

## 12.2 Price Discrimination and Arbitrage

So far, our analysis of firms has been based on several simplifying assumptions. In particular, firms were using the simplest mechanism of pricing we could think of: **uniform pricing**. This type of pricing is such that:

[1] different consumers pay the same price, and

[2] each consumer pays the same price for all the units purchased.

Uniform pricing rules out situations such as discounts by units bought or by type of customer (e.g. students, underage people)—the price is the same for everyone and independently of the amount purchased.

In contrast, we say that there is **price discrimination** when at least one of the two conditions mentioned does not hold.[1] Price discrimination is part of the battery of instruments that a firm has at its disposal to charge prices closer to a consumer's valuation. Typical examples are discounts at the cinemas for retired people or kids below a certain age. It also encompasses discounts for wholesale purchases. In fact, there are a plethora of methods to discriminate prices: two-part tariffs (a fee to have access to the good/service, plus some price for each unit consumed), coupons, differences in prices if the good is to take or stay, etc.

Price discrimination is not always possible. And, even when it is feasible, identifying consumer types may only be performed in a restricted way, depending on the industry characteristics. In particular, its application hinges on

[1] the information about demand that the firm has, and

[2] the possibility of arbitrage.

The first item refers to a consumer's valuation for each unit of the good. This information is required to know how much to charge each consumer, and hence it is can be considered a necessary requirement before engaging in a price discrimination strategy.

Even if the firm knows a consumer's valuation for each unit of the good, it also needs to ensure that the consumer pays the price established by the firm. For instance, this does not occur if reselling is possible: consumers facing lower prices would buy the good and then resell it to consumers that can only buy the good at a higher price.

[1] Some authors define price discrimination in a different way. For instance, they argue that consumers could be paying different prices for other reasons, not related to their valuation. For instance, this encompasses differences in costs depending on the consumer's location, or if two varieties are so different that we can rarely think of them as two versions of the same product. To amend this, they define price discrimination as when two different units of a good with similar costs and features are sold at different prices. The main message of this remark is that there is no definition of price discrimination that is exhaustive and exactly delimits the situations covered.

### 12.2.1 Arbitrage

There are two types of arbitrage: commodity and personal. **Commodity arbitrage** refers to a scenario where reselling is possible. Its emergence crucially depends on the characteristics of the good analyzed. For example, services cannot be usually resold (e.g., a person cannot resell a haircut). When reselling cannot be prevented, the firm is forced to adopt uniform pricing. Thus, it does not matter the identity of the consumer or the units she buys, the price would be the same. Why is this? Because otherwise the price in the market would be the lowest at which the good is available. Any consumer paying the lowest price can profit from the firm discriminating prices, by reselling units to consumers paying higher prices. Once the firm incorporates this into its analysis, it is optimal to maximize profits taking as given that only one price will hold in the market. Thus, it maximizes its profit subject to a uniform pricing strategy.

In some cases, even when reselling is possible, it is too costly. For instance, consider Costco. You can only buy products in Costco, if you pay some fee to become a member. In principle, nothing prevents a group of people from paying one membership and then each shopping by using the same membership. However, organizing this can be quite complex and time-demanding.

**Personal arbitrage** can only arise when a firm sells different versions of the same good. In the literature of differentiated goods, each version of the good is alternatively referred to as a ***variety***.[2] The existence of different varieties of the same good assumes implicitly that price is not the only feature that the consumer cares about.

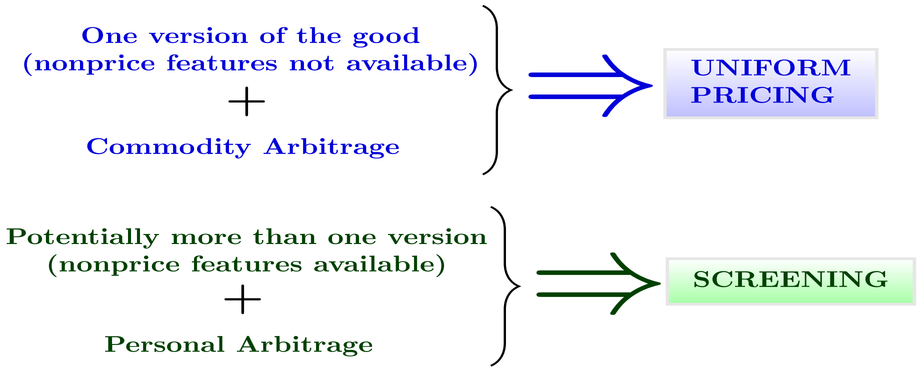
The literature usually distinguishes between non-price features that differentiate a good. **Vertically differentiated aspects** refer to features where a natural order exists. This means that every consumer would agree on what constitutes a more preferred feature. For instance, all consumers would agree that a cell phone with a faster microprocessor or more RAM memory is preferred. This is why vertically differentiated is generally called quality. In addition, there can be **horizontally differentiated aspects,** where consumers do not share the same ranking regarding what is preferred. Continuing with the example of cell phones, consumers have different preferences for colors, or prefer bigger screens at the expense of less portability.

Personal arbitrage is defined as a situation where firms cannot force a specific consumer to buy a specific version of the good. To demonstrate this, suppose that the firm

[2] I will use the terms "varieties" and "versions" of the good interchangeably.

has built two different versions of a good. The goal is that consumers of type A acquire version 1, and consumers of type B buy version 2. When there is personal arbitrage, the firm cannot prevent type-A consumers from choosing version 2 of the good, or type-B consumers from choosing version 1.

Under personal arbitrage, it is optimal for the firm to introduce different versions of the good, so that each consumer self-selects and chooses the version created for her. In the literature of asymmetric information, this is known as a **screening procedure**: the firm chooses each version's features to induce consumers to reveal their own type when they buy the good.




This note considers a model where there is personal arbitrage, and so firms engage in a screening procedure. We derive conclusions by comparing its solution against a baseline situation with no personal arbitrage.

## 12.3 Setup

We keep assuming the existence of only one firm in the industry. Moreover, we consider a good that can be differentiated in terms of quality. In the Industrial-Organization literature, we say that the firm follows a **second-degree price discrimination** strategy when a good is quality differentiated, there is personal arbitrage, and the firm engages in customer screening.

**Remark** *Allowing a firm to choose several quality-price pairs is always more profitable than restricting it to offer only one version. This follows by a revealed-preference argument. The intuition of this result is that, if we ignore strategic interactions,* **giving more options to a firm can never reduce its profit**. *This is easy to see when we analyze the problem mathematically. A firm offering only one version of the good is actually a special case of the two-versions case where the firm produces two identical versions (i.e., with the same quality and price). Ultimately, these versions are indistinguishable for consumers, and so can be considered as the same version of the good.*

### 12.3.1 Consumers

When we studied consumer theory, we were implicitly assuming that each good was available in one version. Due to this, we first need to extend consumer theory to cope with goods available in several qualities.

With this goal, consider a quasilinear utility function defined over good 1 and good 2. We incorporate two modifications to it: there is a parameter that captures a good's quality, and consumers can only buy zero or one unit of the good. The latter is a standard assumption in the literature, but it is actually a simplification—by no means it is necessary to account for quality differences.

Formally, let good 1 be the good under analysis, whose consumption space is $\{0,1\}$. Good 2 is a composite good that represents the rest of the goods in the economy. As we will show, it plays the role of some outside option in case the consumers decides not to buy good 1.

The utility function of each consumer is:

$$U(x_1, x_2; z, \theta) := \theta u(x_1; z) + x_2,$$

where we suppose $u(0; z) = u(x_1; 0) = 0$, $u'_{x_1} > 0$, $u'_z > 0$, and $u''_z < 0$. These assumptions are only added to have a well-defined problem.

The term $\theta u(x_1; z)$ represents the utility derived exclusively from good 1. Likewise, the parameter $z$ represents the quality of good 1, while $\theta$ represents the consumer's subjective valuation of good 1.

Some remarks are in order. First, $\theta$ is similar to the parameter $A$ used in consumer theory to capture appeal under quasilinear preferences. In both cases, a greater value of $\theta$ increases the marginal utility of good 1. This can be seen by noting that $U'_{x_1} := \theta u'_{x_1}(x_1; z)$, and so $\frac{\partial U'_{x_1}}{\partial \theta} = u'_{x_1}(x_1; z) > 0$. But now, ***a greater value of $\theta$ additionally captures a greater valuation for quality***. This is because $U'_z := \theta u'_z(x_1; z) > 0$ and so $\frac{\partial U'_z}{\partial \theta} = u'_z(x_1; z) > 0$, reflecting that a higher value of $\theta$ increases the utility of one unit of quality.

The budget constraint with normalization$p_2 := 1$ entails that $x_2 = Y - p_1 x_1$. Plugging this term into the utility,

$$U(x_1; z, \theta) := \theta u(x_1; z) + Y - p_1 x_1.$$

The agent consumes either one or zero units of good 1. Thus, denoting $u(z) := u(1; z)$, and supposing that $u(0; z) = 0$, the utility function is

$$U(x_1; z, \theta) := \begin{cases} \theta u(z) + Y - p_1 & \text{if } x_1 = 1 \\ Y & \text{if } x_1 = 0 \end{cases}.$$

From the perspective of good 1's industry, $Y$ represents the utility of the consumer's outside option: if she decides not to consume good 1, she spends all her income on good 2 and obtains a utility $\frac{Y}{p_2}$. Given our normalization $p_2 := 1$, this utility is just $Y$. The assumption reflects that when $p_1$ is high enough and/or the quality of good 1 low enough, the consumer might prefer to consume good 2 rather than good 1.

#### 12.3.1.1 The Specific Consumer Setup Used

We have introduced a utility function that accounts for quality differentiated goods. Nonetheless, second-degree price discrimination is commonly studied through a more specific utility. We now add some assumptions to be consistent with this literature.

First, since a good's quality has no intrinsic measure, it is usually supposed that

$u(z) := z$. Second, income $Y$ is not included in the utility function when $x_1 = 1$. This is without loss of generality, since income does not affect the consumer's decision (it acts like a constant, and so plays the same role as a monotone transformation). On the contrary, income is replaced by a parameter $U_0$ in case $x_1 = 0$, which represents the consumer's reservation utility: the utility that the consumer gets if she does not consume good 1. This utility corresponds to the case where good 1 is not consumed, and all income is spent on good 2. Introducing a parameter $U_0$ makes it clear that this is the consumer's outside option.

Incorporating these aspects, we directly consider the following utility function:

$$U(x_1; z, \theta) := \begin{cases} \theta z - p & \text{if } x = 1 \\ U_0 & \text{if } x = 0 \end{cases},$$

where we have dispensed with the subscript of good 1's, since it is the only good under analysis. From now on, we also normalize $U_0 := 0$.

As for aggregate demand, we suppose that there is a population of agents with the same utility function, but with different values of $\theta$. Specifically, we consider that $\theta \in \{\underline{\theta}, \overline{\theta}\}$ where $\overline{\theta} > \underline{\theta} > 0$. Moreover, there is a proportion $\alpha$ of consumers with $\overline{\theta}$, and a proportion $(1-\alpha)$ with $\underline{\theta}$. We refer to them as ***consumer types***. We refer in particular to consumers with $\overline{\theta}$ as high-valuation types, and those with $\underline{\theta}$ as low-valuation types. Occasionally, we simply refer to them as ***high and low types***, respectively.

### 12.3.2 The Firm

To characterize the supply side, we suppose there is only one firm in the industry. This firm chooses the price and quality of each version of the good. In particular, since there are two possible types of consumers, it is enough to consider two versions of the good. These versions are defined as pairs of quality-prices, which we denote by $(\underline{z}, \underline{p})$ and $(\overline{z}, \overline{p})$. The notation distinguishes between varieties according to the type of consumer that the variety is designed for (the first one for low types, and the second one for high types).

**Remark** *The firm only decides the quality and price of each version. However, firms also have to determine the quantities supplied of each, which in this particular model are either zero or one. For each version, we subsume this decision by including the possibility of* $p = \infty$ *(or some price really high), so that the consumers would not buy the variety.*

Producing a good with some quality $z$ entails costs $c(z)$, where $c' > 0$, $c'' > 0$, $c(0) = 0$, and $\lim_{z \to \infty} c'(z) = \infty$. The firm's profit from selling one unit of its product with quality $z$ and price $p$ is $\Pi(p, z) := p - c(z)$ per customer.

## 12.4 Second-Degree Price Discrimination

Once we have established the model setup, we delve into how a firm chooses the price and quality of each version. In particular, we analyze its decision assuming that the firm knows the distribution of preferences in the industry, but not the preferences of each specific consumer. Alternatively, we can suppose that the firm knows the preferences of each consumer, but it cannot prevent them from buying a specific variety. We refer to the solution of this scenario as ***the second-best solution***.

We derive conclusions regarding this solution by comparing it with a benchmark scenario. Comparing solutions between two scenarios can be done through a comparative-statics analysis. This compares an initial situation with another where a parameter has a different value. Another way, which we follow in these notes, is by comparing solutions in different contexts. For instance, this is what we do when we compare a market solution against the solution of a planner that maximizes welfare.

Our benchmark for the second-best solution will be given by the *full-information* case. In this scenario, the firm knows the preferences of each consumer, and can force each to choose a specific variety of the good. We refer to the solution of this problem as ***the first-best solution***.

### 12.4.1 The First-Best Solution

We begin by deriving the optimization problem under full information. In this scenario, the firm knows each consumer's type and offers only one specific variety to each consumer. In other terms, conditional on buying the good, it is the firm, and not the consumer, who decides the variety to be consumed. Nonetheless, we suppose that the firm cannot force a consumer to buy the good, so that she could spend her money on other goods.

The optimization problem is:

$$\max_{\left\{(\overline{z},\overline{p}),(\underline{z},\underline{p})\right\}} \Pi = \alpha\left[\overline{p} - c\left(\overline{z}\right)\right] + (1-\alpha)\left[\underline{p} - c\left(\underline{z}\right)\right]$$

$$\text{subject to} \begin{cases} \overline{\theta}\overline{z} - \overline{p} \geq 0 \\ \underline{\theta}\underline{z} - \underline{p} \geq 0 \end{cases}. \qquad \text{(PC)}$$

The inequalities in PC are known as **participation constraints**. They capture that the transaction is voluntary, in the sense that a consumer decides whether to buy or not the product. Conditional on buying the product, the consumer can only buy the version that the firm offers to her. PC implies that the consumer buys one unit of the good only if consuming that unit provides more utility than her outside option. Keep in mind that the reservation utility represents the utility derived from the outside option, which means that the income is spent on goods from other industries.

The optimization problem is such that it is never optimal to have a slack constraint (i.e. having one of the PC constraints holding with strict inequality). This occurs because, as long as consumers buy the good, a firm always finds it optimal to either increase $p$ or reduce $z$ until each constraint holds with equality.

Since the equations in PC hold with equality, the firm is extracting any potential gain the consumer can have.[3] Thus, each type of consumer ends up with zero utility (or, more generally, its reservation utility). This occurs since PC makes the price of each

[3] In formal terms, we would say that the consumer surplus is zero.

version be equal to the consumer's valuation of the product.[4] Formally,

$$\begin{aligned} \overline{p} &= \overline{\theta}\overline{z}, \\ \underline{p} &= \underline{\theta}\underline{z}. \end{aligned} \tag{PRICE-1}$$

Plugging PRICE-1 into the objective function, the optimization problem reduces to an unconstrained optimization with two variables:

$$\max_{\{\overline{z},\underline{z}\}} \Pi = \alpha \left[\overline{\theta}\overline{z} - c\left(\overline{z}\right)\right] + (1-\alpha)\left[\underline{\theta}\underline{z} - c\left(\underline{z}\right)\right].$$

with prices given by PRICE-1. The solution of this problem admits an interior solution, and so we can use the FOC to identify it:

$$\begin{aligned} \overline{\theta} &= c'\left(\overline{z}^*\right), \\ \underline{\theta} &= c'\left(\underline{z}^*\right). \end{aligned} \tag{FB}$$

From this, we obtain the intuitive result that $\overline{z}^* > \underline{z}^*$, which in turn implies that $\overline{p}^* > \underline{p}^*$. Expressed in words, it states that the variety consumed by the high-valuation consumer under full information has a greater quality and price, relative to the variety consumed by low-valuation consumers.

Since $\overline{\theta} > \underline{\theta}$, by the FOCs we have that $c'\left(\overline{z}^*\right) > c'\left(\underline{z}^*\right)$. Since $c'$ is strictly convex, then that inequality can hold if $\overline{z}^* > \underline{z}^*$.
Regarding prices, by the PC constraints, $\overline{p}^* = \overline{\theta}\overline{z}^*$ and $\underline{p}^* = \underline{\theta}\underline{z}^*$. Since $\overline{\theta} > \underline{\theta}$ and as we have shown that $\overline{z}^* > \underline{z}^*$, then $\overline{p}^* > \underline{p}^*$.

We can also obtain the profit per consumer that the firm gets from each type. The conclusion in this case is that profits are higher if the firm sells to a high type.

By using the FOCs, the optimal profit the firm gets from a type $\theta$ consumer is $\pi^*(z) := c'(z)\,z - c(z)$
Differentiate the expression so that $\frac{d\pi^*(z)}{dz} = c''(z)\,z + c'(z) - c'(z) = c''(z)\,z > 0$. Therefore, the firm gets greater profits per capita from high-valuation consumers.
Or by the Envelope Theorem, $\pi(z;\theta) := \theta z - c(z)$ and so $\frac{\partial \pi(z;\theta)}{\partial \theta} = z$ so that evaluating that derivative at the optimal value we obtain $\frac{\partial \pi^*(z;\theta)}{\partial \theta} = z^* > 0$.
Want to show that $\overline{\theta}\overline{z} - c(\overline{z}) > \underline{\theta}\underline{z} - c(\underline{z})$. For this, we use that the optimization problem is the sum of two different optimization problems. Thus, we know that
$\overline{\theta}\overline{z} - c(\overline{z}) > \overline{\theta}z - c(z)$ for any $z \neq \overline{z}$ by definition of a maximum. Then, in particular for $z = \underline{z}$, we have that

[4] Notice that this refers to the *distribution* of gains between the firm and the consumer, without implying that it is an inefficient solution. In fact, the levels of quality chosen by the firm are the same that a planner would choose. Remember that efficiency is not necessarily related to distributional matters.

$\overline{\theta}\overline{z} - c(\overline{z}) > \overline{\theta}\underline{z} - c(\underline{z})$ and, in turn, $\overline{\theta}\underline{z} - c(\underline{z}) > \underline{\theta}\underline{z} - c(\underline{z})$. Hence, the two inequalities imply that $\overline{\theta}\overline{z} - c(\overline{z}) > \underline{\theta}\underline{z} - c(\underline{z})$.

### 12.4.2 Insights from the First-Best Solution

Before deriving the second-best solution, let's first analyze the first-best solution when the firm now cannot distinguish between consumer types. When this occurs, the firm cannot force a consumer to choose one specific variety. Consequently, the consumer has to decide whether to buy the good, and also which variety would she consume.

To grasp some intuition, consider that the firm keeps offering the varieties of the first-best solution $\left\{(\overline{z}^*, \overline{p}^*), (\underline{z}^*, \underline{p}^*)\right\}$. This design of varieties is actually not optimal for the firm, since each type of consumer would select the same variety. In particular, while low types would still consume the variety with $(\underline{z}^*, \underline{p}^*)$, high types would have incentives to deviate and choose the variety designed for the low type. Let's show this.

As for the low types, she would get zero utility if she buys the bundle with low quality. This follows by how the bundle was designed, ensuring that PC holds. Now suppose that she chooses the variety $(\overline{z}^*, \overline{p}^*)$, in which case her utility would be $U = \underline{\theta}\overline{z} - \overline{p}$. We know that $\overline{\theta}\overline{z} - \overline{p} = 0$ by PC, and so $\underline{\theta}\overline{z} - \overline{p} < 0$ since $\overline{\theta} > \underline{\theta}$. Given that this utility is lower than than the utility given by variety $(\underline{z}^*, \underline{p}^*)$, ***a low-valuation consumer would still choose the low-quality variety.***

Now, consider a high-valuation consumer. If she chooses the variety with high-quality, her utility would be zero due to PC. On the contrary, if she chooses the variety $(\underline{z}^*, \underline{p}^*)$, she would a get a level of utility $U = \overline{\theta}\underline{z} - \underline{p}$. But we know that $\underline{\theta}\underline{z} - \underline{p} = 0$ by PC, and then $\overline{\theta}\underline{z} - \underline{p} > 0$ since $\overline{\theta} > \underline{\theta}$. Therefore, ***high-valuation consumers would consume the low-quality good***. Intuitively, the price premium charged for a quality upgrade is too high, given the differences in quality between both goods. Since high types have incentives to mimic the behavior of low types, the firm needs to redesign the varieties offered. We summarize this outcome in the following result.

**Result 12.1** *Suppose that the firm does not know the type of each consumer, and offers to each the menus* $\left\{(\overline{z}^*, \overline{p}^*), (\underline{z}^*, \underline{p}^*)\right\}$ *of the first-best solution. Then,*

- *low-valuation consumers would choose the low-quality variety, and*
- *high-valuation consumers would choose the low-quality variety.*

*Therefore, high-valuation consumers have incentives to mimic the behavior of low-valuation consumers.*

To illustrate the result, suppose that Apple launches a new iPhone, and this comes in two versions: basic and deluxe. Both versions are identical, except that the deluxe version has one additional gigabyte of RAM memory. In this sense, although the quality of the deluxe version is greater, the difference is not quite pronounced. Instead, the prices charged by Apple for each version are astronomically different: the basic version is just 800 dollars, while the deluxe version costs 2,500 dollars. If this were Apple's pricing strategy, high-valuation consumers would consider that the greater quality of the deluxe version does not justify paying that exorbitant price. Thus, all types of consumers would end up buying the basic version of the iPhone.

We might suspect that the optimal strategy for Apple would require reoptimizing the levels of quality and price of each version. In other words, it seems plausible that Apple may get higher profits by reoptimizing, under the constraint that consumers can choose which variety to consume. We will show below that this is indeed the case.

Two types of solutions can arise, depending on the values of the parameters. Either it is optimal to design only one variety to be exclusively consumed by the high types, or to design varieties to serve both consumer types but making each self-select. In the latter case, we say that the optimal strategy is establishing a ***screening procedure***: offer different varieties, such that high-valuation consumers choose the high-quality variety and low-valuation consumers choose the low-quality one.

Mathematically, relative to the full information case, the optimization requires adding **incentive-compatible constraints**. They ensure that no agent prefers the variety that was not designed for her, by asking that the utility of the variety designed for her type is greater than the utility of buying the other variety. For example, the incentive-compatible constraint for the high type is $\overline{\theta}\overline{z} - \overline{p} \geq \overline{\theta}\underline{z} - \underline{p}$, so that the utility of choosing the variety $(\overline{z}, \overline{p})$ is greater than consuming the variety $(\underline{z}, \underline{p})$.

### 12.4.3 The Optimization Problem for the Second-Best Solution

The firm's optimization problem is now:

$$\max_{\{(\overline{z},\overline{p}),(\underline{z},\underline{p})\}} \Pi = \alpha \left[\overline{p} - c\left(\overline{z}\right)\right] + (1-\alpha)\left[\underline{p} - c\left(\underline{z}\right)\right]$$

subject to

$$\begin{cases} \overline{\theta}\overline{z} - \overline{p} \geq 0 \\ \underline{\theta}\underline{z} - \underline{p} \geq 0 \end{cases} \tag{PC}$$

$$\begin{cases} \overline{\theta}\overline{z} - \overline{p} \geq \overline{\theta}\underline{z} - \underline{p} \\ \underline{\theta}\underline{z} - \underline{p} \geq \underline{\theta}\overline{z} - \overline{p} \end{cases}. \tag{IC}$$

We will not use the Kuhn-Tucker technique to solve the optimization problem. Instead, we use an alternative approach, commonly employed in the textbooks of this topic. The procedure follows three steps. First, we show that some of the constraints are either redundant or hold with equality. After this, we optimize the objective function incorporating the constraints that hold with equality, but ignoring those holding with a strict inequality. Finally, we obtain the solution, and check that the constraints with strict inequality are indeed satisfied.

**Remark** *To keep notation simple, we refer to the constraint of each type by using a bar above or below. Thus, for example, the constraints for the high type are* $(\overline{PC})$ *and* $(\overline{IC})$*, which are given by the first inequalities in PC and IC.*
*For the derivation of the properties we present next, it is assumed that any previous property already proved holds.*

**Property 1.** The $(\overline{PC})$ does not bind. Moreover, if we check that $(\overline{IC})$ and $(\underline{PC})$ hold, then $(\overline{PC})$ holds automatically.

(OPTIONAL) It is enough to show that when $\left(\overline{IC}\right)$ and $(\underline{PC})$ hold, then $\left(\overline{PC}\right)$ cannot bind. This is because $\left(\overline{IC}\right)$ and $(\underline{PC})$ have to be satisfied at the solution.

Formally, we want to show that since $\overline{\theta}\overline{z}-\overline{p}\geq\overline{\theta}\underline{z}-\underline{p}$ and $\underline{\theta}\underline{z}-\underline{p}\geq 0$ then $\overline{\theta}\overline{z}-\overline{p}\geq 0$.

By $\left(\overline{IC}\right)$, we know that $\overline{\theta}\overline{z}-\overline{p}\geq\overline{\theta}\underline{z}-\underline{p}$. Also, since $\overline{\theta}>\underline{\theta}$, then the LHS of the inequality is such that $\overline{\theta}\underline{z}-\underline{p}>\underline{\theta}\underline{z}-\underline{p}$. But, given that $(\underline{PC})$ has to hold, the RHS of the last inequality is such $\underline{\theta}\underline{z}-\underline{p}\geq 0$. So we can conclude that $\overline{\theta}\overline{z}-\overline{p}\geq\overline{\theta}\underline{z}-\underline{p}>0$ so that $\overline{\theta}\overline{z}-\overline{p}>0$ which implies that $\left(\overline{PC}\right)$ cannot hold with equality.

**Property 2.** The $\left(\overline{IC}\right)$ binds.

(OPTIONAL) I provide a proof by contradiction. We are going to show that if $\left(\overline{IC}\right)$ holds as a strict inequality, then the firm can always find a better solution (that is, a solution that gives more profits and at the same time satisfies all the other constraints). Thus a nonbinding $\left(\overline{IC}\right)$ is inconsistent with a maximization program.

If $\left(\overline{IC}\right)$ does not bind, it means that $\overline{\theta}\overline{z}-\overline{p}>\overline{\theta}\underline{z}-\underline{p}$. I will show that the firm could charge a higher price to the high-valuation consumers and this would not violate any of the constraints. This implies that $\overline{p}$ cannot be the solution to the problem since it does not maximize profits.

Thus, consider that instead of charging $\overline{p}$ now it considers charging $\overline{p}+\delta$ where $\delta>0$ is small enough so that $\left(\overline{IC}\right)$ still holds as an inequality. Thus $\overline{\theta}\overline{z}-(\overline{p}+\delta)>\overline{\theta}\underline{z}-\underline{p}$. Since $\left(\overline{PC}\right)$ does not bind, we can also choose $\delta$ such that $\left(\overline{PC}\right)$ holds.

Moreover, we will show that $(\underline{IC})$ is not violated when $\overline{p}+\delta$ is set. With a price $\overline{p}$, the constraint was holding so that $\underline{\theta}\underline{z}-\underline{p}\geq\underline{\theta}\overline{z}-\overline{p}$ and so the LHS is such that $\underline{\theta}\overline{z}-\overline{p}>\underline{\theta}\overline{z}-(\overline{p}+\delta)$, implying that $\underline{\theta}\underline{z}-\underline{p}>\underline{\theta}\overline{z}-(\overline{p}+\delta)$. The intuition is that if the low-valuation consumer was not choosing the high-quality good with prices $\overline{p}$, she will have even less incentives to do so if now the price is $\overline{p}+\delta$.

Finally, $(\underline{PC})$ is independent of the price charged for the high quality good, so it will hold irrespective of the value of $\overline{p}$.

Therefore, the firm would get more profits with $\overline{p}+\delta$ and this is feasible. Thus, a $\overline{p}$ that makes $\left(\overline{IC}\right)$ hold with strict inequality cannot be part of a solution.

**Property 3.** The $(\underline{IC})$ holds iff $\overline{z}\geq\underline{z}$.

(OPTIONAL) We show that $(\underline{IC})$, which is $\underline{\theta}\underline{z}-\underline{p}\geq\underline{\theta}\overline{z}-\overline{p}$, is equivalent to $\overline{q}\geq\underline{q}$. This follows because $\left(\overline{IC}\right)$ binds and so $\overline{\theta}\overline{z}-\overline{p}=\overline{\theta}\underline{z}-\underline{p}$.

$\underline{\theta}\underline{z}-\underline{p}\geq\underline{\theta}\overline{z}-\overline{p}\Leftrightarrow\underline{\theta}\underline{z}-\underline{p}-\left(\overline{\theta}\underline{z}-\underline{p}\right)\geq\underline{\theta}\overline{z}-\overline{p}-\left(\overline{\theta}\overline{z}-\overline{p}\right)$ because both terms in brackets are equal. Thus

$\Leftrightarrow\underline{\theta}\underline{z}-\overline{\theta}\underline{z}\geq\underline{\theta}\overline{z}-\overline{\theta}\overline{z}$

$\Leftrightarrow\left(\underline{\theta}-\overline{\theta}\right)(\underline{z}-\overline{z})\geq 0$

and, since $\left(\underline{\theta}-\overline{\theta}\right)<0$,

$\Leftrightarrow\overline{z}\geq\underline{z}$.

**Property 4.** The $(\underline{PC})$ binds.

(OPTIONAL) I also provide for this case a proof by contradiction. Suppose that $(\underline{PC})$ does not bind so that $\underline{\theta}\underline{z}-\underline{p}>0$. We will show that this is not consistent with a solution. Specifically, we will show it is always possible to choose a $\underline{p}+\delta$ that is feasible and constitutes a better solution relative to choosing $\underline{p}$.

First, since $\underline{\theta}\underline{z}-\underline{p}>0$, we can always define a $\delta>0$ small enough so that the $(\underline{PC})$ would still hold. Also, $\underline{p}$ does not affect $\left(\overline{PC}\right)$. Besides, since $(\underline{IC})$ holds iff $\overline{z}\geq\underline{z}$, then $(\underline{IC})$ is independent of the price charged $\underline{p}$. Finally, $\left(\overline{IC}\right)$

is such that $\overline{\theta}\overline{z} - \overline{p} \geq \overline{\theta}\underline{z} - \underline{p}$ which implies that $\overline{\theta}\overline{z} - \overline{p} \geq \overline{\theta}\underline{z} - \left(\underline{p} + \delta\right)$ because $\delta > 0$.
But by charging $\underline{p} + \delta$ instead of $\underline{p}$, the firm would obtain greater profits. Thus, $(\underline{PC})$ cannot hold as an strict inequality at the solution.

Incorporating these properties, the optimization problem becomes

$$\max_{\{(\overline{z},\overline{p}),(\underline{z},\underline{p})\}} \Pi = \alpha\left[\overline{p} - c\left(\overline{z}\right)\right] + (1-\alpha)\left[\underline{p} - c\left(\underline{z}\right)\right]$$

$$\text{subject to } \begin{cases} (\underline{PC}): \ \underline{\theta}\underline{z} - \underline{p} = 0 \\ \left(\overline{IC}\right): \ \overline{\theta}\overline{z} - \overline{p} = \overline{\theta}\underline{z} - \underline{p} \\ \overline{z} \geq \underline{z} \end{cases}.$$

We have not added the constraint $\left(\overline{PC}\right)$, since it is automatically satisfied by Property 1 once we check that $\left(\overline{IC}\right)$ and $(\underline{PC})$ hold.

Before we solve the optimization problem, we state a conclusion that follows from a simple inspection.

**Result 12.2** *Low-valuation consumers obtain the same utility as in the first-best solution. On the contrary, high-valuation consumers are always better off relative to the first-best solution.*

Low-valuation consumers obtain the same utility since $(\underline{PC})$ holds with equality in both problems. Regarding high-valuation consumers, notice that $\underline{p} = \underline{\theta}\underline{z}$ by $(\underline{PC})$. Plugging this into $\left(\overline{IC}\right)$ we obtain that:

$$\overline{\theta}\overline{z} - \overline{p} = \left(\overline{\theta} - \underline{\theta}\right)\underline{z} > 0$$

This implies that the firm cannot extract the whole surplus from this consumer type. Rather, the firm has to let this type obtain some additional utility $\left(\overline{\theta} - \underline{\theta}\right)\underline{z}$, so that she she self-selects. In the literature's terminology, high-valuation consumers have private information, and the firm can only induce them to elicit their type by providing some ***information rent***. Only by letting them enjoy some extra utility is that the consumer would stop mimicking the behavior of low-valuation consumers.

### 12.4.4 The Second-Best Solution

The optimization problem can be further simplified by plugging the two constraints that hold with equality into the objective function. Thus

$$\max_{\{\overline{z},\underline{z}\}} \Pi = \alpha\left[\overline{\theta}\overline{z} - c\left(\overline{z}\right) - \left(\overline{\theta} - \underline{\theta}\right)\underline{z}\right] + (1-\alpha)\left[\underline{\theta}\underline{z} - c\left(\underline{z}\right)\right]$$

$$\text{subject to } \overline{z} \geq \underline{z}.$$

We use that $\underline{p} = \underline{\theta}\underline{z}$ and so $\left(\overline{IC}\right)$ can be reexpressed as $\overline{\theta}\overline{z} - \overline{p} = \overline{\theta}\underline{z} - \underline{\theta}\underline{z}$ so that $\overline{p} = \overline{\theta}\overline{z} - \left(\overline{\theta} - \underline{\theta}\right)\underline{z}$. Plugging both into the profits function, $\Pi = \alpha\left[\left(\overline{\theta}\overline{z} - \left(\overline{\theta} - \underline{\theta}\right)\underline{z}\right) - c\left(\overline{z}\right)\right] + (1-\alpha)\left[\underline{\theta}\underline{z} - c\left(\underline{z}\right)\right]$ which gives the result.

The problem has two possible types of solutions, depending on the value of $\underline{\theta}$ relative to $\alpha\overline{\theta}$. The first solution holds when $\underline{\theta} > \alpha\overline{\theta}$. It provides an interior solution where both types of consumers are served. The second solution occurs when $\underline{\theta} < \alpha\overline{\theta}$ and entails that only the high-valuation consumers are served.

We show that when $\underline{\theta} < \alpha\overline{\theta}$, then there is a corner solution for the low quality good., take $\frac{\partial \Pi}{\partial \underline{z}} = -\alpha\left(\overline{\theta} - \underline{\theta}\right) + \underline{\theta} - \alpha\underline{\theta} - (1-\alpha)\,c'\left(\underline{z}\right)$ which can be reexpressed as
$\frac{\partial \Pi}{\partial \underline{z}} = \underline{\theta} - \alpha\overline{\theta} - (1-\alpha)\,c'\left(\underline{z}\right)$. By assumption, $c'\left(\underline{z}\right) = 0$ iff $\underline{z} = 0$, and for any $\underline{z} > 0$ we have that $c'\left(\underline{z}\right) > 0$. Therefore, if $\underline{\theta} < \alpha\overline{\theta}$ then $\frac{\partial \Pi}{\partial \underline{z}} < 0$ for any $\underline{z}$ and so the firm has incentives to set the lowest value of $\underline{z}$ possible which is zero.

Since the first case has an interior solution, we can characterize it by the FOCs:

$$\overline{\theta} = c'\left(\overline{z}\right),$$

$$\underline{\theta} - c'\left(\underline{z}\right) = \frac{\alpha}{1-\alpha}\left(\overline{\theta} - \underline{\theta}\right). \tag{SDPD-HL}$$

On the contrary, the solution where only the high-valuation consumers are served is:

$$\overline{\theta} = c'\left(\overline{z}\right), \tag{SDPD-H}$$

$$\underline{z} = 0,$$

where we suppose that $\underline{p} > 0$, so that low-valuation consumers prefer their outside option rather than to consume the good.[5]

[5] Any combination $\left(\underline{z}, \underline{p}\right)$ such that the low type gets a utility lower than her reservation utility would work. The solution only has to ensure that low types do not consume any of the varieties available.

Comparing the solutions FB and SDPD-HL, we obtain the following conclusion:

**Result 12.3** *Suppose that firms cannot distinguish between consumer types and engage in second-degree price discrimination. Then, the variety for high-valuation consumers has the same quality as under full information, but the quality of the variety for low-valuation customers is reduced. This result is known in the literature as* **no distortion at the top**.

Suppose the solution given by equations SDPD-HL. $\overline{\theta} = c'(\overline{z})$ is the FOC for high-valuation consumers in both scenarios. Hence, $\overline{z}^* = \overline{z}^{**}$. Regarding low-valuation consumers, under complete information the FOC is $\underline{\theta} - c'(\underline{z}^*) = 0$ while under SPDP $\underline{\theta} - c'(\underline{z}^{**}) = \frac{\alpha}{1-\alpha}\left(\overline{\theta} - \underline{\theta}\right)$. This implies that $\underline{\theta} - c'(\underline{z}^{**}) > \underline{\theta} - c'(\underline{z}^*)$ and so $c'(\underline{z}^*) > c'(\underline{z}^{**})$. Given that $c$ is strictly convex, then $c'' > 0$ and so the first derivative is increasing. This implies that $c'(\underline{z}^*) > c'(\underline{z}^{**}) \Leftrightarrow \underline{z}^* > \underline{z}^{**}$.

Now, suppose the solutions given by equations SDPD-H. Then, $\overline{\theta} = c'(\overline{z})$ is the FOC for high-valuation consumers in both scenarios, so that $\overline{z}^* = \overline{z}^{**}$. In the case of low-valuation consumers, the FOC $\underline{\theta} - c'(\underline{z}^*) = 0$ under complete information determines that $\underline{z}^* > 0$ while if the firm does not serve the consumer as in SDPD-H then $\underline{z}^{**} = 0$. Thus, the firm decreases the low quality of the good such that no consumer wants to get that version of the good.

## 12.5 An Application: Damaged Goods

We provide several examples of a strategy commonly pursued by firms: the introduction of damaged goods. This means that a company launches some lower quality versions of the original good, but without these varieties having lower costs. Usually, it occurs when all the versions of a good are simultaneously produced, but the company subsequently disables some of the good's features.

In the 90s, Sony created the MiniDisc, which was a smaller version of the compact disc. These discs were not able to store any other type of information that was not music. Moreover, its format had a storage capacity measured by minutes of music. Sony offered two versions of blank MiniDiscs: one of 60 minutes and another of 74 minutes. Both disc versions were produced identically and were able to potentially store 74 minutes of music. However, Sony was deliberately adding some line in the firmware that prevented recording beyond 60 minutes in one of the versions. Thus, the 60-minutes disc was identical to the 74 minutes version, but disabling part of the storage space. [6]

A similar strategy was pursued by IBM with its LaserPrinter E. One of its versions was identical to the regular model, but with a firmware slowing down the printing speed. Specifically, the code was modified to introduce waiting times between pages, resulting in about half the speed of the regular printer.

Another example involves the statistical software Stata. The company exploits that it is not an open-source product, so that the software's code cannot be modified by the customer. Its cheapest versions are actually damaged versions of the full version: they add restrictions in the number of core processors that the computer are allowed to use, and has restrictions in the number of variables that it is allowed to handle.

[6] One of the reasons for Sony to do this was to engage in a second-degree price discrimination. Another motive could be the existence of economies of scale when each version is produced using the same technology. Suppose the case we studied in the course, where there exists a fixed cost but, conditional on producing positive quantities, marginal costs are constant. Then, Sony could be able produce both varieties in the same process, and thus reduce average costs by exploiting a greater scale of production.

# Lecture Note 13

# Game Theory

## 13.1 Introduction

In this lecture note, we start the study of Noncooperative Game Theory. In particular, we prepare the ground for the analysis of *simultaneous-move games*. They represent games where each player only moves once and without knowing what other players have chosen. Formally, we will see that simultaneous-move games refer to static games with imperfect information. In subsequent lecture notes, we'll focus on our ultimate goal: to model games and derive a prediction of how they will be played.

## 13.2 Describing a Game

What is a game? Although we may be tempted to associate it with ludic interactions, its definition is in fact more general. It refers to any situation where agents are aware of the externalities that they impose on each other. In other terms, each player obtains a payoff that depends on her own actions, but also on what others choose. Once agents internalize this, they behave strategically. In this sense, strategic behavior is a consequence, rather than a defining element of a game.

There are four elements that describe a game:

[1] The players

[2] The rules

[3] The outcomes

[4] The payoffs

The rules of the game comprise any information describing the interaction between agents. It includes elements such as the choices available, the information they have at the moment of making a move, the timing of moves, etc.

Te literature additionally distinguishes between outcomes and payoffs. The latter represents an assessment of the outcomes through a utility function. Basically, the same outcome might entail different payoffs, depending on the player's preferences. However, this distinction is not particularly relevant for our purposes—it only affects games involving uncertainty, which are not part of this course.

The usual approach to account for uncertainty in games is to allow for a fictitious player called Nature. This "player" assigns some probability distribution to each possible sub-tree that could be played. A move by Nature gives rise to the so-called games of incomplete information.

Next, we establish some formal terminology regarding games. We suppose there is a number of $I$ players. Let player $i$'s space of strategies be $S_i$. Given a strategy $s_i \in S_i$ for each player $i$, a profile of strategies is a vector $(s_i)_{i=1}^{I}$, with the space of strategy profiles being $S := \times_{i=1}^{I} S_i$.

It is common to use the notation $s_{-i} := (s_i)_{i \neq i}$, which refers to the vector strategies chosen by each player that is not $i$. This allows us to denote a profile of strategies by $(s_i, s_{-i})$ and the space of strategy profiles by $S_i \times S_{-i}$. The notation comes in handy when we focus on player $i$'s choices.

Each profile of strategies determines an outcome for player $i$, assessed by a utility function $u_i$. Given a profile $(s_i, s_{-i})$, the payoffs are $u_i(s_i, s_{-i})$.

**Remark** *In this lecture note, I will not distinguish between* actions *and* strategies*. Although they are conceptually different, these concepts can be considered equivalent in games of simultaneous moves.*

To illustrate the terminology introduced to represent a game, we consider the Prisoner's Dilemma as an example. Since we are not interested in the game itself, we will just characterize its properties, rather than the situation itself.

**Remark** *In general, we will denote players with numbers in brackets. Thus, for instance,* (1) *refers to player 1. In the case of two players, I will use "she" for player* (1) *and "he" for player* (2).

There are two players, (1) and (2), which have to make a decision without knowing what the other player has chosen. The possible actions are “to cooperate” (denoted $C$) or “not to cooperate” (denoted $NC$).

### 13.2.1 Extensive Form Representation

There are two types of representations of games. The first one is the **extensive form**. This is illustrated in Figure 13.1, which is the tree representation of the Prisoner's Dilemma.

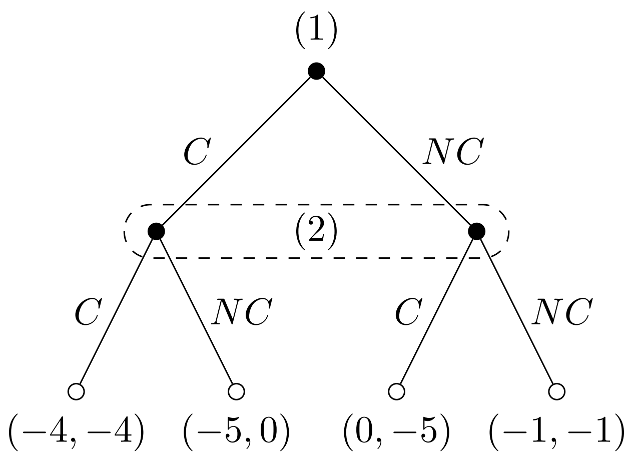



**Figure 13.1**

Each solid dot is a **decision node**. It represents an instance of the game where a player has to make a move. Nodes where the game ends are called terminal nodes. They are represented by non-filled dots and indicate the payoff that each player gets. Recall that a payoff is an outcome evaluated through a player's utility function.

> **Remark** *We use the convention that each payoff $(a, b)$ is such that $a$ represents (1)'s payoff, while $b$ is (2)'s payoff. This also applies to the normal form representations that we present below.*

In the game under analysis, each player has to make a move, without knowing what the other player has decided. In Figure 13.1, this can be clearly observed for (1): the game starts when she makes a move, and so she does not know what (2) will choose. The fact that neither (2) knows what (1) chose is indicated by the dashed oval. Formally, the player is unable to distinguish between the nodes comprised by the dashed oval, because he does not observe (1)'s choice.

More technically, we call **information sets** to the possible set of nodes where the player makes a choice. For instance, (1) has an information set where only one node is possible. On the contrary, (2) has an information set with two nodes, since (1) could have chosen either $C$ or $NC$. This gives rise to the definition of perfect and imperfect information in a game.

**Definition 13.1:** *A game of* ***perfect information*** *is such that all the information sets have only one node (i.e., the information sets are singletons). A game of* ***imperfect information*** *has at least one information set with more than one decision node.*

Based on this, we define a *simultaneous-move game* as one where all players make their moves under imperfect information. This captures that players do not know what others have chosen when they decide.

### 13.2.2 Normal Form Representation

Another game representation is its normal form. This is the one we will use for simultaneous-move games, and is visually illustrated in Figure 13.2 for the Prisoner's Dilemma.

| | | (2) $C$ | (2) $NC$ |
|---|---|---|---|
| (1) | $C$ | $-4, -4$ | $-5, 0$ |
| | $NC$ | $0, -5$ | $-1, -1$ |

**Figure 13.2**

The normal form provides information about each player's strategies and their corresponding payoffs. It abstracts from other pieces of information, such as the timing, knowledge of other players' moves, etc. These details are irrelevant in simultaneous games, but are crucial for more complex games. This is why normal form representations are an inaccurate description for games that are not simultaneous. However, only the mapping between choices and payoffs is relevant for simultaneous-move games, which are our focus.

## 13.3 Solutions Concept

We next endow players with some behavioral assumptions, with the goal of predicting how the game will be played. A set of assumption and its corresponding predictions are referred to as a **solution concept**.

We start by imposing some mild assumptions, and show that they are enough to get a unique prediction in simple games. For other games, the predictions under these assumptions are too broad, and in some cases we cannot even rule any possible solution. Due to this, we refine the solution concept by sequentially adding more assumptions. Following this procedure, we will end up with the so-called Nash equilibrium as a solution concept.

The examples we will use are not particularly relevant at this point. They are abstract games, and so we do not describe the type of situation captured. Their goal is just to show the role of the assumptions in each solution concept.

### 13.3.1 Rationality

Let's start by solving the game of Figure 13.2. We focus on player (1)'s decision, since the analysis for player (2) is analogous.

We will show that this game is quite simple, and we can predict a solution with a minimal assumption: each player maximizes her own utility. We will refer to this assumption as saying that each player is *rational*.

The payoff that (1) gets depends not only on what she chooses, but also on (2)'s choice. She does not know (2)'s decision, and notice (1) is not making any assumption regarding (2)'s behavior. Put it differently, we assume that **each player is rational, but we do not assume that each player considers her rival as rational**.

Under the assumption that (1) is rational, we can analyze what she chooses, depending on her conjectures about (2)'s choice. In this way, we analyze (1)'s choice that maximizes her utility, for each possible decision of (2).

Suppose that (2) chooses $C$. Then, (1) knows she would get $-4$ if her choice is $C$, and 0 if she chooses $NC$.. Therefore, (1) would prefer to choose $NC$ if she knew that (2) chooses $C$. Suppose now that (2) chooses $NC$. Then, (1) would $-5$ if her choice is $C$, and $-1$ if she chooses $NC$. Thus, it would be optimal for (1) to choose $NC$ if (2) chooses $NC$.

The previous analysis has shown that, even when (1)'s payoff depends on what (2) chooses, it is always optimal for (1) to choose $NC$. That is, irrespective of what (2) chooses, the best she can do is to choose $NC$. Consequently, we can predict what (1) will play if we only assume that she is rational, even when (1)'s payoffs depend on what (2) does. A similar analysis could be made for (2)'s choice. The conclusion would be that $NC$ maximizes his utility, irrespective of what (1) chooses.

The main conclusion is that rationality allows us to obtain a prediction for this game in particular. **Rationality** in the context of game theory means that agents do not play strictly dominated strategies. Strategy $s_i''$ is **strictly dominated** for player $i$ if there exists a strategy $s_i'$ such that $s_i'$ provides a greater payoff, irrespective of what the rest of players choose. Formally, $s_i''$ is strictly dominated for player $i$ by $s_i$' when

$$u_i\left(s_i', s_{-i}\right) > u_i\left(s_i'', s_{-i}\right) \text{ for all } s_{-i} \in S_{-i}$$

**Remark** *Notice that a strictly dominated strategy $s_i''$ has to provide a* strictly *lower payoff. We do not consider that $s''$ is weakly dominated, i.e. that there exists some strategy providing the same utility for at least one rivals' strategy). If $s''$ is only weakly dominated, we cannot rule out that $s_i''$ will be played. In fact, eliminating weakly dominated strategies from a player's considerations can lead to some issues. For instance, the game prediction could end up depending on the order in which players eliminate strategies.*

Based on the concept of strictly dominated strategies, we define our first solution concept.

**Solution Concept 1** **Rationality of Each Player**. We suppose that each player is rational. Agents are rational when they do not play strictly dominated strategies.

The fact that agents do not play strictly dominated strategies only rules out strategies that will be played. In other terms, the solution concept is not based on what players will choose, but what they will not choose. Thus, we can only have a prediction for the game if the elimination of dominated strategies results in a unique strategy for each player.

In cases like the Prisoner's Dilemma, or any game with a similar structure to **??**, rationality is enough for having a unique prediction. This is so because each player $i$ has a strategy $s_i^*$ that strictly dominates any other strategy $s_i' \in S_i$ with $s_i' \neq s_i^*$. When this occurs, we say that $s_i^*$ is a **strictly dominant strategy** for player $i$. Notice that the existence of a strictly dominant strategy $s_i^*$ is equivalent to all the strategies $s_i' \neq s_i^*$ being strictly dominated.

In summary, the result in the Prisoner's Dilemma is quite robust because *we get a unique prediction by simply assuming rationality of players—no need to make any assumption on the players' conjectures about rivals*. This does not imply that what rivals do is irrelevant. The final outcome still depends on what others do. However, players can choose optimally without information on how rivals behave. This is why whether rivals are rational or not (or, actually, any other feature of them) is irrelevant for identifying a player's best action.

The downside of Solution Concept 1 is that we could end up with a plethora of possible solutions. Indeed, games where each player has a dominant strategy are the exception, rather than the rule. And it could be argued that Game Theory is actually irrelevant for these situations, since the approach would not differ from a simple utility-maximization procedure.

To demonstrate what occurs when not all players have a strictly dominant strategy, consider the game in Figure 13.3.

| | | (2) D | (2) E |
|---|---|---|---|
| (1) | A | $2,-2$ | $-2,2$ |
| | B | $-2,2$ | $2,-2$ |
| | C | $-3,6$ | $-4,4$ |

**Figure 13.3**

In this game, we can rule out that (1) plays $C$, because $C$ is strictly dominated by $A$ and $B$. However, this is the only choice we can eliminate for any player. Hence, the game prediction is that (1) will choose either $A$ or $B$, and (2) either $D$ or $E$.

### 13.3.2 Iterated Elimination of Strictly Dominated Strategies

To get sharper predictions of a game, we **refine the solution concept**. This means that we start from Solution Concept 1 and add assumptions. Thus, we will still assume the rationality of players. Consider the game in Figure 13.4

| | | (2) L | (2) M | (2) R |
|---|---|---|---|---|
| (1) | U | $1,0$ | $1,2$ | $0,1$ |
| | D | $0,3$ | $0,1$ | $2,0$ |

**Figure 13.4**

Let's see what we can eliminate by only considering the rationality of each player. Player (1) has no strictly dominated strategies. If (2) chose $L$ or $M$, the best move would be $U$, while $D$ is (1)'s best choice if (2) played $R$. On the contrary, $R$ is strictly dominated by $M$ for player (2). This is because if (1) plays $U$, then $M$ provides 2 as payoff, while $R$ just 1; instead, (1) plays $D$, then $M$ provides 1 and $R$ just 0.

Consequently, by only assuming rationality, we can predict that (1) plays either $U$ or $D$, and (2) plays either $L$ or $M$. The normal representation of the game incorporating this is as in Figure 13.5.

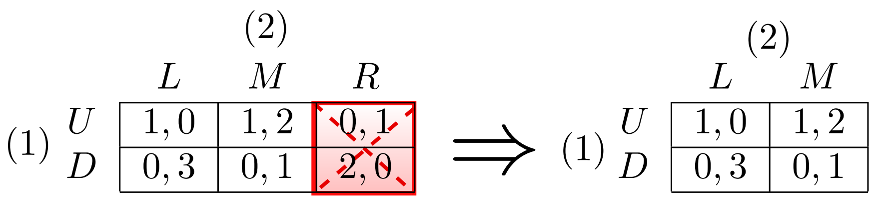



**Figure 13.5**

If our goal is to obtain sharper predictions for this game, it is necessary to make further assumptions. Starting from Figure 13.5, there are no strictly dominated strategies, meaning that a player's optimal choice depends on what others choose. Due to this, we need to assume what players conjecture regarding what rivals will choose. This entails that each player must assume the rivals' goals, how they process information, etc. In other words, a player needs to have conjectures about some matters related to how rivals decide.

Since we are already assuming rationality of each player, it is natural to extend this assumption as a player's conjecture about her rivals. In other words, let's consider that each player is rational and conjectures that their rivals are too.

To see how this affects the game's prediction, let's start considering that player (1) assumes (2) is rational. Now, (1) will analyze the game as (2) did before, knowing that (2) would not play $R$. Thus, (1) will decide by taking Figure 13.6 as the relevant game. Incorporating this, $D$ is strictly dominated by $U$ for player (1), implying that the game in its normal form would be:

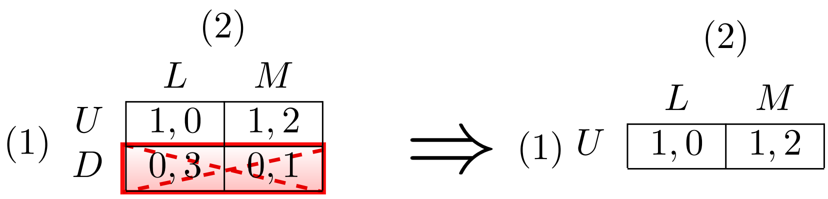



**Figure 13.6**

So, we could predict that (1) would choose $U$. What about player (2)? So far, we have only assumed that he is rational, and that (1) is rational and supposes that (2) is rational. But, we have not established that (2) makes any assumption regarding (1).

Suppose that (2) conjectures that (1) is rational and that (1) knows that he is rational. In that case, (2) considers that (1) does not play strictly dominated strategies. Thus, (2) will conclude that (1) expects to play the game in Figure **??**, and so (1) will play $U$. Once (2) arrives at that conclusion, his choice would be $M$, which gives him a payoff of 2 relative to playing $L$ and get 0. In summary, under the assumptions made, the game prediction is (1) playing $U$ and (2) choosing $M$.

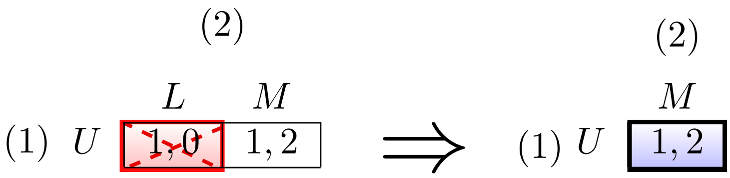



**Figure 13.7**

Let's translate the intuition of this example into a formal solution concept. After eliminating strategies strictly dominated in Figure **??**, each step of the game pushes further the assumption of the rival's rationality. A generalization needs to allow us to iteratively eliminate strictly dominated strategies as many times as we want. This can be accomplished by requiring that rationality and the structure of the game are common knowledge. A fact is common knowledge when the fact holds, players know that the fact holds, players know that players know the fact holds, and so on ad infinitum.

**Solution Concept 2** **Iterated Elimination of Strictly Dominated Strategies** (IES for short). When agents are rational, and this and the structure of the game are common knowledge, strictly dominated strategies can be iteratively removed.

**Remark** *One appealing feature of IES is that the order in which we eliminate strategies does not affect the set of surviving strategies.*
*Even when IES can narrow the prediction of games, it still might lead us to inaccurate predictions about a game. One example is the following game, where any action ends up being possible.*

| | | (2) L | C | R |
|---|---|---|---|---|
| (1) | T | 0, 4 | 4, 0 | 5, 3 |
| | M | 4, 0 | 0, 4 | 5, 3 |
| | B | 3, 5 | 3, 5 | 6, 6 |

**Figure 13.8**

### 13.3.3 Rationalizable Strategies

So far, the solution concepts used were based on strictly dominated strategies. Next, we derive a solution based on an alternative concept, referred to as best response. This will take us to the concept of rationalizable strategies, which rely on the existence of conjectures to justify playing a strategy.

#### 13.3.3.1 Some Preliminaries

Let's begin by adding some definitions. The set of **best response strategies** (BR) to a strategy $s'_{-i}$ is defined as:

$$BR_i\left(s'_{-i}\right) := \left\{s_i \in S_i : u_i\left(s_i, s'_{-i}\right) \geq u_i\left(s'_i, s'_{-i}\right) \text{ for all } s'_i \in S_i\right\}$$

This means that, if $i$ considers that her rivals will choose $s'_{-i}$, she will opt for a strategy $s_i \in BR_i\left(s_{-i}\right)$. This strategy provides the greatest utility *given* the conjecture $s'_{-i}$.

Quite related to this, it is the set of **never best responses** (NBR). This is the set of $i$'s strategies with no conjecture about what others do that justifies its use. Formally,

$$NBR_i := \left\{s_i \in S_i : \nexists s'_{-i} \text{ such that } u_i\left(s_i, s'_{-i}\right) \geq u_i\left(s'_i, s'_{-i}\right) \text{ for all } s'_i \in S_i\right\}.$$

The definition of $s_i \in NBR_i$ is equivalent to $s_i \notin BR_i\left(s'_{-i}\right)$ for any $s'_{-i} \in S_{-i}$. This leads us to an important definition: a player is **Bayes rational** if she does not play strategies that are NBR.

Notice that *a strictly dominated strategy is NBR, but a NBR strategy is not necessarily strictly dominated*. We illustrate this through the following example.

Example

Consider the game in Figure 13.9, where we have only specified the payoffs of player (1). The payoffs of (2) are irrelevant for making the point.

| | | (2) D | E |
|---|---|---|---|
| (1) | $A$ | $3, \cdot$ | $0, \cdot$ |
| | $B$ | $0, \cdot$ | $3, \cdot$ |
| | $C$ | $2, \cdot$ | $2, \cdot$ |

**Figure 13.9**

Figure 13.9 establishes that $BR_1(D) = A$ while $BR_1(E) = B$. In plain words, $A$ is (1)'s BR if she thinks that (2) plays $D$: while $A$ provides a utility of 3, $B$ gives a utility of 0 and $C$ a utility of 2. A similar reasoning shows that $B$ is (1)'s BR to (2) choosing $E$.

From this, we conclude that $C \in NBR_1$, since there is no conjecture about what (2) chooses that justifies (1) playing $C$. However, $C$ is not a strictly dominated strategy: $C$ provides a higher payoff than $B$ if (2) plays $D$, and a higher payoff than $A$ if (2) plays $E$.

#### 13.3.3.2 Determining the Set of Rationalizable Strategies

Based on the concept of NBR, we introduce a new solution concept: rationalizable strategies. Just like with IES, the idea is to iteratively eliminate strategies that are NBR.

**Solution Concept 3** **Rationalizable Strategies**. When agents are Bayes rational, and this and the structure of the game are common knowledge, NBR strategies can be iteratively removed.

**Remark** *Similar to IES, the set of rationalizable strategies does not depend on the order in which we eliminate the NBR strategies.*

Consider the game in Figure 13.10.

| (1) \ (2) | $b_1$ | $b_2$ | $b_3$ | $b_4$ |
|---|---|---|---|---|
| $a_1$ | $0,7$ | $2,5$ | $7,0$ | $0,1$ |
| $a_2$ | $5,2$ | $3,3$ | $3,2$ | $0,1$ |
| $a_3$ | $7,0$ | $2,5$ | $0,7$ | $0,1$ |
| $a_4$ | $0,2$ | $0,0$ | $0,0$ | $5,1$ |

**Figure 13.10**

You can check that this game has no strictly dominated strategies. Furthermore, all the strategies of (1) are rationalizable, which means there are no NBR strategies for (1). On the contrary, $b_4$ is a NBR for (2), since there is no conjecture that leads (2) to play $b_4$.

The set of BR for (1) is:

- $BR_1(b_1) = a_3$
- $BR_1(b_2) = a_2$
- $BR_1(b_3) = a_1$
- $BR_1(b_4) = a_4$

and so all the strategies are BR for some strategy of the rival player.
The set of BR for (2) is:

- $BR_2(a_1) = b_1$
- $BR_2(a_2) = b_2$
- $BR_2(a_3) = b_3$
- $BR_2(a_4) = b_1$

and so $b_4$ is NBR because there is no action of (1) such that it justifies that (2) chooses $b_4$.

Incorporating that $b_4$ is a NBR, the representation of the game is as in Figure 13.11.

(2)

| (1) | $b_1$ | $b_2$ | $b_3$ | $b_4$ |
|---|---|---|---|---|
| $a_1$ | $0,7$ | $2,5$ | $7,0$ | $0,1$ |
| $a_2$ | $5,2$ | $3,3$ | $3,2$ | $0,1$ |
| $a_3$ | $7,0$ | $2,5$ | $0,7$ | $0,1$ |
| $a_4$ | $0,2$ | $0,0$ | $0,0$ | $5,1$ |

$\Longrightarrow$

(2)

| (1) | $b_1$ | $b_2$ | $b_3$ |
|---|---|---|---|
| $a_1$ | $0,7$ | $2,5$ | $7,0$ |
| $a_2$ | $5,2$ | $3,3$ | $3,2$ |
| $a_3$ | $7,0$ | $2,5$ | $0,7$ |
| $a_4$ | $0,2$ | $0,0$ | $0,0$ |

**Figure 13.11**

Given the assumption of common knowledge, (1) knows that (2) will not play $b_4$. This allows us to eliminate $a_4$ from (1)'s consideration: $a_4$ was a BR only to the conjecture $b_4$, but this will not be played by (2). In other words, $a_4$ is not a BR to any strategy that survived the elimination of NBRs. Incorporating this result, the relevant game is Figure 13.12.

(2)

| (1) | $b_1$ | $b_2$ | $b_3$ |
|---|---|---|---|
| $a_1$ | $0,7$ | $2,5$ | $7,0$ |
| $a_2$ | $5,2$ | $3,3$ | $3,2$ |
| $a_3$ | $7,0$ | $2,5$ | $0,7$ |
| $a_4$ | $0,2$ | $0,0$ | $0,0$ |

$\Longrightarrow$

(2)

| (1) | $b_1$ | $b_2$ | $b_3$ |
|---|---|---|---|
| $a_1$ | $0,7$ | $2,5$ | $7,0$ |
| $a_2$ | $5,2$ | $3,3$ | $3,2$ |
| $a_3$ | $7,0$ | $2,5$ | $0,7$ |

**Figure 13.12**

Since now all the strategies are BRs to some conjecture, each player's set of rationalizable strategies is:

- $\{a_1, a_2, a_3\}$ for player (1)
- $\{b_1, b_2, b_3\}$ for player (2).

#### 13.3.3.3 Interpreting the Rationalizable Strategies

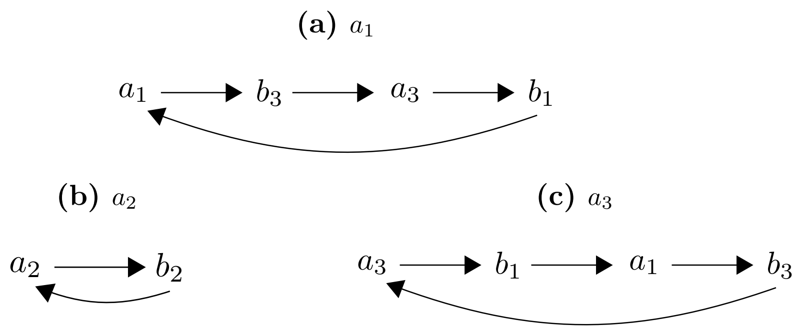
Rationalizable strategies are the set of strategies remaining after an iterative removal of NBR. This has two implications. First, a player can always "justify" her action, since she is giving a BR to some conjecture. However, notice there is another important property of rationalizable strategies: any player can "justify" her conjectures. In fact, we can keep this reasoning ad infinitum, in the sense that the conjectures can be justified by some rationalizable conjecture. To illustrate this, consider the same again as before, whose relevant strategies after the iterative deletion of NBR are presented in Figure 13.13.

| (1) \ (2) | $b_1$ | $b_2$ | $b_3$ |
|---|---|---|---|
| $a_1$ | $0,7$ | $2,5$ | $7,0$ |
| $a_2$ | $5,2$ | $3,3$ | $3,2$ |
| $a_3$ | $7,0$ | $2,5$ | $0,7$ |

**Figure 13.13**

Consider how (1) can justify her choice $a_1$. She plays $a_1$ because she conjectures that (2) will play $b_3$. But why does (1) believe that (2) will play $b_3$? Because (1) thinks that (2) believes that (1) will play $a_3$. And (1) could be thinking this, if she expects that (2) will play $b_1$, which in turn can be justified if (1) plays $a_1$. After this, a loop of justification is created. Graphically, the chain of justifications is as in Figure 13.14.

**Figure 13.14.** *Rationalizable Strategies- Chain of Justifications*




### 13.3.4 Nash Equilibrium

Consider Figure 13.14, which provides several examples of rationalizable strategies. Figure 13.14b reveals that the actions $a_2$ and $b_2$ have a distinctive feature. Player (1) can justify $a_2$ because she expects that (2) plays $b_2$. In turn, (2) would play $b_2$ because he expects that (1) plays $a_2$. This determines that the conjectures are "correct", in the sense that what players conjecture is what actually takes place. Put it differently, each strategy of the pair $(a_2, b_2)$ is a BR to what the rival is choosing. This leads to the most relevant solution concept we will use.

**Solution Concept 4** **Nash Equilibrium**. A strategy profile is a Nash equilibrium when each strategy is rationalizable and based on "correct" conjectures.

## 13.4 Exercises

[1] Suppose the following game:

| (1) \ (2) | $D$ | $E$ |
|---|---|---|
| $A$ | $(\alpha, \beta)$ | $(\gamma, 2)$ |
| $B$ | $(1, 1)$ | $(1, 0)$ |
| $C$ | $(3, 2)$ | $(0, 1)$ |

where all Greek letters are parameters. Determine the range of values of $\alpha, \beta$ and $\gamma$ such that:

(a) $A$ is a strictly dominated strategy.

(b) $C$ is a rationalizable strategy.

(c) $(A, D)$ is a NE.

(d) All the strategies are rationalizable.

[2] Suppose the following game:

| (1) \ (2) | $L$ | $M$ | $R$ |
|---|---|---|---|
| $U$ | $1, 0$ | $1, 2$ | $0, 1$ |
| $D$ | $0, 3$ | $0, 1$ | $2, 0$ |

Establish the game prediction, according to the following assumptions.

(a) Each player is Bayes rational

(b) Each player is Bayes rational, and (1) supposes that (2) is Bayes rational

(c) Each player is Bayes rational, (1) supposes that (2) is Bayes rational, and (2) knows all this (that is, (2) knows that (1) is Bayes rational and that (1) supposes that (2) is Bayes rational).

(d) Identify the rationalizable strategies and the NE in this game.

[3] We'll consider a situation that resembles the Tug of War in several respects. The aim is that you learn how to partition strategy profiles to find a game's Nash

equilibria. To be as intuitive as possible, let's consider a game that has me and all of you as players.

Suppose I sent you an email last Monday with an exercise, and offered you extra points to boost your grade. You can submit a solution to my mailbox by Friday, and there's no penalty if a student does not submit a solution. However, the extra points are only available for one student. Thus, I establish the following rule. I'll be checking my mailbox on Friday. If I find only one correct solution, I'll give the extra points to that person. However, if there's more than one student with a correct solution, I won't give any extra points to anyone—I'm worried about the possibility that students actually cheated and copied the solution.

The actions available to you are $S$ and $NS$, which stands for "to submit" and "not to submit". Suppose that you want to play fairly and won't discuss the solution with your classmates. Thus, you have to make a decision without knowing what the rest of the students have chosen.

Assume also that everyone knows how to solve the exercise, but it's long and requires putting a lot of effort, thus generating disutility. For this reason, the students will only submit a solution if they expect to get the extra points. Formally, let $E$ and $NE$ stands for "extra points" and "no extra points". The preferences of each student are such that $u(S; E) > u(NS; NE)$.

Assuming there are $n \geq 2$ students in the course, find all the NE of the game.

[4] Consider the Tug of War game, but where we dispense with the assumption of an equal number of players in each team. Specifically, suppose that team 1 has $m$ players, while team 2 has $n$ where $m > n$. Determine the NE for the cases we studied in class, according to:

(a) the baseline preferences

(b) The alternative preferences

*Hint*: Assuming one of these preferences determines there's no NE.

**Some Answers to the Exercises**:

**1)** a) either $\alpha < 1$ and $\gamma < 1$, or $\alpha < 3$ and $\gamma < 0$, b) $\alpha \leq 3$, c) $\alpha \geq 3$ and $\beta \geq 2$, d) $\beta \leq 2$ and either $\alpha = 3$ and $\gamma \leq 1$, or $\alpha \leq 3$ and $\gamma = 1$.

**2)** a) for (1) $\{U, D\}$ and for (2) $\{L, M\}$, b) for (1) $\{U\}$ and for (2) $\{L, M\}$, c) for (1) $\{U\}$ and for (2) $\{M\}$, d) they coincide with c)